

# Combinatorial Chemistry

Synthesis, Analysis, Screening

Edited by Günther Jung



# Combinatorial Chemistry

**Synthesis, Analysis, Screening**

Edited by  
Günther Jung

 **WILEY-VCH**

Weinheim · New York · Chichester · Toronto · Brisbane · Singapore

This Page Intentionally Left Blank

# **Combinatorial Chemistry**

Edited by Günther Jung

 **WILEY-VCH**



This Page Intentionally Left Blank

# Combinatorial Chemistry

**Synthesis, Analysis, Screening**

Edited by  
Günther Jung

 **WILEY-VCH**

Weinheim · New York · Chichester · Toronto · Brisbane · Singapore

Prof. Dr. Günther Jung  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

This book was carefully produced. Nevertheless, authors, editor and publisher do not warrant the information contained therein to be free of errors. Readers are advised to keep in mind that statement, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Reprint 2001

Library of Congress Card No: applied for

British Library Cataloguing-in-Publication Data: A catalogue record for this book is available from the British Library

Die Deutsche Bibliothek – CIP Einheitsaufnahme

Ein Titeldatensatz für diese Publikation ist bei der Deutschen Bibliothek erhältlich.

© WILEY-VCH Verlag GmbH, D-69469 Weinheim (Federal Republic of Germany), 1999

Printed on acid-free and chlorine-free paper.

All right reserved (including those of translation in other languages). No part of this book may be reproduced in any form – nor transmitted or translated into machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Composition: Datascan GmbH, D-64295 Darmstadt

Printing: Strauss Offsetdruck, D-69509 Mörlenbach

Bookbinding: Osswald & Co. D-67433 Neustadt (Weinstraße)

# Preface

Since I edited *Combinatorial Peptide and Nonpeptide Libraries* – the first comprehensive book on the subject – in 1996, progress in the field of combinatorial chemistry has been truly dramatic. However, before I began to write this preface, I looked back and found the early work by no means outdated – as evidenced by the number of complimentary letters received from satisfied readers. A number of excellent books on combinatorial chemistry have appeared since then, covering different aspects of the subject. However, despite this earlier ‘boom’, I was encouraged to edit a second volume, and recognized that soon there will be very little overlap of *Combinatorial Chemistry* with the contents of other such books.

The past five years have seen the development and introduction of new and improved approaches to the combinatorial sciences, their practical applications having resulted in an almost ‘explosive’ increase in the number of experimental publications and new journals devoted exclusively to these areas. Thus, a comprehensive book which emphasizes the state of the art of most of the methods and instrumentation of combinatorial chemistry may be of value to those who may consider entering this exciting field. It is possible also that experts in combinatorial organic chemistry may find it interesting to examine more closely exciting new fields such as catalyst research, chemosensor development, new analytical tools, and the creation of biodiversity.

As a logical consequence of the demand for combinatorial chemists in industry, teaching the subject and related key technologies should be a major topic at our universities. It is hoped that some of the chapters of this book are sufficiently convincing as to encourage future scientists not only to use combinatorial methods in their research, but also to offer training for their students or co-workers. Indeed, for some years, I have offered public courses in multiple methods in peptide chemistry, and presently my group teaches theoretical and practical methods in combinatorial organic chemistry.

As we approach the end of the 20th century, we observe two major trends in the drug discovery process: first, the production of large numbers of single compounds and complex libraries in small scale (microgram) quantities; and second, the production of preparative-scale amounts (mg) of single compounds with complete analytical characterization and defined purity criteria, and less complex mixtures. At present, the nanoscale synthesis and on-line screening of 100 000 compounds per day remains a dream for organic chemists – except perhaps for primitive compounds and rather simple biological targets. Nonetheless, such wishes may be realized for oligomers accessible by repetitive synthetic steps.

*Combinatorial Chemistry* begins with a general overview on recent advances in the subject, with expert surveys provided on selected solid-phase organic reactions and – for the first time – also on solution-phase combinatorial chemistry. These chapters are followed by a detailed survey of the broad research into multicomponent reactions.

The chapter on solid supports in the first book has now been supplemented by a complete listing of the various solid-phase anchors and linkers used in organic chemistry – information which may be of particular value to newcomers in the field.

As a consequence of the progress and success of diverse oligomer libraries, experts in combinatorial synthetic oligomers, glycopeptide and oligosaccharide libraries, RNA- and DNA-aptamers have each contributed their news in these fields, while novel experimental examples of the use of templates in combinatorial chemistry for the solid-phase synthesis of multiple core structure libraries provide some insight to praxis-relevant work.

A complete chapter is devoted to novel combinatorial approaches to highly selective chemoreceptors designed for parallel analysis by sensor devices. An assay-oriented chapter on peptide libraries in T-cell-mediated immune response illustrates how libraries of extreme complexity ( $10^{13}$ ) can lead directly to single molecules with activities that may be up to 100 000-fold higher than those of natural epitopes.

Results of combinatorial biosynthesis using gene clusters responsible for metabolites of potential pharmaceutical interest illustrate the state of the art in this field. Library design and diversity analysis are of increasing importance, and the present programs and approaches are discussed, together with expert views and opinions. In particular, the chapter on a miniaturized ultra-high throughput screening system highlights the immediate future in assay systems of this type.

Combinatorial approaches in materials sciences and solid-state catalysts are relatively new, and these are reviewed, together with expert contributions in the field.

Among the last five chapters, four are devoted to the important instrumentation used for analytical on- and off-bead controls. Whereas tagging concepts appear increasingly less interesting due to excessive time constraints and inherent problems introduced in the drug discovery process, a number of methods have proved suitable for routine single-bead analysis, including infra-red microscopy, separation techniques coupled to electrospray-mass spectrometry, and ion-cyclotron resonance-FT-mass spectrometry to determine the elemental composition of each component in a library. In addition, MAS-NMR has become a popular means of examining resin-bound intermediates during feasibility and optimization studies.

Finally, *Combinatorial Chemistry* closes with a critical discussion of the various attempts to automate combinatorial synthesis, with a presentation of a newly developed workstation used to perform fully automated parallel synthesis, including all steps through work-up to sample preparation for screening modules.

I would like to close this Preface by expressing my gratitude to all authors and co-authors who spent their precious time in contributing to this work. I also acknowledge the present and former co-workers of my group who contributed novel experimental results.

I am very grateful to Mrs. Ursula Sanzenbacher for her help during the editorial work, and to Dr. Gudrun Walter and Dr. Peter Göllitz from Wiley-VCH for their continuous interest on this edition.

During the past few years – and in parallel with combinatorial chemistry – communication via the Internet and the availability of published material from databases have emerged exponentially. However, I am confident that a broad readership may still – from time to time – enjoy reading a scientific book such as *Combinatorial Chemistry*.

Tübingen, September 1999

Günther Jung

This Page Intentionally Left Blank

# Contents

1.	Combinatorial Chemistry	
	<i>Dominika Tiebes</i>	
1.1	Introduction	1
1.2	Principles of Combinatorial Chemistry	2
1.3	Methods and Techniques of Combinatorial Synthesis	4
1.3.1	Synthetic Strategies Towards Combinatorial Libraries	4
1.3.1.1	Split-Pool Synthesis Towards Combinatorial Libraries	4
1.3.1.2	Parallel Synthesis Towards Combinatorial Libraries	7
1.3.1.3	Reagent Mixture Synthesis Towards Combinatorial Libraries	10
1.3.2	Synthetic Methodology for Organic Library Construction	12
1.3.2.1	Solid-Phase Organic Synthesis	12
1.3.2.2	Synthesis in Solution and Liquid-Phase Synthesis	13
1.4	Characterization of Combinatorial Libraries	14
1.4.1	Analytical Characterization	14
1.4.1.1	Analytical Characterization of Compound Mixtures	14
1.4.1.2	Analytical Characterization of Single Substances	15
1.4.2	Hit Identification in Combinatorial Libraries by High-Throughput Screening	15
1.4.2.1	Strategies for Libraries of Compound Mixtures	15
1.4.2.2	Strategies for Libraries of Separate Single Compounds	23
1.5	Automation and Data Processing	24
1.5.1	Synthesis Automation and Data Processing	24
1.5.2	Automated Purification	25
1.6	Library Design and Diversity Assessment	26
1.6.1	Diversity Assessment for Selection of Building Blocks or Compounds	27
1.6.2	Iterative Optimization Methods	28
1.7	Economic Aspects	28
1.8	Acknowledgements	30
1.9	References	30



## 2. Survey of Solid-Phase Organic Reactions

*Susan E. Booth, C. Marijke Dreef-Tromp, Petro H.H. Hermkens,*

*Jos A.P.A. de Man and H.C.J. Ottenheijm*

- 2.1 Introduction 35
- 2.2 Observed Trends 36
  - 2.2.1 The Synthetic Repertoire 36
    - 2.2.1.1 Robust, Reliable Solid-Phase Reactions 36
    - 2.2.1.2 Emerging Solid-Phase Reactions 36
  - 2.2.2 Linkers and Cleavage Step 37
  - 2.2.3 Reaction-Monitoring 37
  - 2.2.4 Highlights 37
- 2.3 Conclusions 37
- 2.4 Reaction Tables 38
  - 2.4.1 Substitution Nucleophilic and Electrophilic Type of Reaction:  
Amide Bond Formation and Related Reactions 38
    - 2.4.1.1 Sulphonamide 39
    - 2.4.1.2 (Thio)urea 39
    - 2.4.1.3 Carbonate 40
    - 2.4.1.4 Urethane 40
    - 2.4.1.5 Guanidine 41
    - 2.4.1.6 Imide 41
    - 2.4.1.7 Amide 42
    - 2.4.1.8 Lactam 42
  - 2.4.2 Type of Reaction: Aromatic Substitution;  
Electrophilic Carbon-Carbon Bond Formation 42
    - 2.4.2.1 Suzuki 43
    - 2.4.2.2. Stille 43
    - 2.4.2.3 Heck 44
    - 2.4.2.4 Other 45
  - 2.4.3 Type of Reaction: Aromatic Substitution; Nucleophilic (N-Arylation) 45
  - 2.4.4 Type of Reaction: Cleavage 44
    - 2.4.4.1 Cyclative Cleavage 46
      - 2.4.4.2 Functional Group: None (traceless) 47
      - 2.4.4.3 Functional Group: Halogens 48
      - 2.4.4.4 Functional Group: Alkenes 48
      - 2.4.4.5 Functional Group: Alcohols, Phenols 49
      - 2.4.4.6 Functional Group: Primary Amine 49
      - 2.4.4.7 Functional Group: *sec*-Amine 50
      - 2.4.4.8 Functional Group: *tert*-Amine 50
      - 2.4.4.9 Functional Group: Aldehyde/Ketone 51
      - 2.4.4.10 Functional Group: Hydroxamic Acid 51
      - 2.4.4.11 Functional Group: Amidine 52
      - 2.4.4.12 Functional Group: Guanidine 52
      - 2.4.4.13 Functional Group: *sec* Amide/*tert* Amide/Sulfonamide 52

2.4.5	Type of Reaction: Condensation	53
2.4.6	Type of Reaction: Cycloaddition	54
2.4.6.1	[2+2] Cycloaddition	54
2.4.6.2	[3+2] Cycloaddition	54
2.4.6.3	[4+2] Cycloaddition	55
2.4.7	Type of Reaction: Grignard and Related Reactions	55
2.4.8	Type of Reaction: Heterocycle Formation	55
2.4.8.1	Nitrogen-Containing Heterocycles	56
2.4.8.2	Multiple Nitrogen-Containing Heterocycles	57
2.4.8.3	Oxygen-Containing Heterocycles	58
2.4.8.4	Sulphur-Containing Heterocycles	58
2.4.9	Type of Reaction: Michael Addition	58
2.4.9.1	Immobilized Nucleophile	59
2.4.9.2	Thiol Addition	59
2.4.9.3	Amine Addition	59
2.4.10	Type of Reaction: Miscellaneous	60
2.4.11	Type of Reaction: Olefin Formation	61
2.4.11.1	Wittig	61
2.4.11.2	Horner-Emmons	62
2.4.11.3	Cross-Metathesis	62
2.4.11.4	$\beta$ -Elimination	62
2.4.12	Type of Reaction: Oxidation	62
2.4.12.1	Alcohol to Aldehyde/Ketone	62
2.4.12.2	Sulfide to Sulfoxide/Sulfone	63
2.4.12.3	Epoxidation	64
2.4.12.4	Ozonolysis	64
2.4.12.5	Other	64
2.4.13	Type of Reaction: Reduction	65
2.4.13.1	Aldehyde/Ketone to Alcohol	65
2.4.13.2	Nitro to Aniline	65
2.4.13.3	Azide to Amine	66
2.4.13.4	Amide to Amine	66
2.4.13.5	Reductive Alkylation/Amination	66
2.4.13.6	Other	67
2.4.14	Type of Reaction: Substitution	67
2.4.14.1	C-Alkylation (Aldol, Anion)	67
2.4.14.2	O-Alkylation (Mitsunobu, Anion)	68
2.4.14.4	N-Alkylation (Mitsunobu, Anion)	69
2.4.14.4	N-Alkylation (Sulfonamide, Amide)	69
2.4.14.5	S-Alkylation	70
2.5	References	70

### 3. Solution-Phase Combinatorial Chemistry

*Volkhard Austel*

- 3.1 Comparison with Solid-Phase Combinatorial Synthesis 77
- 3.2 Synthesis of Mixtures 79
- 3.3 Reactions Applied to Solution Phase Combinatorial Chemistry 80
  - 3.3.1 Acylation of Alcohols and Amines 81
  - 3.3.2 Sulfonation of Amines 86
  - 3.3.3 Formation of Ureas, Thioureas and Carbamates 86
  - 3.3.4 Alkylation and Addition Reactions 86
  - 3.3.5 Reductive Amination 88
  - 3.3.6 Arylation of Amines 89
  - 3.3.7 C-C Bond Formation via Condensation Reactions 89
  - 3.3.8 Pd-Catalyzed C-C Bond Formation 89
  - 3.3.9 Hydrogenations and Reductions 93
  - 3.3.10 Multicomponent Reactions 93
  - 3.3.11 Cyclisation Reactions 95
  - 3.3.12 Miscellaneous Reactions 95
  - 3.3.13 Reaction Sequences 103
- 3.4 Purification 106
  - 3.4.1 Solid-Phase-Bound Reagents 106
  - 3.4.2 Solid-Phase Extraction 109
  - 3.4.3 Liquid-Phase Extraction 114
  - 3.4.4 Fluorous Synthesis 116
  - 3.4.5 Synthesis on Soluble Polymers 117
- 3.5 Conclusions 119
- 3.6 References 119

### 4. Combinatorial Chemistry of Multicomponent Reactions

*Ivar Ugi, Alexander Dömling and Birgit Ebert*

- 4.1 Classical and Modern Chemistry of Isocyanides and MCRs 125
- 4.2 Early Studies and Concepts of MCR Chemistry 126
- 4.3 Conceptual Differences between Conceptual Chemical Reactions and MCRs 127
- 4.4 The Different Types of MCRs 128
- 4.5 The First Century of Isocyanide Chemistry 130
- 4.6 Complementary Aspects of Natural Product Syntheses by Tandem-Domino Reactions and MCR Chemistry 132
- 4.7 Modern Synthesis of the Isocyanides 134
- 4.8 The Introduction of the New Isocyanide MCRs and their Libraries 135
- 4.9 The Great Variability of the U-4CR 137
- 4.10 The Usual and Unusual Peptide Chemistry 141
- 4.11 Stereoselective U-4CRs and their Secondary Reactions 144

- 4.12 The MCRs of Educts with Two or Three Functional Groups 149
- 4.13 Unions of MCRs and Related Reactions 153
- 4.14 The MCRs of More than Four Functional Groups 154
- 4.15 New Unions of the U-4CR and Further Chemical Reaction 156
- 4.16 Progress in Combinatorial MCR Chemistry 160
- 4.17 Perspectives 161
- 4.18 Acknowledgements 162
- 4.19 References 162

## 5. Solid-Phase Anchors in Organic Chemistry

*Ralf Warrass*

- 5.1 Introduction 167
- 5.2 Acid-Labile Anchors 171
- 5.3 Anchors Cleaved by Nucleophiles 187
- 5.4 Photolysis-Labile Anchor 200
- 5.5 Allyl-Functionalized Anchors 204
- 5.6 Safety-Catch Anchor 207
- 5.7 Silicon-Anchor/Traceless-Anchor 212
- 5.8 Miscellaneous 216
  - 5.8.1 Multifunctional Anchors 216
  - 5.8.2 Anchors Cleaved by Reduction 217
  - 5.8.3 Hydrogenolysis-Labile Anchors 218
  - 5.8.4 Anchors Cleaved by Oxidation 219
  - 5.8.5 Enzymatically Cleavable Anchors 219
- 5.9 Acknowledgements 220
- 5.10 References 221

## 6. The Use of Templates in Combinatorial Chemistry for the Solid Phase Synthesis of Multiple Core Structure Libraries

*Hartmut Richter, Axel Trautwein, Tilmann Walk and Günther Jung*

- 6.1 Introduction 229
- 6.2 Squaric Acid as Template 231
- 6.3 Templates Derived from the Baylis–Hillman Reaction 233
  - 6.3.1 Baylis–Hillman Reaction on the Solid Phase 233
  - 6.3.2 The Use of 3-Hydroxy-2-Methylidene Propionic Acids as Templates 234
    - 6.3.2.1 Michael Addition of Amines 235
    - 6.3.2.2 Mitsunobu Reaction 236
    - 6.3.2.3 1,3-Dipolar Cycloaddition with Nitrile oxides 237
    - 6.3.2.4 Synthesis of 2-Methylidene- $\beta$ -Alanines 238
    - 6.3.2.5 Synthesis of Pyrazolones 240

## XIV *Contents*

6.3.2.6	Synthesis of 2-Diethoxy-Phosphorylmethyl Acrylic Acids	242
6.3.3	Alkylation of 2-Arylsulfonylaminomethyl Acrylic Acids	243
6.4	5-(2-Bromoacetyl)Pyrroles as Templates	244
6.4.1	Synthesis of Polymer-Bound 5-(2-Bromoacetyl)pyrroles	244
6.4.2	Synthesis of 5-(2-Aminoacetyl)pyrroles	246
6.4.3	Synthesis of Thiazolpyrroles, Aminothiazolpyrroles and Selenazolpyrroles	246
6.4.4	Synthesis of Imidazol[1,2-a]Pyri(mi)dylpyrroles	247
6.4.5	Synthesis of Benzofurylpyrrolylketones	248
6.4.6	Examples	248
6.5	The Use of Enones as Templates	249
6.5.1	Synthesis of Polymer-Bound Enones	250
6.5.2	Michael Addition of Aryl Thiolates	250
6.5.3	Synthesis of Pyridines and Pyrido[2,3-d]pyrimidines	251
6.5.4	Synthesis of Pyridones	252
6.5.5	Synthesis of Pyrrolidines	252
6.5.6	Synthesis of Pyrimidines and Pyrimidones	253
6.5.7	Synthesis of Pyrazoles	254
6.6	Conclusions	255
6.7	References	255

## 7. Combinatorial Synthetic Oligomers

*Alberto Bianco*

7.1	Introduction	257
7.2	Peptoids	258
7.2.1	Submonomer Approach to Peptoid Synthesis	258
7.2.2	Monomer Approach to Peptoid Synthesis	259
7.2.3	Peptomers	260
7.2.4	Peralkylated Peptides	261
7.2.5	$\beta$ -Peptoids	262
7.3	Oligocarbamates	263
7.4	Sulfonopeptides and Vinylogous-Sulfonamidopeptides	264
7.5	Poly-N-Acylated Amines	266
7.6	Oligoureas	268
7.6.1	Linear Ureas	268
7.6.2	Thioureas	270
7.6.3	Oligocycloureas and Cyclothioureas	270
7.6.4	Ureapeptoids	271
7.7	Heterocyclic Ring-Containing Oligomers	271
7.7.1	Poly-N-Methyl-Pyrroles and Imidazoles	271
7.7.2	Thiazole and Oxazole Ring-Containing Peptides	272
7.7.3	Polyisoxazolines	273
7.7.4	Oligothiophenes	273

7.7.5	Oligotetrahydrofurans	273
7.7.6	Pyrrolinone-Containing Oligomers	274
7.8	Other Examples of Synthetic Oligomers	275
7.8.1	Vinylogous Polypeptides	275
7.8.2	Retro-inverso Pseudopeptides	276
7.8.3	Azatides and Azapeptides	277
7.8.4	Polyketides	278
7.8.5	$\beta$ -Polypeptides	278
7.8.6	$\alpha,\alpha$ -Tetrasubstituted Amino Acids-Containing Peptides	279
7.8.7	Peptide Nucleic Acids	280
7.9	Single Position Peptide Bond Modifications	281
7.9.1	Thioamide Pseudopeptides	281
7.9.2	Reduced-Amide Bond Peptides	282
7.9.3	N-Hydroxyamide Bond-Containing Peptides	282
7.9.4	Hydroxyethylamine Peptide Bond Isosteres	283
7.9.5	Methylene Ether Isosteres	284
7.9.6	Phosphono- and Phosphinopeptides	285
7.10	Summary	286
7.11	Acknowledgements	286
7.12	References	287
8.	<b>Glycopeptide and Oligosaccharide Libraries</b>	
	<i>Phaedria M. St. Hilaire and Morton Meldal</i>	
8.1	Parallel Arrays versus Libraries of Compounds	291
8.2	Proteins Recognising Carbohydrates	291
8.3	The Carbohydrate Ligands	293
8.4	Supports for Solid-Phase Libraries	296
8.5	Analytical Tools for Oligosaccharide and Glycopeptide Libraries	298
8.5.1	Tagging Techniques in Oligosaccharide Libraries	298
8.5.2	Analysis by Mass Spectrometry	298
8.5.3	Structural Analysis of Compounds Linked to Single Beads by MAS-NMR	300
8.6	Introduction to Array and Library Synthesis of Oligosaccharides	301
8.6	Parallel Synthesis of Oligosaccharide Arrays	301
8.8	Synthesis of Oligosaccharide Libraries	302
8.9	Glycopeptide Templates as Oligosaccharide Mimetics	305
8.10	Parallel Synthesis of Glycopeptide Arrays	307
8.11	Preparation and Analysis of Solid-Phase Glycopeptide Template Libraries	310
8.12	Screening of a Solid-Phase Glycopeptide Library	312
8.13	Conclusions	313
8.14	Acknowledgements	313
8.15	References	314

9. **RNA and DNA Aptamers**

*Michael Famulok*

- 9.1 Introduction 319
- 9.2 Aptamers for Small Molecules 320
- 9.3 Functional Aptamers for Proteins and their Application in Biotechnology, Molecular Medicine, and Diagnostics 324
- 9.4 Functional Aptamers In Vivo 327
- 9.5 Conclusions 330
- 9.6 Acknowledgements 331
- 9.7 References 331

10. **Combinatorial Approaches to Molecular Receptors for Chemosensors**

*Dietmar Leipert, Jürgen Mack, Rolf Tünnemann and Günther Jung*

- 10.1 Introduction 335
- 10.2 Supramolecular Recognition Sites 336
- 10.3 Macrocyclic Peptides 339
- 10.4 Combinatorial Receptor Libraries 340
- 10.5 Cyclopeptides as Supramolecular Recognition Sites for Chemosensors 342
  - 10.5.1 Quartz Microbalance Measurements in the Liquid Phase 342
  - 10.5.2 Reflectometric Interference Spectroscopy (RIfS) in the Liquid Phase 346
- 10.6 Outlook 351
- 10.7 References 351

11. **Peptide Libraries in T-Cell-Mediated Immune Response**

*Burkhard Fleckenstein, Günther Jung and Karl-Heinz Wiesmüller*

- 11.1 Introduction 355
  - 11.1.1 Generation and Presentation of Antigens 355
  - 11.1.2 Combinatorial Peptide Libraries 356
  - 11.1.3 Peptide Libraries for the Investigation of TAP 356
  - 11.1.4 MHC Class I and Class II Molecules 357
    - 11.1.4.1 Peptide Ligands of MHC Class I Molecules 358
    - 11.1.4.2 Peptide Ligands of MHC Class II Molecules 359
  - 11.1.5 T-Cell Response 360
    - 11.1.5.1 Recognition of MHC Class I-Bound Peptides by TCR 360
    - 11.1.5.2 Recognition of MHC Class II-Bound Peptides by TCR 361

11.2	Methods	362
11.2.1	Synthesis of Peptides and Peptide Libraries	362
11.2.2	Analytical Characterization of Peptide Libraries	363
11.2.3	HPSEC-Competition Assay	363
11.2.4	Isolation of HLA Class II Molecules	363
11.3	Results and Discussion	366
11.3.1	Optimal Length of Peptides for MHC Class II Interaction	366
11.3.2	Activity Pattern Describing DR1/Peptide Interaction	366
11.3.2.1	Tolerance to Amino Acid Variations in a HLA Class II Ligand	370
11.3.2.2	Competition of Peptides Deduced from the Activity Pattern of the Undecapeptide Library	371
11.3.2.3	Predictions of HLA Class II-Ligands and T-Cell Epitopes by the Algorithm, Actipat	371
11.3.2.4	Peptide Libraries and the Antigen Recognition of CD4 <sup>+</sup> T Cells	372
11.3.2.5	T-Cell Response to Completely Randomized Peptide Libraries	373
11.3.2.6	Complete Dissection of the Epitope for TCC 5G7	374
11.3.2.7	Evaluation of TCR Contact Positions	375
11.3.2.8	Superagonists and Prediction of Epitopes	375
11.4	Conclusions	376
11.5	Acknowledgements	376
11.6	References	376

## 12. Combinatorial Biosynthesis of Microbial Metabolites

*Andreas Bechthold and José Antonio Salas Fernández*

12.1	Introduction	381
12.2	Cloning of Biosynthetic Gene Clusters	381
12.3	New Drugs by Genetic Engineering	382
12.3.1	New Drugs by Targeted Gene Disruption	384
12.3.2	New Drugs by Expression of Single Genes	385
12.3.3	New Drugs by Expression of Gene Clusters	388
12.3.4	New Drugs by Variation of the Starter Units	390
12.3.5	New Drugs by Recombinant Assembly of Enzymatic Subunits	390
12.3.5.1	Polyketide Synthases	391
12.3.5.2	Peptide Synthases	399
12.3.5.3	Proteins Involved in Deoxysugar Biosynthesis	402
12.4	Future Perspectives	403
12.4.1	Use of Genes Involved in the Biosynthesis of Oligosaccharide Antibiotics	403
12.4.2	Use of Genes from Other Sources – New Compounds from Earth	404
12.4.3	Changing the Substrate Specificity of an Enzyme	404
12.5	References	404



### 13. Design and Diversity Analysis of Compound Libraries for Lead Discovery

*Hans Matter and Matthias Rarey*

- 13.1 Introduction 409
- 13.2 Concepts and Issues for Combinatorial Library Design 410
- 13.3 The Similarity Principle 412
- 13.4 Molecular Descriptors and Selection Techniques 413
- 13.4.1 Two-Dimensional Fingerprints 413
- 13.4.2 Pharmacophore Definition Triplets 414
- 13.4.3 Other Descriptors 414
- 13.4.4 Compound Selection Techniques 415
- 13.5 Selected Examples of New Approaches to Molecular Similarity 416
- 13.5.1 Affinity Fingerprints 417
- 13.5.2 Feature Trees 418
- 13.5.3 Automated Structural Superposition of Fragments 420
- 13.6 Descriptor Validation Studies 421
- 13.6.1 2D Versus 3D Descriptors for Global Diversity 421
- 13.6.2 Random Versus Rational Design for Global Diversity 422
- 13.6.3 3D Pharmacophore Definition Triplets Versus 2D Fingerprints 424
- 13.6.4 Local Similarity – The Radius of Similarity 426
- 13.7 Designing Combinatorial Libraries 428
- 13.7.1 Strategy 428
- 13.7.2 Practical Issues 429
- 13.7.3 Analysis of Diverse Libraries 432
- 13.8 Conclusions 433
- 13.9 Acknowledgement 435
- 13.10 References 435

### 14. How to Scale-up While Scaling Down: EVOscreen™, a Miniaturized Ultra High Throughput Screening System

*Rodney Turner, Sylvia Sterrer, Karl-Heinz Wiesmüller and Franz-Josef Meyer-Almes*

- 14.1 Introduction to Drug Discovery and Screening 441
- 14.1.1 A Rationale for New Methods in Drug Discovery 441
- 14.1.2 Chemistry, Biology and Technology 442
- 14.1.3 Sources of Compounds for HTS 444
- 14.1.4 Demands on HTS 444
- 14.2 Combinatorial Chemistry 445
- 14.2.1 Design and Production of Chemical Diversity 445
- 14.2.2 Microreactor Systems 446
- 14.2.3 Procedures for the Synthesis of Combinatorial Compounds 446
- 14.2.4 Solution-Phase Combinatorial Chemistry 447

14.2.5	Limitations on Combinatorial Design Strategies	448
14.3	Screening Technology	448
14.3.1	The Ideal Read-out Technology: Confocal Fluorescence	448
14.3.2	Examples of Specific Fluorescence Detection Technologies	450
14.3.2.1	Fluorescence Correlation Spectroscopy (FCS)	450
14.3.2.2	Fluorescence Resonance Energy Transfer (FRET)	450
14.3.2.3	Fluorescence Polarization (FP)	451
14.3.2.4	Fluorescence Imaging	451
14.3.3	EVOscreen™	451
14.3.3.1	Liquid Handling	452
14.3.3.2	Scanner/Picker	453
14.3.4	Blurring the Lines Between Primary, Secondary and Tertiary Screening	454
14.3.5	Using Primary Cells	455
14.4	Practical Application	455
14.4.1	Association Kinetics of Gene Product Fragments Derived from <i>E. coli</i> $\beta$ -Galactosidase	455
14.4.1.1	Association Kinetics	457
14.4.2	Comparison of Native (EA) <sub>2</sub> and Indicated (EA) <sub>2</sub>	459
14.4.3	FCS Adaption	459
14.5	Summary	459
14.6	Acknowledgements	460
14.7	References	460

## 15. High Throughput Experimentation in Catalysis

*Ferdi Schüth, Christian Hoffmann, Anke Wolf, Stephan Schunk, Wolfram Stichert and Armin Brenner*

15.1	Introduction	463
15.2	General Considerations	464
15.2.1	Motivation for and Problems of HTE Approaches in Catalysis	464
15.2.2	Strategies for Library Design and Testing	467
15.3	Approaches to Synthesis	469
15.4	Approaches to Testing	471
15.5	Conclusions	476
15.6	Acknowledgements	477
15.7	References	477

## 16. FT-Infrared Spectroscopy and IR-Microscopy for On Bead Analysis of Compound Libraries

*Holger Bandel, Wolfgang Haap and Günther Jung*

16.1	Introduction	479
16.2	Analytical Methods Using FT-IR Spectroscopy	479

16.2.1	KBr Pellet Method	479
16.2.2	ATR-Spectroscopy	482
16.2.3	FT-IR Microscopy	484
16.2.3.1	Single Bead Reaction Monitoring	484
16.2.3.2	Examination of the Interaction between Resin-Bound Reactive Groups via IR-Microscopy	485
16.2.3.3	FT-IR Mapping: A New Tool for Spatially Resolved Characterization of Polymer-Bound Combinatorial Compound Libraries with IR-Microscopy	490
16.2.4	Other IR-Spectroscopy Methods	496
16.2.4.1	DRIFTS (Diffuse Reflectance Infrared Fourier Transform) Spectroscopy	496
16.2.4.2	Photoacoustic Spectroscopy	497
16.2.4.3	FT Raman Spectroscopy	497
16.3	Conclusions	497
16.4	References	498

## 17. Mass Spectrometric Analysis of Combinatorially Generated Compounds and Libraries

*Roderich Süßmuth, Axel Trautwein, Hartmut Richter,  
Graeme Nicholson and Günther Jung*

17.1	Introduction	499
17.2	Significance of Mass Spectrometry in the Analytical Concept of Combinatorial Chemistry	499
17.2.1	Decisional Pathway for Combinatorial Synthesis and Analysis	499
17.2.2	Mass Spectrometric Techniques for Combinatorial Compound Analysis	502
17.2.2.1	Ionization Techniques	503
17.2.3	Analysis of Peptide- and Oligonucleotide-Libraries	504
17.2.4	Analysis of Non-Oligomeric Compounds and Libraries	505
17.2.4.1	Direct Analysis of Non-Oligomeric Compound Libraries	505
17.2.4.2	Coupling of Separation Techniques with Mass Spectrometry	505
17.2.5	“On-Bead” versus “Off-Bead” Analysis	506
17.2.6	High-Throughput Systems for Sample Analysis and Purification	506
17.2.6.1	High-Throughput Analysis	506
17.2.6.2	High-Throughput Sample Purification using MS-Detection	507
17.2.6.3	Automated Software-Aided Spectra Evaluation	508
17.3	Application of ES-MS in Combinatorial Chemistry Analysis	508
17.3.1	ES-MS-Analysis of Single Compounds	509
17.3.1.1	Influence of Resin Types on the MS-Analysis	509
17.3.1.2	Influence of Reagents on the MS-Analysis	509
17.3.1.3	Influence of Fragmentation on the MS-Analysis	511
17.3.1.4	High-Throughput Analysis	511

17.3.2	ES-MS-Analysis of Non-Oligomeric Libraries	514
17.3.2.1	ES-MS-Analysis of a Pyrrol-Library	514
17.3.2.2	MS-Analysis of an Isoxazoline-Library	523
17.4	Conclusions	528
17.5	Materials and Methods	528
17.5.1	Electrospray-Mass-Spectrometry	528
17.5.2	GC-EI Mass Spectrometry	529
17.6	Acknowledgements	529
17.7	Appendix	529
17.8	References	531
18.	<b>High Resolution Magic Angle Spinning (MAS) NMR Spectroscopy for On-Bead Analysis of Solid-Phase Synthesis</b>	
	<i>Ralf Warrass and Guy Lippens</i>	
18.1	Introduction	533
18.2	Methods Used for On-Bead NMR Analysis	533
18.3	Sample Preparations, Observations and Quantifications by HR MAS NMR	535
18.4	Use of Protonated Solvent in HR MAS NMR	539
18.5	Summary and Perspectives	540
18.6	References	542
19.	<b>Automated Combinatorial Chemistry</b>	
	<i>Martin Winter, Karl-Heinz Wiesmüller and Günther Jung</i>	
19.1	Introduction	543
19.2	Parallel Synthesis of Individual Compounds	544
19.3	Productivity in the Research Laboratory	545
19.4	Automation Concepts: How Multicomponent Systems are Organized	545
19.4.1	General Considerations	545
19.4.2	Decentralized Automation Systems	549
19.4.3	Central-Controlled, Function-Oriented Multicomponent Systems	550
19.4.4	Central-Automated, Sample-Oriented Multicomponent Systems	551
19.5	Collateral Technologies for High-Throughput Purification and Analysis	556
19.6	Summary and Prospects	557
19.7	Acknowledgements	558
19.8	References	558

20. **Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS) for Analysis of Compound Libraries**

*Tilman B. Walk, Axel Trautwein, Holger Bandel and Günther Jung*

20.1	Introduction	561
20.2	Physical Basis of FT-ICR Mass Spectrometry	562
20.2.1	Instrument	562
20.2.1.1	Superconducting Magnet	562
20.2.1.2	Vacuum System	563
20.2.1.3	Analyzer Cell	563
20.2.2	Detection Principle	564
20.2.2.1	Cyclotron Motion	564
20.2.2.2	Signal Generation	565
20.2.2.3	Mass Accuracy	567
20.2.2.4	MS <sup>n</sup> -Experiment	567
20.3	FT-ICR-MS Analysis	567
20.3.1	Combinatorial Compound Collections	567
20.3.1.1	Synthesis of Pyrrole Amide Collections	568
20.3.1.2	Automated Direct Measurement	569
20.3.1.3	MicroLC/ES-FT-ICR-MS-Coupling	573
20.3.2	Single Bead Analysis	575
20.3.2.1	Synthesis of Pyrrolidine Libraries	575
20.3.2.2	Result of Single Bead Analysis	577
20.4	Materials and Methods	578
20.5	References	579

Colour Plates 583

Index 589

# List of Abbreviations

A-3CR	Asinger – three component reaction
Aba	2-aminobenzoic acid
Ac	acetyl
ACN	acetonitrile
ACP	acyl carrier protein
ADME	absorption, distribution, metabolism, excretion
ADPV	5-(4-aminomethyl-3,5-dimethoxy-phenoxy)-valeric acid
Alloc	allyloxycarbonyl
AM	aminomethyl
AMCA	7-amino-4-methyl-coumarin-3-acetic acid
ANP	3-amino-3-(2-nitrophenyl) propionic acid
APCI	atmospheric pressure chemical ionization
ATP	adenosine triphosphate
ATR-IR	attenuated total reflection infrared spectroscopy
BAL	backbone amide linker
BAP	borane/pyridine complex
BB-4CR	Berg-Bucherer – four component reaction
BBTO	bis-(tributyl)tin oxide
BEMP	2-tert-butylimino-2-diethylamino-1,3-dimethylperhydro-1,3,2-diazaphosphorine
BHA	benzyhydrilamine
BINAP	2,2'-bis(diphenylphosphino)-1,1'-binaphthyl
Bn, Benz	benzyl
Boc	tert.-butyloxycarbonyl
BOP	benzotriazolyl-oxy-tris(dimethylamino)phosphonium hexafluorophosphate
Bz	benzoyl
CAD	collision activated dissociation
CALD	computer assisted library design
CAN	ceric ammonium nitrate
CBP	carbohydrate binding proteins
Cbz	benzyloxycarbonyl
CD	circular dichroism
CDI	N,N-carbonyldiimidazol
CE	capillary electrophoresis

XXIV *List of Abbreviations*

CHC	a-cyano-4-hydroxy-cinnamic acid
CHO	chinese hamster ovary (cells)
Chx	cyclohexyl
CID	collision induced decay
CID	collision induced decomposition
CLEAR	cross-linked ethoxylate acrylate resin
CLF	chain length factor protein
CLIP	class II-associated invariant chain peptide
CMPP	chloromethyl phenyl pentyl polystyrene
CoA	coenzyme A
Cod	cyclooctadienyl
CPG	controlled pore glass
CRD	carbohydrate recognition domain
CTL	cytotoxic T lymphocytes
cHex	cyclohexyl
Cy	cyclohexyl
CZE	capillary zone electrophoresis
Da	dalton
DABCO	1,4-diazabicyclo[2.2.2]octane
DBE	double bond equivalent
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCI	desorption chemical ionization
DCM	dichloromethane
Dde	1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl
DDQ	dichlorodicyano-p-benzochinon
DEAD	diethylazodicarboxylate
DEBS	6-deoxyerythronolide B synthase
DH	dehydratase
DIAD	diisopropyl azodicarboxylate
DIBAH	diisobutylaluminium hydride
DIBAL	diisobutylaluminium hydride
DIC	diisopropylcarbodiimide
DIEA	diisopropylethylamine
DMAE	dimethylaminoethanol
DMAP	4-N,N-dimethylaminopyridine
DMF	N,N-dimethylformamide
DMSO	dimethylsulfoxide
DMTST	dimethyl(methylthio)sulfonium-trifluoromethylsulfonate
DNA	desoxyribonucleic acid
DRIFTS	diffuse reflectance infrared Fourier transform spectroscopy
DSB	4-(2,5-dimethyl-4-methylsulfinylphenyl)-4-hydroxybutyric acid
DVB	divinylbenzene
EA	enzyme acceptor
EBV	Epstein-Barr virus
ED	enzyme donor

EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
EI	electron impact
ELISA	enzyme-linked immunosorbent assay
ER	endoplasmatic reticulum
ER	enoyl reductase
ESI	electrospray ionization
ES-MS	electrospray-mass spectrometry
Et	ethyl
FAB	fast atom bombardment
FCS	fluorescence correlation spectroscopy
FITC	fluoresceinisothiocyanate
FMN	flavinemononucleotide
Fmoc	fluorenylmethyloxycarbonyl
FP	fluorescence polarization
FRET	fluorescence resonance energy transfer
FT	Fourier transform
FT-IR	Fourier transform infrared spectroscopy
Fuc	L-fucose
GalNAc	N-acetyl-D-galactosamine
GC	gas chromatography
GlcNAc	N-acetyl-D-glucosamine
HFIP	hexafluoroisopropanol
HHR	hammer head ribozyme
HIV	human immunodeficiency virus
hKGF	human keratinocyte growth factor
HLA	human leukocyte antigen (MHC)
HMBA	4-hydroxymethylbenzoic acid
HMPB	4-hydroxymethyl-3-methoxyphenoxybutyric acid
HMP-PS	p-benzyloxybenzyl alcohol polystyrene
HMQC	heteronuclear correlation quantum coherence
HOBT	hydroxybenzotriazole
HPLC	high performance liquid chromatography
HPSEC	high-performance size-exclusion chromatography
HQN	hydroxy-quinuclidine
HTE	high-throughput experimentation
HTS	high-throughput screening
HR	high resolution
iPr	isopropyl
IR	infrared
ICR-MS	ion cyclotron resonance mass spectrometry
ICR	ion cyclotron resonance
IRMDP	multiphoton infrared photodissociation
kDa	kilodalton
KHMDS	potassium bis(trimethylsilyl)amide
KR	ketoreductase



XXVI *List of Abbreviations*

KS	ketosynthase
Lau	lauric acid
LC	liquid chromatography
LCAA-CPG	long-chain alkyl amino controlled pore glass
LC-MS	liquid chromatography-mass spectrometry
LD	laser desorption
M-3CR	Mannich – three component reaction
MALDI	matrix assisted laser ionization
MAMP	alpha-methoxy-phenyl (4-methoxy-benzhydryl)
Man	mannose
MAS-NMR	magic angle spinning-nuclear magnetic resonance
MBHA	4-methylbenzhydramine
MBP	myelin basic protein
mCPBA	meta-chloroperoxobenzoic acid
MCPS	multiple column peptide synthesis
MCR	multicomponent reaction
MCSL	multiple core structure libraries
Me	methyl
Me-Du-Phos	1,2-bis(2,5-dimethylphospholano)benzene
MHC	major histocompatibility complex
MOBHA	4-methoxybenzhydramine
MPR	mannose-phosphate receptor
Ms	mesyl, methanesulfonyl
MS	mass spectrometry
MsCl	methanesulfonylchloride
MTBD	7-methyl-1,5,7-triazabicyclo[4,4,0]elec-5-ene
MTP	microtiter plate
Mtr	4-methoxy-2,3,6-trimethylphenyl-sulfonyl
Nasyl	naphthylsulfonyl
NBA	4-aminomethyl-3-nitrobenzoic acid
Nbb-PS	aminobenzhydryl-polystyrene
NBHA	ortho-nitrobenzylhydramino (2-nitrobenzhydramine)
NeuAc	N-acetyl-neuraminic acid
NIS	N-iodosuccinimide
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
NPE	4-hydroxyethyl-3-nitro-benzoic acid
Npss	2-methoxy-5-(2-(2-nitrophenyl)dithio-1-oxypropylphenylacetic ac
NSG	N-substituted oligoglycine
NBS-Cl	o-nitrobenzenesulfonyl chloride
P	partition coefficient
P-3CR	Passerini – three component reaction
PAAAs	poly-N-acylated amines

PAL	peptide amide linker: 5-[4[9-fluorenylmethoxycarbonylaminoethyl]-3,5-dimethoxyphenoxy]valeric acid
PAM	phenylacetamidomethyl
PBS	phosphate buffered saline
PCA	principal component analysis
PCR	polymerase chain reaction
PDC	pyridinium dichromate
PDTs	pharmacophore definition triplets
PEG	polyethyleneglycol
PEGA	polyethylene glycol-polyacrylamide co-polymer
PEG-PS	polyethylene glycol-polystyrene
PET	photo-induced electron transfer effect
Pfp	pentafluorophenyl
Ph	phenyl
PhFl Resin	9-phenylfluorene-9-yl-PS/DVB
Pip	piperidine
PK	pharmacokinetics
PKSs	polyketide synthases
PNAs	peptide nucleic acids
PNP	para-nitrophenyl
POEPOP	polyoxyethylene-epichlorohydrine derivative
POEPS	polyoxyethylene-cross-linked-oligostyrene
Ppa	phenylpropionic acid
PS	polystyrene
PS-DVB	polystyrene-divinylbenzene copolymer
PSD	post-source decay
PS-PEG	polystyrene-polyethyleneglycol graft polymer
PyBroP	bromo-tris-pyrrolidino-phosphonium-hexafluorophosphate
QSAR	quantitative structure-activity relationship
R	alkyl
RCM	ring closure metathesis
relSI	relative stimulation index
REMPI	laser-induced resonant multiphoton ionization
Rev	reverse transcriptase (HIV)
RIFS	reflectometric interference spectroscopy
RP-HPLC	reversed phase-high performance liquid chromatography
RRE	Rev-responsive element
S-3CR	Strecker – three component reaction
SAR	structure-activity relationship
SASRIN	super acid sensitive resin
SDS-PAGE	sodium dodecylsulfate-polyacrylamide gel electrophoresis
SI	stimulation index
S-Le	sialyl – Lewis
SMSI	string metal support interaction
SPOC	solid-phase organic chemistry

XXVIII *List of Abbreviations*

SPOCC	polyoxyethylene-oxetane derivative
SPOS	solid phase organic synthesis
SPPS	solid phase peptide synthesis
ssRNA	single stranded ribonucleic acid
TAR RNA	trans-activation response region
TAP	transporter associated with antigen processing
TBAF	tetrabutylammonium fluoride
Tbf	tetrabenz[a,c,g,i]fluorene
tBu	tert.-butyl
TCC	CD4 <sup>+</sup> T-cell clones
TCR	T-cell receptor
TE	thioesterase
TEA	triethyl amine
Tf	trifluoromethylsulfonyl
TFA	trifluoroacetic acid
TFE	trifluoroethanol
TFMSA	trifluoromethanesulfonic acid
THF	tetrahydrofurane
THP	tetrahydropyranyl
TIC	total ion current
TIPS	triisopropylsilane (TIS)
TIS	triisopropylsilane (TIPS)
TLC	thin layer chromatography
TMS	trimethylsilyl
TMOF	trimethylsilyl trifluoromethanesulfonate
TMSOTf	trimethylsilyl trifluoromethanesulfonate
Tn	tags for labeling
TOCSY	total correlation spectroscopy
TOF	time-of-flight
TPAP	thrombin receptor activating peptide
Ts	tosyl, para-toluenesulfonyl
U-4CR	Ugi – four component reaction
uHTS	ultra high-throughput screening
UV	ultra-violet
UVDD	ultra-violet photodissociation
VEGF	vascular endothelial growth factor
VLSIPS	very large-scale immobilized polymer synthesis
VPF	vascular permeability factor
WHIM	weighted holistic invariant molecular (indices)
Z	benzyloxycarbonyl

Further special abbreviations of linkers and resins are found in Chapter 5.

# List of Contributors

V. Austel  
Boehringer Ingelheim Pharma KG  
Postfach 1755  
D-88397 Biberach/Riss  
Germany

H. Bandel  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

S.E. Booth,  
N.V. Organon  
Lead Discovery Unit  
Organon Laboratories  
Newhouse, ML1 5SH  
Scotland, UK

A. Bianco  
IBMC  
Institut de Biologie Moléculaire et Cellu-  
laire  
CNRS  
UPR 9021  
15 Rue René Descartes  
F-67084 Strassbourg  
France

A. Bechthold  
Pharmazeutische Biologie  
Universität Tübingen  
Auf der Morgenstelle 8  
D-72076 Tübingen  
Germany

A. Brenner  
HTE GmbH  
Kurpfalz Ring 104  
D-69123 Heidelberg-Pfaffengrund  
Germany

A. Dömling  
Institut für Organische Chemie und Bio-  
chemie  
Technische Universität München  
Lichtenbergstrasse 4  
D-85747 Garching  
Germany

C.M. Dreef Tromp  
N.V. Organon  
Lead Discovery Unit  
P.O.Box 20  
NL-5340 BH OSS  
The Netherlands

A.P.A. de Han  
N.V. Organon  
Lead Discovery Unit  
P.O.Box 20  
NL-5340 BH OSS  
The Netherlands

B. Ebert  
Institut für Organische Chemie und Biochemie  
Technische Universität München  
Lichtenbergstrasse 4  
D-85747 Garching  
Germany

M. Famulok  
Institut für Biochemie- Genzentrum  
TMU - München  
Feodor-Lynen-Str. 25  
D-81377 München  
Germany

B. Fleckenstein  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

W. Haap  
Ciba Spezialitätenchemie  
F \* E Antimicrobials  
Postfach 1266  
D-79630 Grenzach-Wyhlen  
Germany

P.H.H. Hermkens,  
N.V. Organon  
Lead Discovery Unit  
P.O.Box 20  
NL-5340 BH OSS  
The Netherlands

P.M.St. Hilaire  
Carlsberg Lab.  
Dept. of Chemistry  
Gamle Carlsberg Yej 10  
DK-2500 Valby, Copenhagen  
Denmark

C. Hoffmann  
Max-Planck-Institut für Kohleforschung  
Kaiser-Wilhelm-Platz 1  
D-45470 Mülheim / Ruhr  
Germany

G. Jung  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

D. Leipert  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

H. Matter  
Hoechst Marion Roussel Deutschland  
GmbH  
Computational Chemistry  
Building G-838  
Chemical Research  
D-65926 Frankfurt/Main  
Germany

J. Mack  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

M. Meldal  
Carlsberg Lab.  
Dept. of Chemistry  
Gamle Carlsberg Yej 10  
DK-2500 Valby, Copenhagen  
Denmark

F.-J. Meyer-Almes  
EVOTEC BioSystems AG  
Schnackenburgallee 114  
D-22525 Hamburg  
Germany

G. Nicholson  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

H.C.J. Ottenheijm  
N.V. Organon  
Lead Discovery Unit  
P.O.Box 20  
NL-5340 BH OSS  
The Netherlands

M. Rarey  
German National Research Center  
for Information Technology  
Institute for Algorithms and Scientific  
Computing (SCAI)  
Schloß Birlinghoven  
D-53754 Sankt Augustin  
Germany

H. Richter  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

J.A. Salas  
Universidad de Oviedo  
Departamento de Biología Funcional  
E-33006 Oviedo  
Spain

S. Schunk  
HTE GmbH  
Kurpfalz Ring 104  
D-69123 Heidelberg-Pfaffengrund  
Germany

S. Sterrer  
EVOTEC BioSystems GmbH  
Schnackenburgallee 114  
D-22525 Hamburg  
Germany

W. Stichert  
HTE GmbH  
Kurpfalz Ring 104  
D-69123 Heidelberg-Pfaffengrund  
Germany

R. Süßmuth  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

F. Schüth  
Max-Planck-Institut für Kohleforschung  
Kaiser-Wilhelm-Platz 1  
D-45470 Mühlheim / Ruhr  
Germany

D. Tiebes  
DECHEMA  
Theodor-Heuss-Allee 25  
D-60486 Frankfurt/Main  
Germany

A. Trautwein  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

R. Turner  
EVOTEC BioSystems AG  
Schnackenburgallee 114  
D-22525 Hamburg  
Germany

R. Tünnemann  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

I. Ugi  
Institut für Organische Chemie und Bio-  
chemie  
Technische Universität München  
Lichtenbergstrasse 4  
D-85747 Garching  
Germany

R. Warrass  
Groupe RMN  
Institut Pasteur de Lille  
1 Rue du Professeur Calmette  
F-59021 Lille Cedex  
France

T. Walk  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

K.-H. Wiesmüller  
EMC microcollections  
Sindelfingerstrasse 3  
D-72070 Tübingen  
Germany

M. Winter  
Institut für Organische Chemie  
Universität Tübingen  
Auf der Morgenstelle 18  
D-72076 Tübingen  
Germany

A. Wolf  
Max-Planck-Institut für Kohleforschung  
Kaiser-Wilhelm-Platz 1  
D-45470 Mülheim / Ruhr  
Germany

# 1 Combinatorial Chemistry

*Dominika Tiebes*

## 1.1 Introduction

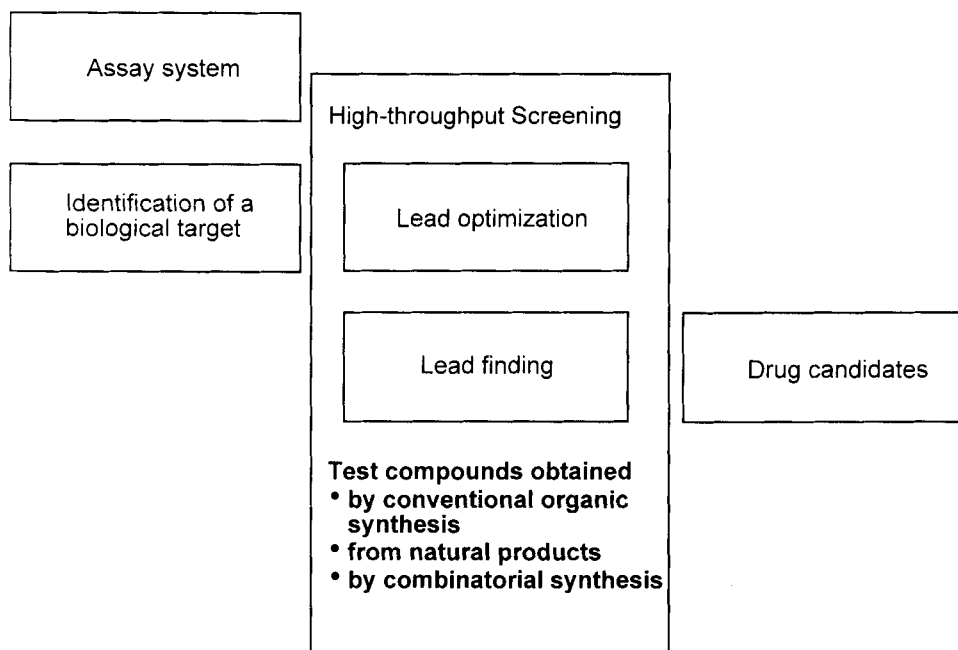
Molecular medicine has changed the way that new drugs are discovered and optimized. Scientists no longer mainly view disease on a large scale but, by investigating the molecular pathways involved, are identifying compounds that relate to the disease in question. Having identified a potential biological target, they can join the race to find appropriate leads that can interact with the target, resulting in a causal therapy. Thousands of leads and modified leads are tried against each biological target within high-throughput screening, and it is in this area where combinatorial chemistry becomes invaluable.

Conventionally, drug development involved the individual synthesis and biological evaluation of hundreds to thousands of organic compounds in an attempt to enhance the biological activity, selectivity and bioavailability, while at the same time decreasing its toxicity (Fig. 1.1.). On average, such a drug discovery process takes 6.5 years before a suitable drug candidate is found for further clinical evaluation.

Combinatorial synthesis as a high-throughput method enables the rapid production of hundreds to thousands times more compounds than conventional serial organic synthesis. This in conjunction with high-throughput screening, and the increasing number of new biological targets is expected to accelerate the process for discovering new and improved drug candidates [1–11]. Thus, since the early 1990s, combinatorial chemistry has attracted the attention of companies as a means of cutting the time and high costs associated with the serial drug discovery process.

The underlying principles used in combinatorial chemistry in the synthesis of organic compounds have their origin in the field of peptide chemistry. The design of Geysen's multipin apparatus in 1984 [12, 13] and Houghten's tea-bag system in 1985 (tea-bag method) [14] addressed methods for multiple (parallel) synthesis of peptides, based on the techniques of solid-phase peptide synthesis according to Merrifield [15] that was presented first in 1963. Further progress was achieved in this field, among others by Frank (spot synthesis) [16] and Fodor (VLSIPS method) [17, 18]. However, combinatorial synthesis of peptides in the literal sense was made possible by the work of Furka (split-pool method) [19–22] in 1988, and this can be regarded as the 'birth' of combinatorial chemistry. These studies have clearly demonstrated the power of combinatorial synthesis in medical chemistry. But peptides and oligonucleotides generally have low bioavailability and are rapidly degraded proteolytically; therefore, their use as therapeutic agents is often limited [23].





**Figure 1.1.** Drug discovery process to new or improved drug candidates in the R & D laboratories of pharmaceutical and agrochemical companies.

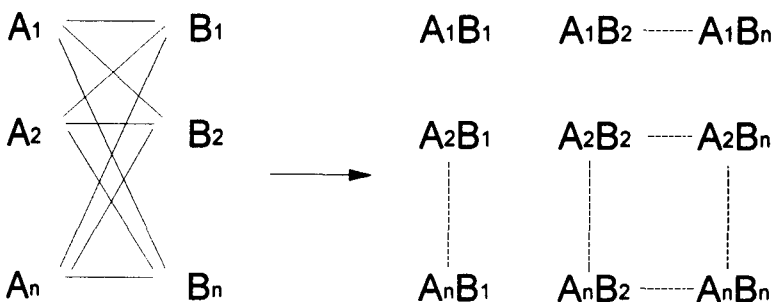
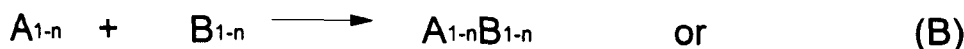
Due to the more favorable pharmacokinetic properties of many organic compounds (generally with a molecular mass < 750 Da), termed small molecules, combinatorial libraries of small molecules have rapidly become an important tool in organic chemistry.

However, it is not only the drug discovery process that might benefit from combinatorial chemistry, as the principles are being applied increasingly in the search for new materials [24–32] and for better catalysts [33–43].

## 1.2 Principles of Combinatorial Chemistry

The basic principle of combinatorial chemistry is to prepare a large number of different compounds at the same time – instead of synthesizing compounds in a conventional one-at-a-time manner – and then to identify the most promising compound for further development by high-throughput screening.

The characteristic of combinatorial synthesis is that different compounds are generated simultaneously under identical reaction conditions in a systematic manner, so that ideally the products of all possible combinations of a given set of starting materials (termed building blocks) will be obtained at once (Fig. 1.2.). The collection of these finally synthesized compounds is referred to as a *combinatorial library*. (The term ‘library’ is not used consistently in the literature. Often this designation is reserved for compound



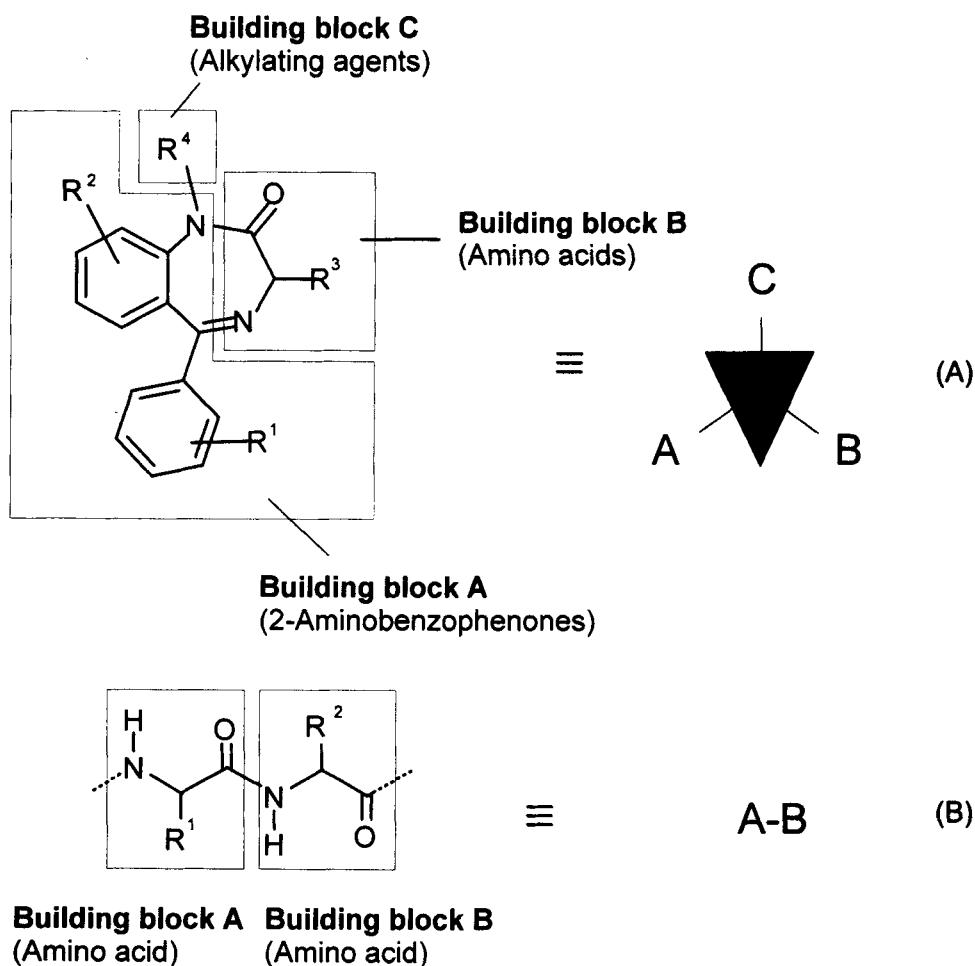
**Figure 1.2.** (A) In general, in a conventional synthesis one starting material A reacts with one reagent B resulting in one product AB. (B) In a combinatorial synthesis different building blocks of type A ( $A_1$ – $A_n$ ) are treated simultaneously with different building blocks of type B ( $B_1$ – $B_n$ ) according to combinatorial principles, i.e. each starting material A reacts separately with all reagents B resulting in a combinatorial library  $A_{1-n}B_{1-n}$ .

mixtures, whereas collections of separate single substances are called ‘arrays’.) The library is then screened for the property in question and the active compounds are identified.

The combinatorial libraries can be structurally related by a central core structure, termed scaffold, or by a common backbone, as shown in Figure 1.3. In both cases, the accessible dissimilarity of the compounds within the library depends on the building blocks which are used for the construction (see also Chapter 1.6.).

The size of the combinatorial library ranges from a few tens to many hundreds of thousands compounds. In a multistep combinatorial synthesis, generally required for drug-relevant compounds, the total number of all synthesized compounds within the combinatorial library is determined by two factors: the number of building blocks used per reaction, and the number of reaction steps, in which a new building block is incorporated. The number of all synthesized compounds ( $N$ ) is given by the equation  $N = bx$ , where  $b$  is an equal number of building blocks used in each reaction ( $1, 2, \dots, x$ ) and  $x$  is the number of reaction steps, in which a new building block is introduced. If the number of building blocks for each reaction step varies (e.g.,  $a, b, c$  in a three-step synthesis), then  $N = a \cdot b \cdot c$ . Thus, a large compound library will be obtained rapidly from only a few building blocks.

Depending on the synthetic strategy used the created combinatorial library can comprise compound mixtures, or separate, single substances. (The term ‘combinatorial synthesis’ was strictly used only when different building blocks are treated in one reaction vessel, and compound mixtures are thus obtained. Since the boundary becomes blurred, herein syntheses of both libraries of compound mixtures and of separate, single substan-



**Figure 1.3.** Scaffold-based (A) or backbone-based (B) combinatorial libraries.

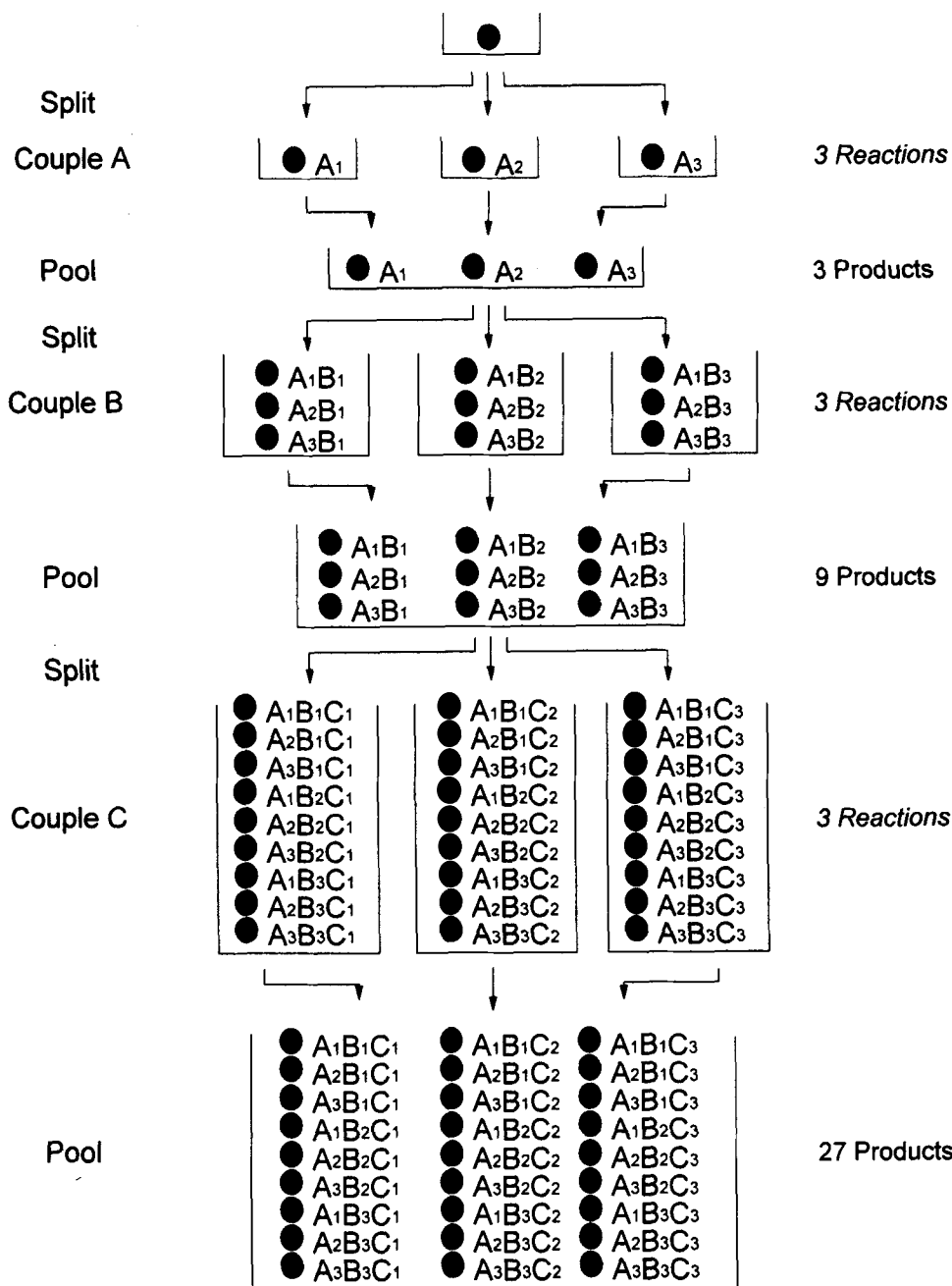
ces are included, as long as the procedure can be described as combinatorial.) Combinatorial libraries may consist of compounds free in solution, or linked to a solid phase.

### 1.3 Methods and Techniques of Combinatorial Synthesis

#### 1.3.1 Synthetic Strategies towards Combinatorial Libraries

##### 1.3.1.1 Split-Pool Synthesis towards Combinatorial Libraries

Furka and co-workers pioneered the split-pool synthesis method [19–22] for the synthesis of large peptide libraries in 1988; this approach is termed divide, couple, and recombine synthesis by other workers [44–47].



**Figure 1.4.** Split-pool synthesis to prepare combinatorial libraries of compound mixtures. (Spheres represent resin beads, A, B, C represent the sets of building blocks, borders represent the reaction vessels.) In the case of three building blocks used, in each coupling step after three stages a total number of 27 different compounds, one on each resin bead, are formed using nine individual reactions (ignoring deprotection).

The split-pool synthesis procedure allows combinatorial libraries to be created in just a few reaction vessels (Fig. 1.4.). In the first step, a quantity of resin beads is split into multiple, equally sized portions in separate reaction vessels, each of which is coupled with a single building block. After the first reaction step has been carried out, the resin-bound compounds from all reaction vessels are pooled together in one vessel. In this vessel the common steps such as resin washing and deprotection are performed in one batch. By re-portioning the resin-bound compounds into the requisite number of separate reaction vessels, the second solid-phase reaction provides compounds that incorporate all of the possible combinations of the two sets of building blocks. These split and pool operations are repeated until the desired combinatorial library has been assembled. Through this process, each resin bead in a library ends up with (ideally) just one single compound bound to it. Combinatorial libraries resulting from split-pool synthesis are referred to as ‘one-bead-one-compound’ libraries according to Lebl [44–46]. Thus, the library consists of quasi separate compounds as long as they are resin-bound.

Currently, the split-pool synthesis is the most popular method for the synthesis of large combinatorial libraries of compound mixtures comprising thousands to hundred thousands of compounds. A variety of methods have been developed to determine the structure of bioactive compounds within libraries of compound mixtures (see Section 1.4.2.1).

#### 1.3.1.1.1 Techniques using Resin Beads

The split-pool protocol is normally carried out on resin beads. There are limitations when generating mixtures of compounds. Due to the statistical distribution of the solid support at each splitting step, the synthesis will lead to over- and under-representation within the library. In order to ensure that 95 % of all possible compound members of the library are included with a probability greater than 99 % [48, 49], the split-pool synthesis should be carried out with an approximately threefold amount of resin beads (termed 3-fold redundancy). For the commonly used resins (about 100  $\mu\text{m}$  diameter bead), 1 g of the support material corresponds to several million resin beads, so that from a statistical point of view, libraries of the order of  $> 10^5$  different compounds are possible in practice [48–50]. Depending on the loading capacity of the resin bead, quantities of about 200 pmol (0.1 mg compound,  $M_r = 500$ ) can be obtained per resin bead.

Analytical methods for the quality control of combinatorial libraries comprising resin-bound compound mixtures are described in Section 1.4.1.

#### 1.3.1.1.2 Tea-Bag Method

In the so-called tea-bag method, originated in 1984 by Houghten et al. [14] for multiple peptide synthesis, the split-pool protocol occurs batchwise on  $15 \times 22$  mm polypropylene mesh packets with  $\mu\text{m}$ -sized pores known as tea bags, sealed with resin beads for solid-phase synthesis. This method offers the advantage that a greater quantity of each compound of the library is available at once (up to 500  $\mu\text{mol}$ ), which is sufficient for a complete biological and structural characterization. Furthermore, the structural identity of

each synthesized compound can be preserved during the split-pool protocol by labeling the tea-bags.

More recently, Nicolaou and Xiao [51], as well as Moran [52], developed a radiofrequency encoding system. In practice, a glass-encased semiconductor memory microchip capable of receiving, storing and emitting radiofrequency signals [53, 54] is placed in a porous polypropylene capsules along with resin beads. Taking advantage of the radiofrequency-encodable microchips, each of the capsules (tea-bags) can undergo the split-pool protocol in a precise manner, wherein each capsule is radioscanned between solid-phase synthesis rounds and a specific radiofrequency signal is recorded on the memory microchip for each reaction step. Ultimately, the device can be scanned to record the identity of a compound bound on each batch of resin beads. Depending on the size of the capsule and the resin filling quantities of about 7–15 mg compound can be obtained in MicroKanReactors.

### 1.3.1.2 Parallel Synthesis towards Combinatorial Libraries

Combinatorial libraries can also be prepared by parallel synthesis [55]. Here, compounds are synthesized in parallel using ordered arrays of spatially separated reaction vessels adhering to a traditional 'one vessel–one compound' philosophy (Fig. 1.5.). This offers the advantage that each compound, when evaluated for some desired performance, is substantially 'pure' in its local area, provided that the synthesis has proceeded with high efficiency in each stage (see Section 1.4.2.2). Furthermore, in parallel synthesis the defined location of the compound in the array provides the structure of the compound. A commonly used format for parallel synthesis is the 96-well microtiter plate.

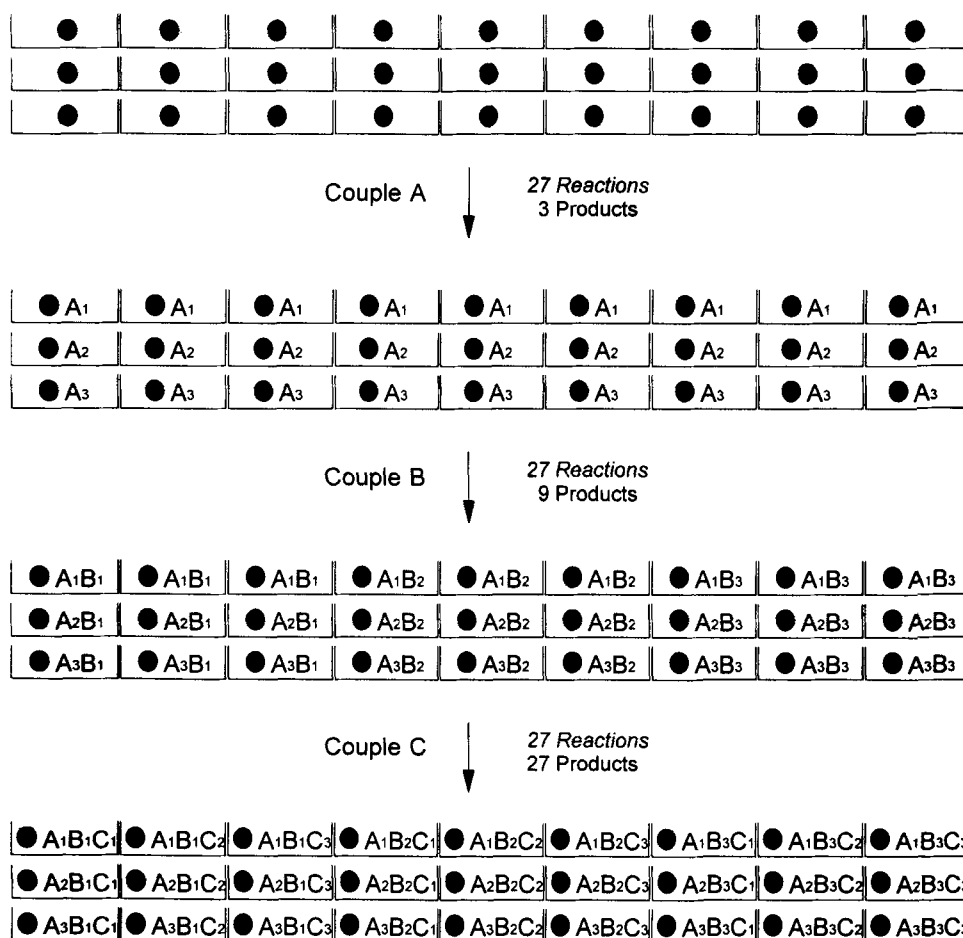
In general, combinatorial libraries comprising hundreds to thousands of compounds are synthesized by parallel synthesis, often in an automated fashion (see Section 1.5).

A number of different solid supports and uniquely designed reaction vessels are adopted for the parallel synthesis of organic compound libraries. The yields of the individual compounds synthesized vary widely from nanomoles to millimoles. Unlike split-pool synthesis, which requires a solid support, parallel synthesis can be done either on solid phase or in solution.

#### 1.3.1.2.1 Reaction Apparatus using Resin Beads

Various reaction apparatus [56] have been developed for parallel synthesis on the basis of a 96-deep-well polypropylene microtiter plate. The resin beads and reaction solutions are placed into each well of the microtiter plate. During the reaction step, the plate is clamped against a Viton gasket in order to seal the hole at the bottom of the well. Between reaction steps the reaction solution can be drained and the resin can then be rinsed by removal of the Viton gasket.

In 1993, DeWitt et al. described the Diversomer apparatus [57–59], which is one of the first reaction apparatus designed exclusively for the parallel synthesis of small molecules. Pipetting machines dispense reagents and solvents into vials that are located in a custom-



**Figure 1.5.** Parallel synthesis to prepare combinatorial libraries of separate single compounds. (Spheres represent resin beads, A, B, C represent the sets of building blocks, borders represent the reaction vessels.) If three building blocks are used in each coupling step, after three stages a total of 27 different compounds are formed, one on each resin bead, with 81 individual reactions being used (ignoring deprotection).

made reaction block with 40 positions. The ends of the gas dispersion tubes – that contain the resin beads – are placed into the vials, allowing the reagents to diffuse into the tubes and contact the solid support. The reaction temperature can be controlled by heating and cooling the reaction block. The apparatus is also enclosed in a manifold with an injectable gasket employed for reagents and solvent additions so that reactions can be maintained under an inert atmosphere. The final compounds are each obtained in quantities larger than 1 mg.

As a reaction vessel for resin beads, an array of chromatography tubes or syringes fitted with frits and stoppers can also be used. These can be gently agitated on a shaker ta-

ble or blood tube rotator during the reaction, and the resulting compounds can be filtered using a vacuum manifold.

#### 1.3.1.2.2 Multipin-Technique

In 1984, Geysen and co-workers introduced an apparatus that allowed individual peptides to be produced in parallel in microplates containing 96 wells [12, 13]. The apparatus uses an array of polyethylene pins that are about 40 mm long and 4 mm in diameter, spaced in such a way that they fit in a conventional 96-deep-well polypropylene microtiter plate. Polyacrylamide- or polystyrene-grafted polypropylene pins [40] are functionalized with a variety of linkers to allow flexibility in solid-phase synthesis. The loading levels per pin can range from about 100 nmol to 50  $\mu$ mol of (ideally) one compound, permitting quantities up to 25 mg compound ( $M_r = 500$ ) to be prepared. This apparatus, marketed as Multipin, has been adopted by Chiron in Australia and applications are preliminary in the area of immunology [60].

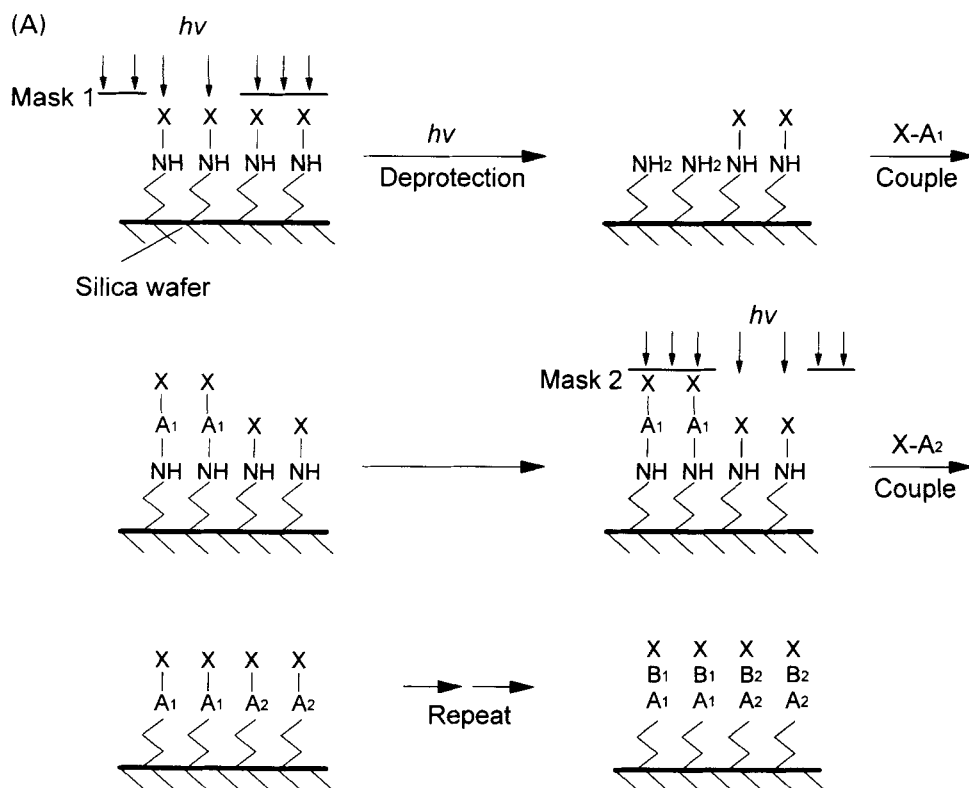
The execution of the parallel synthesis of up to 96 single compounds by the Multipin method involves pipetting the reactants to each well of a 96-well microtiter plate. The pin array is then placed on top of the plate and the resin is allowed to incubate with the reactants to perform the coupling step. The reaction temperature can be raised to 90°C by placing the reaction block into an incubator. Following each reaction step, the pin array is removed and treated in batch to wash the solid support. These operations are repeated until the desired combinatorial synthesis is completed. The resulting compounds can then be removed from the pins into individual wells on a microtiter plate, each of which ideally contains one single compound.

#### 1.3.1.2.3 Spatially Addressable Parallel Synthesis on Silica Wafer

A special case of parallel synthesis is the spatially addressable synthesis pioneered by Fodor et al. [17, 18] in 1991. Here, each member of the library is synthesized at a specific location on a functionalized silica wafer rather than on resin beads in separate reaction vessels. This approach, based initially upon solid-phase peptide synthesis and semiconductor photolithographic techniques by using photolabile amino protecting groups, allows the synthesis of combinatorial libraries containing about 50000 compounds localized to a 50  $\mu$ m square site on a silica wafer ('library on a chip').

The key points of this method are illustrated in Figure 1.6. The photolabile protecting groups on the surface of the silica wafer are selectively cleaved at specified locations by site-specific illumination through a mask. The entire surface is then exposed to the first set of building blocks bearing also the photolabile protecting group. The coupling reaction occurs only at the sites that were addressed by light in the preceding step. The substrate is illuminated through another mask for deprotection and coupling with a second building block. The steps of illumination and coupling are repeated to build up the members of the compound library. This technique which comprises light-directed spatially addressable parallel chemicals is referred to as very large-scale immobilized polymer





**Figure 1.6.** Schematic representation of spatially addressable parallel synthesis on functionalized silica wafer as described in the text (A) and the masking scheme towards combinatorial libraries of a trimer using the generic building blocks A, B, and C (B).

synthesis (VLSIPS). It is mostly applied to oligomeric, peptide and oligonucleotide synthesis.

In the light-directed synthesis, the pattern of masks and the sequence of building blocks define the final structure of the compounds synthesized. The equipment required for light-directed peptide synthesis includes a commercial peptide synthesizer, a flow cell to hold the substrate and deliver the fluids from the synthesizer, a translation stage to move among the masks all located on the same large-format glass plate (which is held against the synthesis substrate in a 'contact' printing mode), and an arc lamp that can deliver collimated long-wavelength UV light to the mask (an instrumentation initially developed for computer microchip construction) [61, 62].

### 1.3.1.3 Reagent Mixture Synthesis towards Combinatorial Libraries

In the so-called reagent mixture synthetic approach towards combinatorial libraries of compound mixtures, each reaction step of the combinatorial synthesis is carried out with

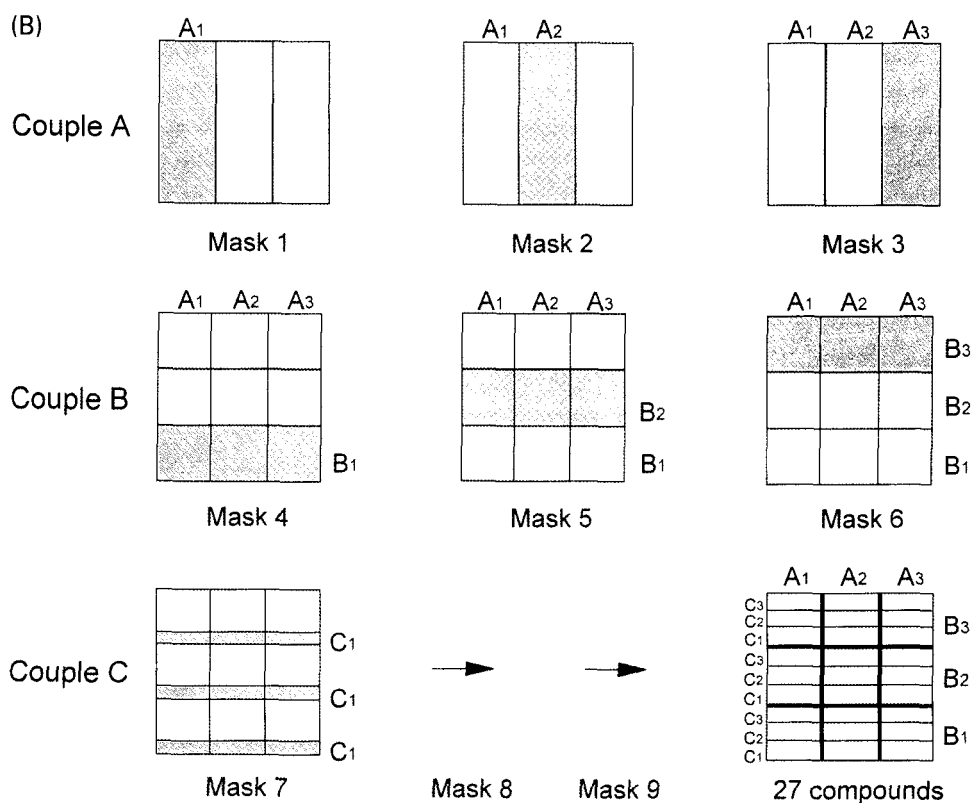


Figure 1.6. Continued

a building block mixture in one reaction vessel either on a solid support or in solution [63]. The use of reagent mixtures requires a thorough knowledge of the mechanism and kinetics involved in the specific reaction being carried out. It is important that the relative reaction rates of the incoming reagents are approximately equal and relatively independent of the resin-bound compounds (*i.e.*, equal nucleophilicity, no significant steric hindrance).

Reasonable results can be obtained for the addition of one set of building blocks. Repetitive cycles using this method can result in large deviations from equimolarity in the final compound mixture, some compounds formed to a lesser extent or not at all.

Combinatorial synthesis towards libraries of compound mixtures can be done either on a solid support or in solution. In both cases an effective decoding strategy is required to extract structural information from the results of the biological assay (see Section 1.4.2.1.2 Deconvolution by Orthogonal Libraries).

### 1.3.2 Synthetic Methodology for Organic Library Construction

In principle, combinatorial synthesis can be performed both in solution and on solid phase. Although chemistry in solution has the advantage of being familiar and well-established as the method of choice in conventional organic synthesis, to date the majority of the compound libraries have been synthesized on solid phases such as resin beads, pins, or chips. This may be attributed to reasons that include easy parallel work-up procedures (filtration); accelerated reactions with higher yields by employing excess of reagents; and amenability to automation. Furthermore, it also facilitates the partitioning of compounds into multiple aliquots in the case of split-pool synthesis.

#### 1.3.2.1 Solid-Phase Organic Synthesis

Solid-phase organic synthesis really began in 1963, when Merrifield [15] used polystyrene resin beads to aid the synthesis of peptides. This was followed in the 1970s by investigations on solid-phase synthesis towards organic compounds by Leznoff, Camps, Fréchet, Rapaport and others [64–67].

Throughout the solid-phase synthesis the compound under construction is covalently attached to a swollen insoluble solid support (usually a resin bead) by a linker that can be cleaved under specific conditions with an appropriate reagent to give the target compound in solution later on (*e.g.*, for assessment of purity, analytical characterization, or biological evaluation). The reactions can be accelerated and driven to completion by using a relatively large molar excess of reagents, resulting in reduced reaction time and higher yields. The support matrix in particular facilitates all steps of a synthesis protocol, such as addition of reagent solutions, agitation, filtration, and washing. Thus, solid-phase synthesis enables full automation, even for multistep synthesis, where the building blocks are added repetitively to build up the desired final compound.

Solid-phase synthesis requires functionalized solid supports and orthogonal chemistry suitable for compound attachment and detachment, and specialized methods for reaction monitoring in a multistep solid-phase synthesis, including the use of attended orthogonal capping strategies for blocking unreacted substrates. Thus, reactions in solid-phase synthesis must be very selective and efficient. Otherwise, the purification of the final products can present a considerable challenge. The scale of solid-phase synthesis is limited and generally restricted by the amount of the solid support and its loading capacity, and the preparation of multimilligram quantities can be cumbersome and expensive for large combinatorial libraries. [Example: For a combinatorial library consisting of 10000 compounds (three-step synthesis at 95% yield/step) to obtain 50 mg of each compound ( $M_r = 500$ ) on a crosslinked polystyrene resin with a typical loading capacity of 1 mmol/g, 1.166 kg solid support is required, *i.e.*, \$ 2500 for solid support or \$ 6 per compound including chemicals, solvents and disposals.]

The synthetic repertoire of solid-phase organic chemistry has been growing rapidly in the past years and has expanded its scope in the complexity of compounds synthesized on solid phase. However, so far only a medium percentage of chemistry in solution has been transferred into reliable solid-supported protocols [64–67]. Two ISIS-based databases

have been established covering the solid-phase synthesis literature: Solid-Phase Organic REaction (SPORE) database (from MDL Information Systems and FIZ Chemie Berlin), and Synopsys database (from Synopsys Scientific Systems together with Oxford Diversity).

In combinatorial synthesis the reactions must operate with reliable yield on a structurally broad set of building blocks to provide a multitude of almost pure final compounds under identical conditions. In the most time- and labor-intensive step, selected building blocks are 'rehearsed' individually through reactions in the solid-phase format, under conditions mimicking those that will be used faithfully in the final combinatorial synthesis. As it will often be impracticable to examine every member of the desired library to confirm its presence, building block combinations that are anticipated to represent worst-case scenarios (*e.g.*, with respect to steric and/or electronic factors) are studied and optimized, with problematic building blocks being excluded from the library construction.

In contrast, using an optimized synthesis sequence the combinatorial libraries can be constructed comparatively rapidly and efficiently as long as the building blocks are commercially available or readily accessible.

### 1.3.2.2 Synthesis in Solution and Liquid-Phase Synthesis

In the literature, few examples exist for combinatorial synthesis in solution [68], with most being one- or two-step parallel synthesis of individual compounds or 'reagent-mixture' synthesis of compound mixtures. Pools of dimeric compounds including esters, amides and carbamates have been successfully prepared without needing further purification. One example of a reaction which is particularly applicable to combinatorial synthesis in solution is a multicomponent reaction, such as the Ugi four-component reaction, as it can produce four points of diversity in a single step [69–71].

The purification problem in the solution-phase synthesis of combinatorial libraries is addressed in various ways. For example, for a multistep synthesis polymer-bound reagents could be used [72–75]. A complementary approach uses scavenger, such as a solid-phase-supported electrophile or nucleophile, to remove unreacted starting material after the synthesis has taken place [76–79].

An extension of the combinatorial synthesis in solution is achieved by the use of soluble polymeric supports [80, 81], which combines some of the advantages of chemistry in solution and on solid phase. The so-called *liquid-phase combinatorial synthesis* is based on the physical properties of poly(ethylene glycol) monomethyl ether. The polymer is soluble in a variety of aqueous and organic solvents, which allows reactions to be conducted in homogeneous phase whereas the propensity to crystallize in appropriate solvents facilitates the isolation and purification of the compound at each step of the combinatorial synthesis.

Another variation uses polyfluorinated compounds as the starting material for synthesis. In this technique, known as fluororous synthesis, the starting material is first derivatized with perfluorinated groups [82]. After the synthesis, a water-immiscible fluorocarbon solvent is added to the solution, effectively separating the product from the reaction mixture. The perfluorinated groups can then be removed to recover the purified product.

## 1.4 Characterization of Combinatorial Libraries

The combinatorial synthesis is followed by the analytical characterization and the biological evaluation of the compound library. Essential elements are the bioassays employed for high-throughput screening and the methods available for structural elucidation of compounds that emerge from the biological evaluation of the compound library, termed as hits. The strategies used depend on the library format – separate single compounds or mixture of compounds, resin-bound compounds or compounds in solution.

### 1.4.1 Analytical Characterization

Since in a combinatorial synthesis the amount of compounds bound to a single resin bead is usually of the order of several hundred picomoles, the analytical methods for direct structural determination must be highly sensitive [83]. Efficient methods exist for the structural elucidation of peptides and oligonucleotides in the form of automatic microsequencing procedures (*e.g.*, based on Edman degradation and DNA sequencing) [84]. For minute amounts of low-molecular mass organic compounds, mass spectroscopic methods are the method of choice [85, 86], despite recent progress in the field of IR [87, 88] and NMR [89–92] spectroscopy. The sensitivity of modern mass spectrometry allows the structural determination of compounds in the high femtomolar range. Especially, electrospray mass spectrometry (ES/MS) or matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF/MS) techniques are used for compound analysis from small resin samples or single beads.

#### 1.4.1.1 Analytical Characterization of Compound Mixtures

If combinatorial libraries consist of compound mixtures (*e.g.*, after cleavage from the resin beads), a reliable analytical characterization is difficult. In this case the utilization of mass spectrometry for analysis is based on the prediction of mass distribution of the library. Computer-generated distribution profiles can be compared with the actual profile obtained from the compound library [93, 94]. Evaluation of mass distribution detects synthetic problems based on incomplete coupling (shift toward lower molecular masses), incomplete deprotection, or unwanted library modification, such as oxidation, acylation, or alkylation (shift toward higher molecular masses). However, for a larger library of compound mixtures many different compounds will have the same molecular mass. This greatly complicates structural determination and even makes it impossible at a certain mixture complexity.

For library analysis of compound mixtures, reliable results are obtained by coupling mass spectrometry with HPLC or capillary electrophoresis (CE). In this arrangement, mixtures of several hundred compounds can be analyzed. A comparison of MS, CE, and NMR for library characterization has been made in [93]. It was concluded that existing analytical techniques can provide sufficient information about prepared libraries.

In addition to the direct methods of structural analysis of libraries of compound mixtures, there are also indirect methods, which are particularly advantageous when working with larger compound libraries made by the split-pool synthesis (see Section 1.4.2.1).

#### 1.4.1.2 Analytical Characterization of Single Substances

The analysis of separately synthesized single substances is much simpler than that of compound mixtures. Here, the yield and purity can be determined in the traditional way: purification by extraction, crystallization, or chromatography, gravimetric determination of the yield, and confirmation of the purity by elemental analysis, chromatography, or NMR.

### 1.4.2 Hit Identification in Combinatorial Libraries by High-Throughput Screening

Organic compounds are usually screened in solution in a 96-well microtiter plate format. Therefore, many bioassays are available, *e.g.*, competitive receptor binding assays with radiolabeled ligands, functional assays, or cell-based assays. All these 'solution-phase' bioassays, in principle, can be adapted to high-throughput screening of combinatorial libraries. Besides the growing number of compounds to be tested, the costs for a single bioassay (\$5 per well) is the main reason for the current trend to miniaturize the assay format towards 384- or 1536-well microtiter plates. In an ultra-high-throughput screening it should be possible to test more than 100 000 compounds per day.

In the case of combinatorial libraries comprising mixtures of solid-bound compounds the compounds can be screened still covalently attached to the solid support ('on-bead' screening, see Section 1.4.2.1.1) and, after cleavage of the solid support, in solution as compound mixture or as single compounds, when the beads are spatially separated (see Sections 1.4.2.1.2–1.4.2.1.4).

The trend in bioassays is clearly towards examining the test compounds in solution. The results are not influenced by undesired interaction with the solid support. Experiences with 'on-bead' assays were sometimes disappointing because even compounds that were known to be biologically active in solution did not bind to a detectable extent to the biological receptor when attached to the bead. Moreover, assaying compounds in solution is probably more relevant from a pharmacological standpoint. In addition, screening on solid support is not applicable to intracellular targets, when a whole-cell assay is used.

#### 1.4.2.1 Strategies for Libraries of Compound Mixtures

##### 1.4.2.1.1 On-Bead Screening

Combinatorial libraries can be screened, while the compounds are still attached to the resin beads or onto a surface [17, 18, 95–97]. Thus, the solid support and its linkers must

be soluble in water and for quantitative results (*e.g.* structure–activity relationships) the beads must be uniform in both size and substitution.

The solid-bound library is treated with a labeled soluble biological target. For many studies a fluorescent label has been employed because of the high sensitivity of fluorescence detection. The labeled receptor binds to those resin beads that are derivatized with compounds that have the highest affinity to the biological receptor. The labeled beads are then selected followed by structural elucidation of the support-bound compound (see Section 1.4.1). The identity of the bioactive substance can be limited to a few alternative structures by mass spectrometric determination of the molecular mass. Highly efficient, automated methods have been developed to isolate the labeled beads, for example, by use of a fluorescence-activated cell sorting instrument [96].

This on-bead screening is particularly useful for libraries of several thousand to a million compounds and the isolation of a few bioactive compounds from many inactive ones. By using an on-bead assay, the screening of a library of  $10^7$  resin-bound compounds can be accomplished routinely by one person in one day. One advantage of this method is that, once the library has been prepared and assayed, the remaining compounds may be reused for different biological assays. However, all of the difficulties of screening on the solid support remain.

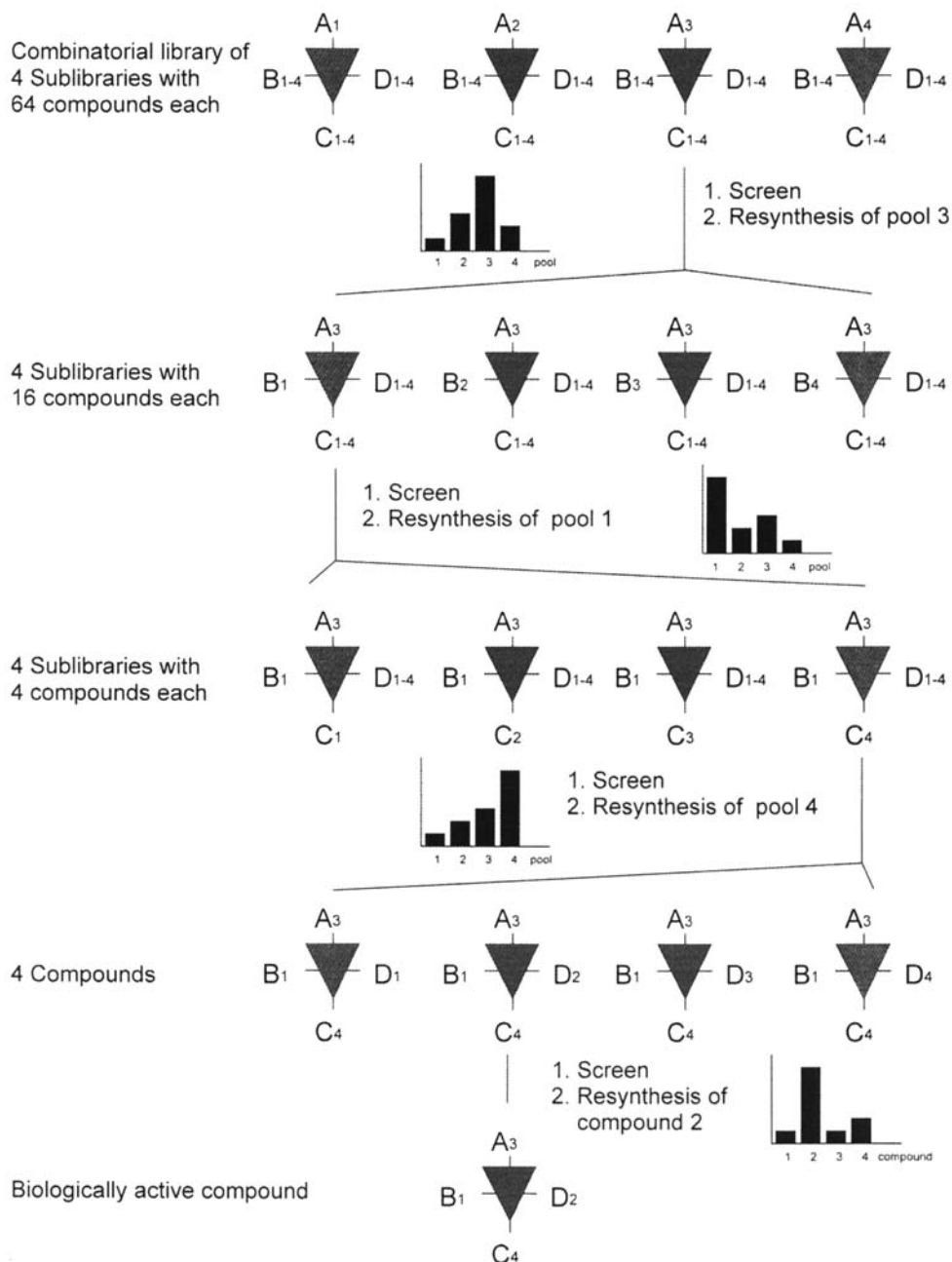
#### 1.4.2.1.2 Deconvolution

Three deconvolution strategies can be used for the structural determination of the biologically most active compounds in a combinatorial library comprising mixtures of up to several thousand compounds: (1) iterative deconvolution [47, 98]; (2) deconvolution by positional scanning [99–101]; or (3) deconvolution by orthogonal libraries [102–104].

##### Iterative Deconvolution

The iterative deconvolution method [47, 98] involves preparation of a series of spatially separated sublibraries comprising compound mixtures (termed pools) in which the identity of the building block at least at one position is known and at the remaining positions all combinations of building blocks are incorporated. This can be achieved, *e.g.*, by omitting the final pooling step of the split-pool synthesis. Each pool is screened and the compound mixture with the highest biological activity indicates the importance of the building block at that defined position. On the basis of this result the split-pool synthesis is repeated with the selected building block at the initial defined position in order to prepare pools where the next defined position is introduced. Each pool is evaluated for biological activity in order to select the optimal building block at the additional defined position. The remaining positions are then identified sequentially through an iterative procedure of (re)synthesis and screening, as illustrated schematically in Figure 1.7, with a generic scaffold.

The iterative deconvolution procedure has been applied successfully to determine the biologically active structure directly, but several issues should be considered. Firstly, as



**Figure 1.7.** Iterative deconvolution procedure illustrated by a combinatorial library consisting of a generic scaffold containing four variable building blocks A, B, C and D and in total 256 compounds.



the number of compounds in a pool increases, lower concentrations of each compound must often be used in order to maintain the solubility of all the compounds in the pool. Compounds that have modest activity therefore may be not detected. Secondly, the biological activity observed for a given compound mixture depends on both the activity and abundance of the active compounds in each pool, *i.e.*, biological activity is generally a result of the sum of more than one biologically active compound. Thus, the compound mixture that shows the greatest biological activity does not necessarily contain the most potent compound(s) [105–108]. Thirdly, the iterative resynthesis and biological evaluation can be a time-consuming and laborious process.

A number of modifications have been reported in order to address these complications. Janda has noted that during the first split-pool procedure, resin can be saved at each cycle immediately prior to resin pooling. This resin can later be used as an intermediate in the iterative resynthesis and deconvolution procedure (recursive deconvolution) [109]. This results in considerable savings in time and effort.

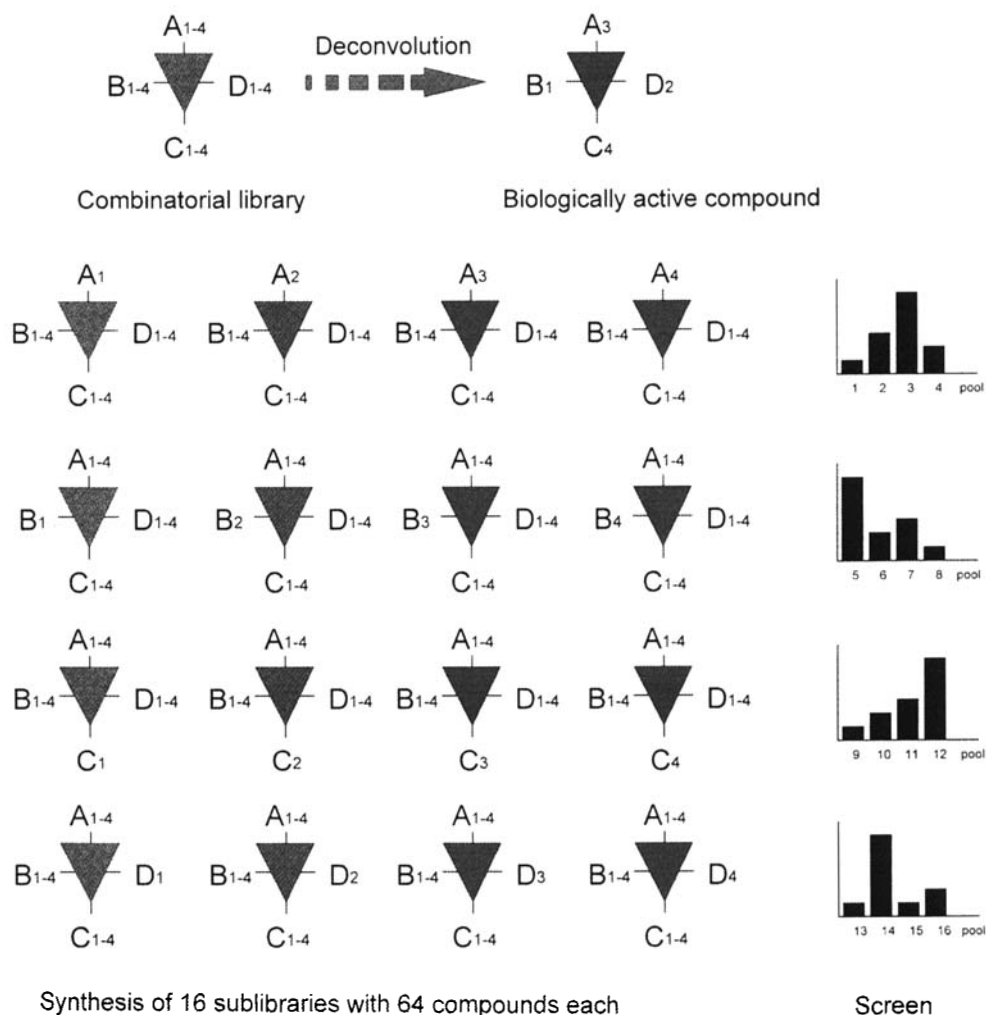
### Deconvolution by Positional Scanning

In the positional scanning approach for deconvolution [99–101], originated by Houghten and coworkers for peptide libraries, separate sublibraries are prepared, each of which contains a single defined building block at one position and a mixture of building block at the others. The number of sublibraries is equal to the number of variable positions in the substitution pattern. This method resorted to a hybrid of split-pool and parallel synthesis. Each positional sublibrary is then screened to directly determine the building block at each defined position that contributes most to biological activity (see Fig. 1.8.).

By combining the positive screening results from all series of sublibraries the whole structure of a highly biologically active compound can be deduced directly. In theory, if only one compound were active in the library, activity corresponding to that compound would be found in the one mixture of each sublibrary containing that compound. In reality, the same result is seen, but the activity generally results from the sum of more than one active compound. Anomalous results are seen if the activity is due to the sum of many weakly active compounds.

However, if in one position there is no clear preference for a definitive substituent, all combinations of the preferred substituent must be synthesized in order to find the most active compound. Although active compounds are identified without iterative synthesis using positional scanning, in comparison with iterative deconvolution there is an increased likelihood that the most potent compound(s) will not be identified [110, 111].

The advantage of positional scanning over iterative deconvolution is that the sublibrary syntheses are carried out at once. But the number of split and pool operations that would be required for library synthesis is rather high so that building block mixtures are used in several of the library synthesis steps in order to reduce the number of operations. A thorough examination of the theoretical and experimental aspects of iterative and positional scanning deconvolution has been published [110].

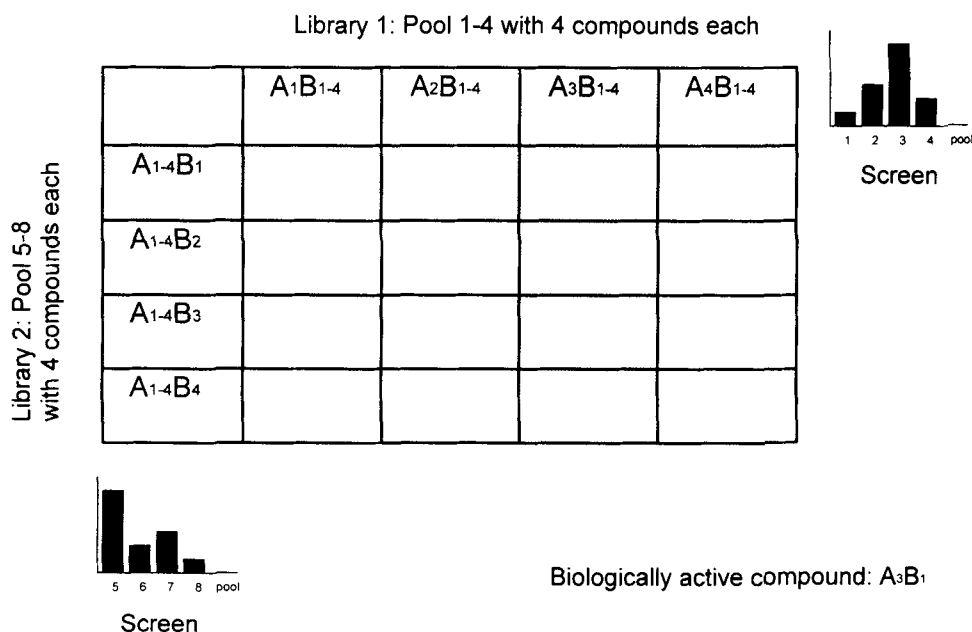


**Figure 1.8.** Deconvolution by positional scanning procedure as described in the text.

### Deconvolution by Orthogonal Libraries

A method similar to the deconvolution by positional scanning was developed independently by Pirrung [102, 103] and Déprez [104]. This method has been termed *orthogonal library approach* or *indexed library approach* by other authors [112]. With respect to deducing structural information from the screening results this approach involves preparing two series of orthogonal sublibraries consisting of compound mixtures (see Fig. 1.9).

The combinatorial synthesis of orthogonal libraries involves the synthesis of two series of sublibraries. In the first series, each of the building blocks  $A_n$  reacts separately with a



**Figure 1.9.** Schematic representation of an orthogonal library synthesized from each of four components A and B. Each column and each row represents a compound mixture consisting of four compounds (for example A<sub>1</sub>B<sub>1</sub> to A<sub>1</sub>B<sub>4</sub>). The single compound responsible for the biological activity in the screening is derived from the correlation of two most active mixtures (shaded rows). In this case, A<sub>3</sub>B<sub>1</sub>.

stoichiometric quantity of an equimolar mixture of the building blocks B<sub>1</sub>–B<sub>x</sub> and in the second series of sublibraries, each building block B<sub>m</sub> reacts separately with an equimolar mixture of the building blocks A<sub>1</sub>–A<sub>x</sub>. This synthetic approach provides compound mixtures, in which the identity of either the building block A or the building block B is fixed. This synthetic procedure results in each compound being formed exactly twice, and every pair of mixture has exactly one compound in common.

A positive result for any given sublibrary in a biological assay identifies the fixed part of an active compound. By combining the positive results from both series of sublibraries the whole structure of an active component can be deduced directly. Moreover, this form of internal control allows the recognition of false positives in the biological evaluation process. Assignment of activity to a certain structure is possible by simple comparison, that is, without time-consuming deconvolution or new synthesis. However, as in the case of the positional scanning method, it seems to be necessary to employ building block mixtures in each of the synthesis steps in order to reduce the number of split-pool operations.

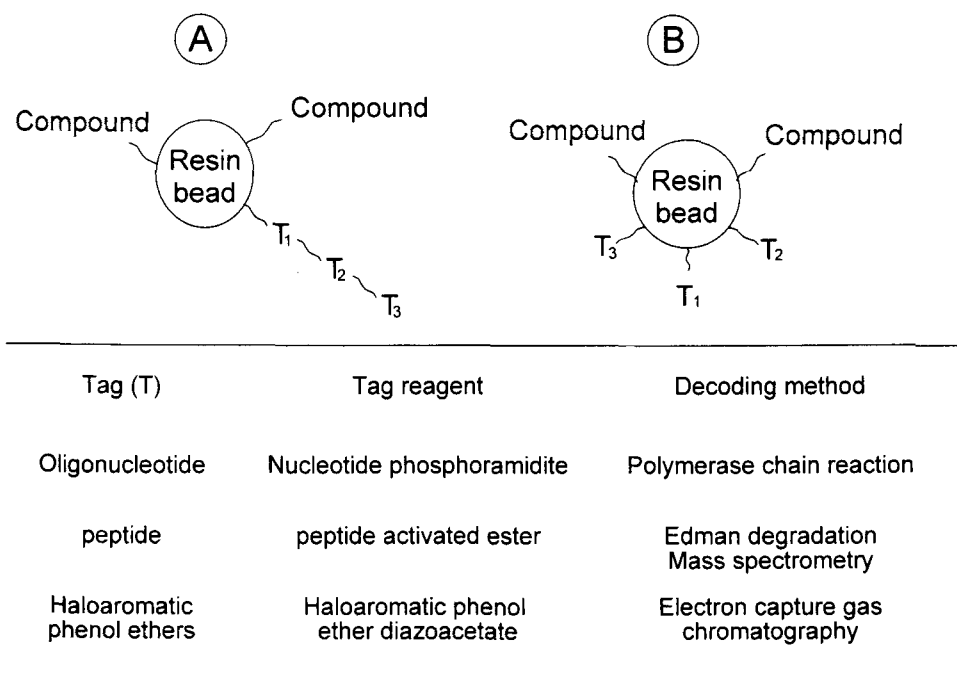
## 1.4.2.1.3 Encoding

An alternative strategy with respect to structural determination of active compounds emerging from the biological evaluation of compound mixtures is to label the compound during the split-pool synthesis (encoding or tagging) [9, 113–117]. Here, instead of the structure of the biologically active compound the corresponding more accessible (chemical) code is determined.

## Chemical Encoding Strategies

In the chemical encoding strategy in general, resin linkers containing two orthogonally protected functional groups are used to allow the concurrent synthesis of both the compound of interest and the encoding compounds (so-called tags) on the bead, which upon cleavage are sequenced or otherwise decoded to determine the structure of the compound of interest.

The tags may be added to provide an encoding sequence whereby the structure of the tag encodes for the building block and the location in the sequence encodes for the library synthesis step (oligonucleotide [9, 113–117] and peptide [118–122] as sequenceable tags; Fig. 1.10.A). The advantage of oligonucleotides and peptides as tags lies in the sensi-



**Figure 1.10.** Encoded combinatorial libraries by (A) sequenceable chemical tags (oligonucleotides and peptides) or by (B) non-sequenceable chemical tags (haloaromatic phenol ethers).

tivity and possibility of sequencing tens or hundreds of samples of nucleic acid codes in parallel. However, its application for coding organic libraries is rather limited due to the tendency of DNA and peptide tags to break down under the often very rough conditions of organic synthesis.

Alternatively, each tag may be added individually to the resin [123–129]. In this case, the tags must encode for both the building block structure and the step in the library synthesis (haloaromatics as non sequenceable tags; Fig. 1.10.B) [123–127].

For each tagging approach, the orthogonal chemistry between the tags and compounds of interest has to be determined. Those compounds are chosen as encoding compounds that can be easily identified by well-established highly sensitive analytical microtechniques, even at very low concentrations.

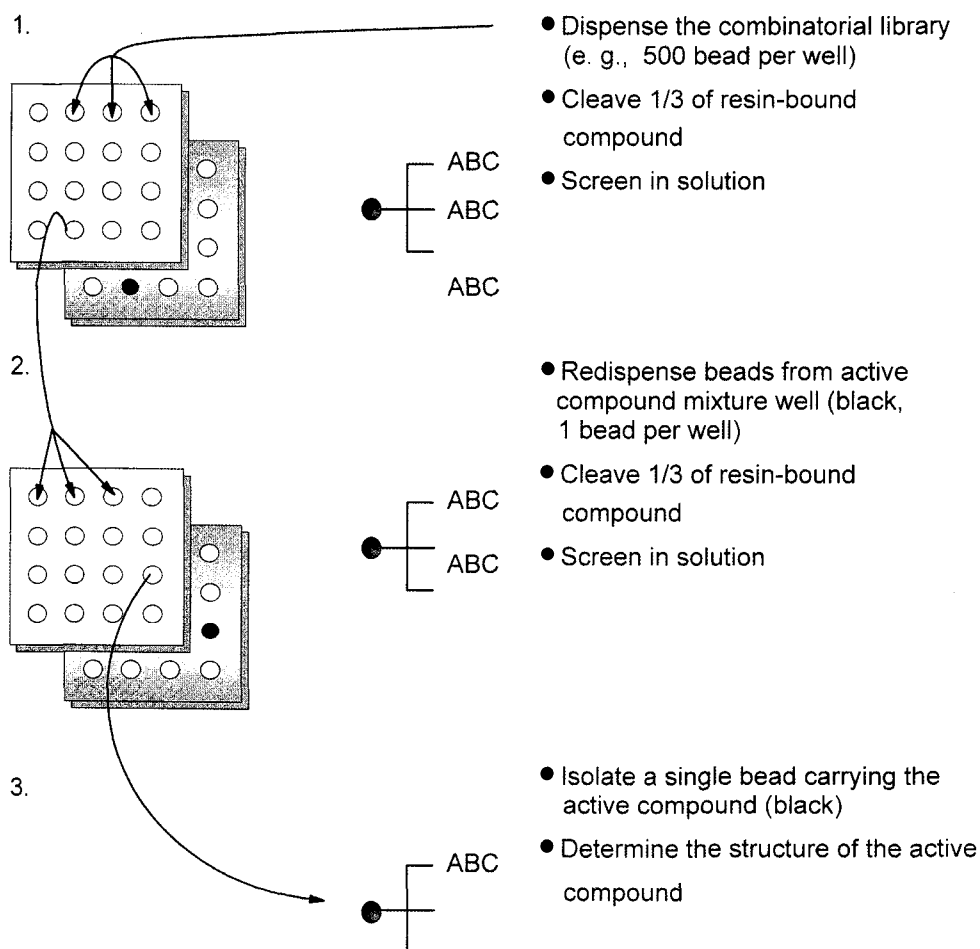
In all cases, such a coding strategy is, of course, valid only if the relationship between support, compound and the code molecules remains intact. This provides no problem if biological testing is carried out with the support-bound compound. After testing and cleavage the active compound is identified by reading the code. However, the assay is usually carried out with dissolved compounds. The synthesized compound can then be selectively released from the support for the biological assay in solution (see Section 1.4.2.1.4), while the encoding strand is retained on the bead for later structural elucidation. This procedure requires that the solution containing the compound be spatially addressable to the solid support from which it originated in order to allow decoding.

Non-Chemical Encoding Strategies (see Section 1.3.1.1.2)

#### 1.4.2.1.4 Multiple Cleavable Linkers

Lebl [121, 130–134] and Baldwin [126, 127] first reported an efficient strategy, in which the advantages of performing assays in solution and structural determination of resin-bound compounds are combined. This method is based on multiple cleavable linkers which can release aliquots of the resin-bound compound, rather than cleaving all the compound bound to one resin bead at once.

The compound library is divided into mixtures of up to several hundred compounds (Fig. 1.11.). In the first step, about one-third of each respective compound bound on a macrobead (about 600–800 pmol compound per bead) is split off into solution for biological evaluation. Experiments performed indicate that an optimal mixture complexity is about 20 compounds per mixture [135]. This can be accomplished, *e.g.*, by the use of a three-armed linker that selectively releases compounds at different pH optima [132, 134] or of a photochemically cleavable linker which releases compounds into solution by controlled irradiation [126, 136]. In a second step, the collection of beads corresponding to the greatest biological activity is then redistributed in smaller mixtures. A further aliquot of the respective compound is then released and assayed. Ideally, the beads are rearranged separately for direct identification of the compound responsible for biological activity ('single-bead assay'). The structure of the bioactive compound can then be determined from the remaining resin-bound compound (often in only femtomole quantity) by the analytical methods described in Section 1.4.1.



**Figure 1.11.** Screening scheme of combinatorial libraries by the use of multiple cleavable linkers: Two-stage screening and then identification of the biologically active compound.

### 1.4.2.2 Strategies for Libraries of Separate Single Compounds

If combinatorial libraries of separate single compounds are prepared by parallel synthesis, then elucidating the structure of biologically active compounds does not pose a problem, because the site of synthesis is clearly associated with the structure being synthesized (see Section 1.3.1.2), given that the desired product was built up during synthesis.

One of the major advantages of parallel synthesis is the possibility to compare the bioactivity of the compounds within the library directly. Thus, the establishment of structure–activity relationships is possible, which is especially important during the drug optimization process. However, the screening capacity may still be a limiting factor, especial-

ly during lead optimization (where quite often assays are not automated), or if multiple measurements for dose–response curves are required.

## 1.5 Automation and Data Processing

### 1.5.1 Synthesis Automation and Data Processing

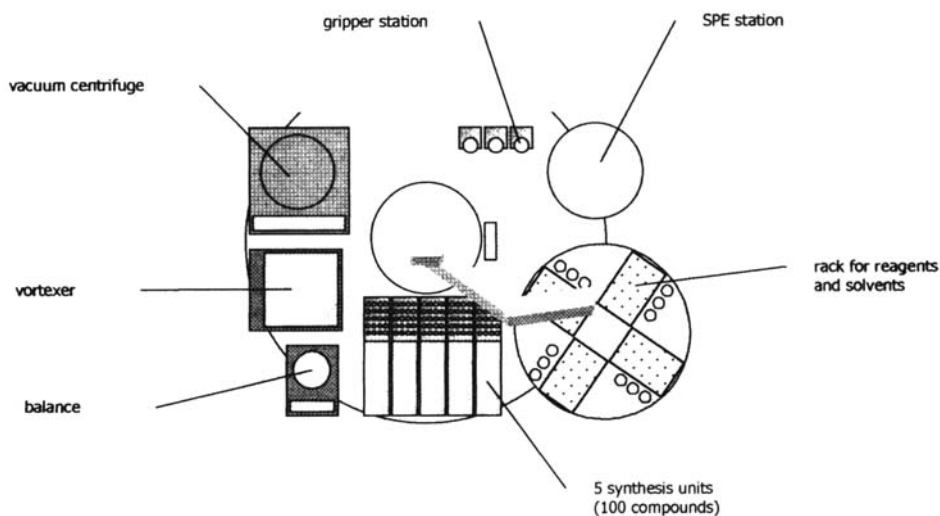
When the size of a combinatorial library reaches thousands of separate single compounds or compound mixtures, manual synthesis and testing against a biological target will hardly be manageable in an acceptable time frame. Therefore, the advent of combinatorial chemistry for the high-throughput synthesis of compounds has driven the advancement of automated methods for synthesis [137, 138] as well as requisite pre- and post-synthesis operations, *i.e.*, resin loading, reagent and resin delivery, compound isolation, purification, and analysis.

Ideally, a fully automated process exists from the combinatorial parallel synthesis to compound isolation, delivering the compounds in a standard 96-well microtiter plate for the following biological high-throughput screening.

Most of the proprietary marketed automated synthesizers derive from instrumentation suitable for multiple peptide synthesis and from pipetting robots. These synthesizers preclude reaction conditions necessary for some organic transformations – especially carbon–carbon coupling reactions using air- or moisture-sensitive reagents that require inert atmosphere – and therefore introduce constraints upon library synthesis. That is why reactor blocks have been elaborated incorporating features most important to synthetic organic chemistry – both solid–liquid-phase synthesis, or synthesis in solution. These features include temperatures between  $-78^{\circ}\text{C}$  and  $150^{\circ}\text{C}$  and inert atmosphere, as well as using a wide variety of organic solvents. Furthermore, reactors for solid-phase synthesis require some mechanism for filtration and washing of the resin-beads (for representative reactors, see Section 1.3.1.2.1).

The architecture of an automated synthesis system can be modular or integrated, open or closed. In a *modular computer-controlled robot system*, all liquid transfers (addition of solvents or reagents and washing solutions) take place in a central module constructed around a robot (Fig. 1.12.). The charged reaction vessels are then transferred to special incubation modules, where they remain for the duration of the reaction, while the central module is again available for further operations. In the *closed (non robotic) concept*, reagent and solvent are permanently interconnected with the reaction vessels by means of tubes and valves to insure an inert, anhydrous environment. Automation in high-throughput synthesis is still being developed by instrument and robotic companies as well as pharmaceutical companies. Examples are ACT496(MOS) (AdvancedChemTech), ARCoSyn (accelab Laborautomations-GmbH), Genesis (Tecan), Nautilus(TM)2400 (Argonaut Technologies, Inc.) [139], RAM(TM)Synthesizer (Bohdan Automation, Inc.), Sophas(TM3) (Zinsser Analytic), Syro(TM) (MultiSynTec), and Zymark robotic workstation (Zymark, Inc.).

Automated combinatorial chemistry will demand computer control of processing instrumentation. A *serial processing system* performs all operations on one sample before



**Figure 1.12.** Schematic representation of a modular computer-controlled robot system (Source: accelab Laborautomations GmbH, Tübingen).

proceeding to the next. On the other hand, automated synthesizers that process samples *in parallel* perform the same operation on multiple samples before proceeding to the next operation. Parallel processing significantly improves the throughput and efficiency of automated synthesis systems. Furthermore, efficient software tools are necessary to program the synthesis run, to retain records of the synthetic operations or the biological testing, and to handle the huge amount of compounds and corresponding data with respect to compound searches or data interpretation [140].

Current automated instrumentation for high-throughput syntheses are still below the capacity of the well-established high-throughput screening of compounds. Some instruments can produce 1000 to 2000 compounds per day by parallel synthesis. The time it will take to achieve a robotic synthesis capability that is equal to screening throughput is not yet known. By performing biological binding experiments in microtiter plates and employing both special pooling strategies and automation, nowadays 10 000–50 000 compounds can be screened against a given biological receptor in a single day, with a good level of reliability (see also Section 1.4.2).

### 1.5.2 Automated Purification

In principle, combinatorial synthesis does not create analytically pure compounds, even though the applied chemistry might be optimized. With respect to structure–activity relationships, the compounds evaluated in lead optimization should be as pure as possible. In high-throughput screening for lead finding, the compounds are usually tested



without purification. If the purity becomes a severe problem, it must be considered that all compounds be purified before screening, or that only the bioactive molecules be purified.

If *HPLC purification* is necessary, a one-column system can purify about 100 samples in milligram scale in one day. In this case, it is advantageous to use a LC-MS system where the fraction collector is driven by the mass spectroscopic information, because all fractions are already analyzed, and ideally a reformatting step in the 96-well format is not required. Commercially available HPLC-systems using more than one column in parallel are few in number. One system from Biotage consists of four columns driven by one pump and an individual fraction collector for each column. However, the system is not connected online to a mass spectrometer, and all fractions have to be analyzed offline by mass spectrometry. This may cause difficulties in sample-tracking and reformatting into the 96-well format, which should be controlled by appropriate software.

*Solid-phase extraction* might be an alternative to HPLC separations, especially when large libraries are synthesized for high-throughput screening. Many of the modules for solid-phase extraction can handle 96 samples in an microtiter plate array in parallel, either manually or automatically, depending on the system. Solid-phase extraction is suitable to eliminate very polar or unpolar byproducts or salts. Examples of equipment include ASPEC XL4 (Gilson and Abimed Analysen-Technik), Baker spe\*12G System (J.T. Baker), Genesis SPE (Tecan), Microlute (Zinsser Analytic), Roliman-96 (EPR Labautomation), and VacMaster (ict).

## 1.6 Library Design and Diversity Assessment

Combinatorial libraries serve two distinct functions in the drug discovery process, namely lead identification and lead optimization (see Fig. 1.1.), which dictate their size and composition.

A compound library suitable for the discovery of a novel lead, where *a priori* no specific structure or substructure is obligatory as a starting point, will initially consist of a large number of compounds within a broad structural range (so-called 'generic', 'unbiased', or 'random' library). Thus, the chance to identify a compound of significant affinity for the novel biological target or an alternative lead for a known biological target will be increased in random screening. On the other hand, in the case of lead optimization the compound library will be invariably of a more limited size with close structural resemblance to the original lead structure (so-called 'focused', 'biased', or 'directed' library). Most drug discovery development is performed through a series of optimization cycles within a targeted screening in order to meet a set of predetermined criteria for a drug candidate.

Thus, the library design (compound selection for combinatorial synthesis) as a central element in combinatorial chemistry depends markedly on the intended use of the library (Table 1.1.). In the case of *generic libraries*, the best overall strategy for library design is rather blurred, and a variety of approaches are under investigation. One approach favors using structural motifs, which have distinguished themselves by appearing frequently in previous potent drugs or drug candidates, so-called 'privileged' structures. One of the first examples of this type was provided by Ellmann and co-workers [141–144] in the prepa-

**Table 1.1.** Library design in the case of lead finding and lead optimization.

	Lead finding	Lead optimization
Library type	generic	focused
Library size	huge ( $\geq 10000$ compounds, < 1 mg per compound)	modest ( $\leq 10000$ compounds, > 1 mg per compound)
Structural goal	no special initial	specific (lead oriented)
Structural diversity	broadest	relatively narrow
Building blocks	any diverse	specific according to retrosynthesis
Synthetic strategy	flexible	well defined
Screening goal	many biological targets	one biological target or target class

ration of libraries of benzodiazepines, an important class of heterocycles that have found widespread use in the pharmaceutical industry. Another approach is to incorporate key recognition elements for receptor binding (pharmacophoric patterns) that are relevant to the particular biological target under investigation [145]. A third approach is to select stable compounds upon which only a few or no therapeutic agents or candidates have been based, but which are straightforward to prepare with multiple sites available for the display of different functionalities.

### 1.6.1 Diversity Assessment for Selection of Building Blocks or Compounds

The synthetic effort towards combinatorial libraries of several hundred thousand compounds is considerable. Therefore, computer-based techniques were introduced to reflect and analyze compound libraries to enhance the composition, while constraining their size.

Several groups have validated physico-chemical or structural descriptors to quantify the similarity or so-called molecular diversity of compounds and compound libraries [146–150]. Representative approaches are based on physico-chemical descriptors, *i.e.*, on  $\log P$ , where  $P$  is the octanol–water partition coefficient, or on molar refractivity, 2D descriptors (2D fingerprints, atom pairs) or 3D descriptors in connection with the 3D model of the biological target (CoMFA, 3D fingerprints). Examples of diversity profiling using 2D versus 3D descriptors are given in [151, 152]. One of the basic flaws in the concept of a diversity parameter is the lack of evidence that it correlates directly with biological activity of the compounds.

The computer-based library design process includes modules for the selection of the building block sets, the creation of the corresponding virtual compound library, and the selection of the compounds to be synthesized and screened. The selection procedure for the generation of a generic library with a broad range of structurally unrelated compounds by comparing the building blocks or the members of the virtual library consists of calculating descriptors for the compounds, weighting them according to an algorithm, and

then calculating a similarity coefficient (*e.g.*, Tanimoto coefficient). Part of the selection process is to determine the availability and cost of reagents used as building blocks. The members of a building block set should reflect a broad range of physico-chemical properties, functionalities, charges, and conformations. Databases of chemicals are available for this purpose, *e.g.*, the Available Chemicals Directory by MDL Information Systems, a collection of chemical products offered by more than 200 chemical suppliers and totaling well over 400 000 compounds.

Modules in software programs for diversity analysis are for example Selector (Tripos), Jarpat (Daylight Chemical Information Systems), C<sup>2</sup>-Diversity (Molecular Simulation), and Chem-X (Chemical Design). In the context of management of compound libraries software programs such as Project Library, Legion, or UNITY (all by Tripos) are available.

### 1.6.2 Iterative Optimization Methods

The interpretation of the biological data resulting from library screening drives the new (secondary, tertiary, etc.) library design, creating an iterative cycle of combinatorial library synthesis and biological evaluation. One challenge in lead optimization is to capture the data and to build structure–activity relationship (SAR) and quantitative structure–activity relationship (QSAR) models. The power of combinatorial chemistry, and in this particular case the efficacy of (automated) parallel synthesis, is to produce hundreds of analogs in a short time frame. This will contribute to the faster development of a drug candidate with the desired biological activity. Furthermore, other main tools of traditional drug development including structure- and mechanism-based drug design will be integrated in the library design.

An alternative computer-based approach minimizing the number of compounds to be synthesized is the iterative use of genetic algorithms [153, 154] in library design. Here, in order to find biologically active compounds from a virtual library of a hundred thousand compounds, only a small number of randomly selected compounds are synthesized, followed by biological evaluation. The most active compounds are selected, which formed the parent generation of computation and synthesis of further compounds. As given in [154], this process is carried out with 20 compounds for each of the 20th generations, so that in total 400 (0.25% of the possible 160 000) reaction products were synthesized and tested. The compounds of the 20th generation were 2000 times more active against the biological target (thrombin inhibitor) than were those of the first generation.

## 1.7 Economic Aspects

The pressure on pharmaceutical companies to produce new drugs is intense. It is estimated that, to remain competitive, a top-ten drug company must quadruple the number of new drugs it launches each year [155]. However, the discovery of new lead compounds and the subsequent development process is extremely high risk and expensive. Obvious-

ly, any technique that reduces the time, cost or risk involved in lead development is useful, and the recent application of combinatorial chemistry and high-throughput screening to drug discovery looks to become a key technology in the mid 1990s. The ongoing introduction of combinatorial chemistry seems to result in a fundamental change in the way the drug discovery process for new or improved therapeutics and agrochemicals is carried out. The importance of this technology is also indicated by several alliances between major pharmaceutical companies and small 'Combi Chem' companies and library suppliers. From 1995 to 1997 alone, more than 60 of these agreements were completed, which typically included upfront payment, milestones, and substantial royalties on the sale of commercialized products [155–158].

The industry now consists of four major types of companies: library suppliers, which supply the chemical building blocks and reagents for library-based screening programs; library value-adders, which are high-throughput automated screening companies; commodity users, which provide targets for high-throughput screening and library generation; and hardware and/or software developers, dedicated to developing instrumentation and information management for library generation and screening, as well as data interpretation and handling.

Combinatorial chemistry can contribute to drug discovery on the one hand by improving the finding of biologically active compounds in combination with high-throughput screening, and on the other hand by the acceleration of lead optimization, resulting in an active drug candidate with a desired pharmacological profile (see Fig. 1.1.).

Historically, the pharmaceutical compound collections range from several tens of thousand to several hundreds of thousands of compounds per company, depending on the size and history of the company. These compounds have long been regarded as the 'crown jewels' of their respective R&D organizations. Assuming one newly synthesized compound per researcher per week and using \$ 250 000 as the fully burdened cost per researcher per year, a 100 000-compound library approximates a \$500 million investment in compound synthesis, patenting, and potential screening against biological targets [156]. In this context, combinatorial chemistry raises the opportunity to build up or amplify the number of compounds for high-throughput screening in a shorter period of time. Combinatorial chemistry companies such as Pharmacoopia report the generation of 100 000 compounds per calendar quarter, and the use of robotic systems capable of analyzing 100 000 compounds per week.

Another aspect in many companies is the diversity of the compound pool. Typically, it contains many molecules with high structural similarity because they were prepared during lead optimization in which only slight variations on the architecture of the molecules were performed. The design and synthesis of diverse combinatorial libraries according to computational methods described in Chapter 1.6, and may contribute towards improving the overall diversity of a compound collection with respect to lead finding.

Thus, combinatorial chemistry is emerging as a time- and cost-efficient tool for drug discovery [155, 159], although there are presently very few compounds undergoing clinical trials that were discovered utilizing combinatorial chemistry (Table 1.2.).

Scores of patents already exist in the field of combinatorial chemistry directed towards methods of making libraries, to devices used in these methodologies, to assay methodologies, and on the libraries themselves [160].

**Table 1.2.** Combinatorial chemistry programmes in development [155].

Company	Area	Stage
Affymax (San Francisco, CA/USA)	Infectious diseases, cancer, hormone disorders	Pre-clinical screening
ArQuele (Medford, MA/USA)	Alzheimer's, cancer, autoimmune disease	Phase I and pre-clinical stages
Cadus (Tarrytown, NY/USA)	Cancer	Pre-clinical screening
Cambridge Combinatorial (Cambridge, UK)	Cancer, autoimmune disease	Pre-clinical library generation
Eli Lilly (Indianapolis, IN/USA)	Central nervous system disorders	Phase II clinical trials
ISIS (Carlsbad, CA/USA)	Anti-sense antivirals	Phase II
Magainin (Exton, PA/USA)	Anti-infectives	Phase I/II
NexStar (Boulder, CO/USA)	Anti-infectives, cancer	Pre-clinical and phase I
Pfizer (Groton, CT/USA)	Cancer	Phase I/II
Trega (San Diego, CA/USA)	Asthma	Phase I/II

Combinatorial chemistry will probably not substitute traditional medicinal chemistry, but it should be considered as one important tool in the process of drug discovery. Large-scale syntheses as well as numerous possibilities for structural variations of a given compound (e.g., replacement of atoms in a ring system, change of ring size, rigidization through introduction of bridging units) are still a domain of traditional synthetic chemistry.

## 1.8 Acknowledgements

The author gratefully thanks Dr. Stefan Peters (Boehringer Ingelheim Pharma KG) for his fruitful discussion concerning combinatorial chemistry and his valuable contribution, especially to Chapters 1.5 and 1.7.

## 1.9 References

- [1] W. H. Moos, M. R. Pavia, A. D. Ellington, B. K. Kay (eds.): *Annual Reports in Combinatorial Chemistry and Molecular Diversity*, vol. 1, ESCOM, Leiden 1997.
- [2] R. Cortes (ed.): *Combinatorial Libraries – Synthesis, Screening and Application Potential*, Walter de Gruyter, Berlin 1996.
- [3] I. Chaiken, K. Janda (eds.): *Molecular Diversity and Combinatorial Chemistry: Libraries and Drug Discovery*, ACS, Washington 1996.
- [4] G. Jung (ed.): *Combinatorial Peptide and Non Peptide Libraries – a Handbook*, VCH Weinheim 1996.
- [5] R. A. Houghten (ed.): *Combinatorial Libraries: The Theory and Practice of Combinatorial Chemistry*, ESCOM, Leiden 1995.
- [6] Special thematic issue 'Combinatorial Chemistry': *Chem. Rev.* 96 (1996).

- [7] Special thematic issue 'Combinatorial Chemistry': *Acc. Chem. Res.* 29 (1996).
- [8] F. Balkenhohl, C. von dem Bussche-Hünnefeld, A. Lansky, C. Zechel, *Angew. Chem. Int. Ed. Engl.* 35 (1996) 2289.
- [9] M. A. Gallop et al., *J. Med. Chem.* 37 (1994) 1233–1251 and 1385–1401.
- [10] A dynamic database of articles in the field of library synthesis (both peptide libraries and organic compound libraries) is maintained on the world wide web by the journal *Molecular Diversity* (<http://vesta.pd.com>).
- [11] A database *Combinatorial Chemistry and Mass Screening Yellow Pages* is maintained on the world wide web (<http://www.netsci.org/Resources/CCPY/top.html>).
- [12] H. M. Geysen, R. H. Meloen, S. J. Barteling, *Proc. Natl. Acad. Sci. USA* 81 (1984) 3998–4002.
- [13] H. M. Geysen, S. J. Barteling, R. H. Meloen, *Proc. Natl. Acad. Sci. USA* 82 (1985) 178–182.
- [14] R. A. Houghten, *Proc. Natl. Acad. Sci. USA* 82 (1985) 5131–5135.
- [15] R. B. Merrifield, *J. Am. Chem. Soc.* 85 (1963) 2149–2154.
- [16] R. Frank, *Tetrahedron* 48 (1992) 9217–9232.
- [17] S. P. A. Fodor et al., *Science* 251 (1991) 767–773.
- [18] US 5 143 854, 1992 (M. C. Pirrung, J. L. Read, S. P. A. Fodor, L. Stryer).
- [19] 'Studies on Possibilities of Systematic Searching for Pharmaceutically Useful Peptides': Á. Furka, notarized Document No. 36237/1982, Budapest, 1982.
- [20] Á. Furka, F. Sebestyén, M. Asgedom, G. Dibo, 1988, Abstr. 14th Int. Congr. Biochem., Prague, Czechoslovakia, Vol. 5., p. 47.
- [21] Á. Furka, F. Sebestyén, M. Asgedom, G. Dibo, *Int. J. Pept. Protein Res.* 37 (1991) 487–493.
- [22] F. Sebestyén, G. Dibo, A. Kovacs, Á. Furka, *BioMed. Chem. Lett.* 3 (1993) 413–418.
- [23] R. Hirschmann, *Angew. Chem. Int. Ed. Engl.* 30 (1991) 1278–1301.
- [24] X.-D. Xiang et al., *Science* 268 (1995) 1738–1742.
- [25] G. Briceno et al., *Science* 270 (1995) 273–275.
- [26] L. C. Hsieh-Wilson, X.-D. Xiang, P. G. Schultz, *Acc. Chem. Res.* 29 (1996) 164–170.
- [27] X. D. Sun, C. Gao, J. Wang, X.-D. Xiang, *Appl. Phys. Lett.* 70 (1997) 3353–3355.
- [28] B. E. Baker, N. J. Klinc, P. J. Treado, M. J. Natan, *J. Am. Chem. Soc.* 118 (1996) 8721.
- [29] E. Danielson et al., *Nature* 389 (1997) 944–948.
- [30] E. Danielson et al., *Science* 279 (1998) 837–839.
- [31] D. E. Akporiaye, I. M. Dahl, A. Karlsson, R. Wendelbo, *Angew. Chem.* 110 (1998) 629–631.
- [32] X.-D. Xiang, *Chemistry & Industry* 19 (1998) 800–802.
- [33] S. Borman, *Chem. Eng. News* 1996, 37–38.
- [34] K. Burgess, H.-J. Lim, A. M. Porte, G. A. Sulikowski, *Angew. Chem. Int. Ed. Engl.* 35 (1996) 220.
- [35] S. R. Gilbertson, S. Wang, *Tetrahedron Lett.* 37 (1996) 6475–6478.
- [36] Jacobsen, *J. Am. Chem. Soc.* 118 (1996) 8983–8984.
- [37] G. Liu, J. A. Ellman, *J. Org. Chem.* 60 (1995) 7712.
- [38] K. D. Shimizu et al., *Angew. Chem.* 109 (1997) 1782–1785; *Angew. Chem. Int. Ed. Engl.* 36 (1997) 1704–1707.
- [39] B. M. Cole et al., *Angew. Chem. Int. Ed. Engl.* 35 (1996) 1668–1671.
- [40] F. C. Moates et al., *Ind. Eng. Chem.* 35 (1996) 4801–4803.
- [41] F. M. Menger, A. V. Eleseev, V. A. Migulin, *J. Org. Chem.* 60 (1995) 6666.
- [42] F. M. Menger, Z. X. Fei, *Angew. Chem.* 106 (1994) 329–331; *Angew. Chem. Int. Ed. Engl.* 33 (1994) 346–348.
- [43] B. Jandeleit, W. H. Weinberg, *Chemistry & Industry* 19 (1998) 795–798.
- [44] K. S. Lam et al., *Nature* 354 (1991) 82–84.
- [45] M. Stankova et al., *Drug Dev. Res.* 33 (1994) 146–156.
- [46] K. S. Lam, M. Lebl, V. Krchnák, *Chem. Rev.* 96 (1996) 411–488.
- [47] R. A. Houghten et al., *Nature* 354 (1991) 82–86.
- [48] K. Burgess, A. I. Liaw, N. Wang, *J. Med. Chem.* 19 (1994) 2985–2987.
- [49] P.-L. Zhao et al., *Proc. Natl. Acad. Sci. USA* 92 (1995) 10212–10216.
- [50] J. A. Boutin, A. L. Fauchère, *Trends Pharmac. Sci.* 17 (1996) 8–12.
- [51] K. C. Nicolaou et al., *Angew. Chem. Int. Ed. Engl.* 34 (1995) 2289–2291.
- [52] E. J. Moran et al., *J. Am. Chem. Soc.* 117 (1995) 10787–10788.
- [53] Biomedic Data Systems, US 5 252 962, 1993.
- [54] Texas Instruments, US 5 351 052, 1994.
- [55] M. C. Pirrung, *Chem. Rev.* 97 (1997) 473–488.

- [56] H. V. Meyers et al., *Mol. Diversity* 1 (1995) 13–20.
- [57] S. Hobbs DeWitt, A. W. Czarnik, *Acc. Chem. Res.* 29 (1996) 114–122.
- [58] S. Hobbs DeWitt, J. S. Kiely, C. J. Stankovic, M. C. Schroeder, D. M. Reynolds Cody, M. R. Pavia, *Proc. Natl. Acad. Sci. U.S.A.* 90 (1993) 6909–6913.
- [59] Warner-Lambert-Company, WO 94/08711, 1994.
- [60] H. M. Geysen, S. J. Rodda, T. J. Mason, G. Triblick, P. G. Schoofs, *J. Immunol. Methods* 102 (1987) 259–274.
- [61] T. Kreiner, *Am. Lab.* 2 (1996) 244–255.
- [62] A. C. Pease et al., *Proc. Natl. Acad. Sci. USA* 91 (1994) 5022–5026.
- [63] J. M. Ostresh, J. H. Winkle, V. T. Hamashin, R. A. Houghten, *Biopolymers* 34 (1994) 1681–1689.
- [64] L. A. Thompson, J. A. Ellman, *Chem. Rev.* 96 (1996) 555–601.
- [65] J. S. Früchtel, G. Jung, *Angew. Chem. Int. Ed. Engl.* 51 (1996) 17–42.
- [66] P. H. H. Hermkens, H. C. J. Ottenheijm, D. Rees, *Tetrahedron* 53 (1996) 4527–4554.
- [67] Solid-Phase Chemistry brochures published by Chiron Mimotopes Pty. Ltd.
- [68] R. Storer, *Drug Discov. Today* 1 (1996) 248–254.
- [69] R. W. Armstrong et al., *Acc. Chem. Res.* 29 (1996) 123–131.
- [70] I. Ugi, A. Dömling, W. Hörl, *Endeavour* 18 (1994) 115–123.
- [71] I. Ugi, S. Lohberger, R. Karl in B. M. Trost, I. Fleming (eds.): *Comprehensive Organic Synthesis*, vol. 2, Pergamon, New York 1991, p. 1083–1109.
- [72] P. Hodge, *Chem. Soc. Rev.* 26 (1997) 417–424.
- [73] S. J. Shuttleworth, *Synthesis* 1997, 1217–1239.
- [74] S. Kobayashi, *Tetrahedron Lett.* 37 (1997) 9221–9224.
- [75] A. Akelah, D. C. Sherrington, *Chem. Rev.* 81 (1981) 557.
- [76] R. J. Booth, *J. Am. Chem. Soc.* 119 (1997) 4882–4886.
- [77] D. L. Flynn, *J. Am. Chem. Soc.* 119 (1997) 4874–81.
- [78] R. M. Lawrence, *Synthesis* 1997, 553–558.
- [79] S. W. Kaldor et al., *Tetrahedron Lett.* 37 (1996) 7193.
- [80] D. J. Gravert, K. D. Janda, *Chem. Rev.* 96 (1996) 489–509.
- [81] H. Han, M. M. Wolfe, S. Brenner, K. D. Janda, *Proc. Natl. Acad. Sci. U.S.A.* 95 (1995) 6419–6423.
- [82] D. P. Curran, S. J. Hadida, *J. Am. Chem. Soc.* 118 (1996) 2531–2532.
- [83] W. L. Fitch: 'Analytical Methods for the Quality Control of Combinatorial Libraries', in W. H. Moos, M. R. Pavia, A. D. Ellington, B. K. Kay (eds.): *Annual Reports in Combinatorial Chemistry and Molecular Diversity*, vol. 1, ESCOM, Leiden 1997, p. 59–68.
- [84] D. Stoll et al., in H. L. S. Maia (ed.): *Peptides 94, Proc. 23. EPS*, ESCOM, Leiden 1995, p. 401.
- [85] A. L. Burlingame, R. K. Boyd, S. J. Gaskell, *J. Anal. Chem.* 68 (1996) 599–601.
- [86] C. L. Brummel et al., *J. Anal. Chem.* 68 (1996) 237–239.
- [87] B. Yan, G. Kumaravel, *Tetrahedron* 52 (1996) 843–848.
- [88] B. Yan et al., *J. Org. Chem.* 60 (1995) 5736–5738.
- [89] P. A. Keifer et al., *J. Magn. Reson. Ser. A.* 119 (1996) 65–75.
- [90] S. K. Sarkar, R. S. Garigapati, J. L. Adams, P. A. Keifer, *J. Am. Chem. Soc.* 118 (1996) 2305–2306.
- [91] R. C. Anderson et al., *J. Org. Chem.* 60 (1995) 2650–2651;
- [92] G. C. Look, C. P. Holmes, J. P. Chinn, M. A. Gallop, *J. Org. Chem.* 59 (1994) 7588–7590.
- [93] J. A. Boutin et al., *Anal. Biochem.* 234 (1996) 126–128.
- [94] J. Metzger, S. Stevanovic, K. H. Wiesmüller, G. Jung, *Methods (San Diego)* 6 (1994) 425–432.
- [95] D. A. Campbell, J. C. Bermak, T. S. Burkoth, D. V. Patel, *J. Am. Chem. Soc.* 117 (1995) 5381–5382.
- [96] D. Vetter, E. M. Tate, M. A. Gallop, *Bioconjugate Chem.* 6 (1995) 319–322.
- [97] K. S. Lam, M. Lebl, *Methods, A Companion of Methods in Enzymology* 6 (1994) 372–380.
- [98] H. M. Geysen, S. J. Rodda, T. J. Mason, in R. Porter, J. Whelan (eds.): *Synthetic peptides as antigens*, John Wiley and Sons, Chichester 1986, p. 130–149.
- [99] C. T. Dooley et al., *Science* 266 (1994) 2019–2021.
- [100] C. T. Dooley, N. N. Chung, P. W. Schiller, R. A. Houghten, *Proc. Natl. Acad. Sci. USA* 90 (1993) 10811–10812.
- [101] C. Pinilla, J. R. Appel, P. Blanc, R. A. Houghten, *Biotechniques* 13 (1992) 901.
- [102] M. C. Pirrung, J. Chen, *J. Am. Chem. Soc.* 117 (1995) 1240–1241.
- [103] M. C. Pirrung, J. H.-L. Chau, *J. Chem. Biol.* 2 (1995) 621–624.
- [104] B. Deprez et al., *J. Am. Chem. Soc.* 117 (1995) 5405–5407.
- [105] C. Pinilla et al., *Drug Dev. Res.* 33 (1994) 133–137.

- [106] J. R. Wyatt et al., *Proc. Natl. Acad. Sci. USA* 91 (1994) 1356–1357.
- [107] C. T. Dooley, R. A. Houghten, *Life Sci.* 52 (1993) 1509–1602.
- [108] R. A. Houghten, *Gene* 137 (1993) 7–13.
- [109] E. Erb, K. D. Janda, S. Brenner, *Proc. Natl. Acad. Sci. USA* 91 (1994) 11422–11423.
- [110] L. Wilson-Lingardo et al., *J. Med. Chem.* 39 (1996) 2720.
- [111] S. M. Freier, D. A. M. Konings, J. R. Wyatt, D. J. Ecker, *J. Med. Chem.* 38 (1995) 344–346.
- [112] P. W. Smith et al., *Bioorg. Med. Chem. Lett.* 4 (1994) 2821–2823.
- [113] K. D. Janda, *Proc. Natl. Acad. Sci. USA* 91 (1994) 10779–10782.
- [114] J. Nielsen, S. Brenner, K. D. Janda, *J. Am. Chem. Soc.* 115 (1993) 9812–9814.
- [115] S. Brenner, R. A. Lerner, *Proc. Natl. Acad. Sci. USA* 89 (1992) 5381–5382.
- [116] J. C. Chabala, *Perspectives in drug discovery and design* 2 (1994) 305.
- [117] M. C. Needels et al., *Proc. Natl. Acad. Sci. USA* 90 (1993) 10700–10701.
- [118] E. R. Felder et al., *Mol. Diversity* 1 (1995) 109–112.
- [119] J. M. Kerr, S. C. Banville, R. N. Zuckermann, *J. Am. Chem. Soc.* 115 (1993) 2529–2531.
- [120] V. Nikolaev et al., *Pept. Res.* 6 (1993) 161–163.
- [121] Selectide, WO 94/28028 (M. Lebl et al.).
- [122] Ciba Geigy, WO 95/16209 (E. R. Felder, H. Rink, I. T. W. Matthews).
- [123] H. P. Nestler, P. A. Bartlett, W. C. Still, *J. Org. Chem.* 59 (1994) 4723–4725.
- [124] M. H. J. Ohlmeyer et al., *Proc. Natl. Acad. Sci. USA* 90 (1993) 10922–10923.
- [125] P. Eckes, *Angew. Chem. Int. Ed. Engl.* 33 (1994) 1573–1575; *Angew. Chem.* 106 (1994) 1649–1651.
- [126] J. J. Baldwin, J. J. Burbaum, I. Henderson, M. H. J. Ohlmeyer, *J. Am. Chem. Soc.* 117 (1995) 5588–5589.
- [127] J. J. Burbaum et al., *Proc. Natl. Acad. Sci. USA* 92 (1995) 6027–6028.
- [128] Z. J. Ni, D. Maclean et al., *J. Med. Chem.* 39 (1996) 1601–1604.
- [129] Z. J. Ni, D. Maclean, C. P. Holmes, M. A. Gallop, in J. N. Abelson (ed.): *Methods in Enzymology*, Academic Press, San Diego 1996, p. 261.
- [130] D. Madden, V. Krchnak, M. Lebl, *Perspect. Drug Discovery Des.* 1 (1995) 269.
- [131] M. Lebl, V. Krchnak, S. E. Salmon, K. S. Lam, *Methods (San Diego)* 6 (1994) 381.
- [132] S. E. Salmon et al., *Proc. Natl. Acad. Sci. U.S.A.* 90 (1993) 11708–11712.
- [133] P. Kocis, V. Krchnak, M. Lebl, *Tetrahedron Lett.* 34 (1993) 7251–7263.
- [134] Selectide, WO 95/00165 (M. Lebl, V. Krchnak, P. Kocis).
- [135] N. K. Terret et al., *Bioorg. Med. Chem. Lett.* 5 (1995) 917.
- [136] C. P. Holmes, D. G. Jones, *J. Org. Chem.* 60 (1995) 2318–2319.
- [137] S. H. DeWitt: 'Automated Synthesis', in W. H. Moos, M. R. Pavia, A. D. Ellington, B. K. Kay (eds.): *Annual Reports in Combinatorial Chemistry and Molecular Diversity*, vol. 1, ESCOM, Leiden 1997, p. 69–77.
- [138] S. Borman, *Chem. Eng. News* 1996, 29–54.
- [139] P. D. Hoepflich, *Nature Biotechnology* 14 (1996) 1311–1312.
- [140] J. H. Krieger, *Chem. Eng. News* 1996, 67–73.
- [141] J. A. Ellman, *Acc. Chem. Res.* 29 (1996) 132–143.
- [142] B. A. Bunin, M. J. Plunkett, J. A. Ellman, *Proc. Natl. Acad. Sci. USA* 91 (1994) 4708–4712.
- [143] B. A. Bunin, J. A. Ellman, *J. Am. Chem. Soc.* 114 (1992) 10997–10998.
- [144] US 5 288 514, 1994 (J. A. Ellman).
- [145] P. L. Myers, J. W. Greene, J. Saunders, S. L. Teig, *Today's Chemists at work* 6 (1997) 47.
- [146] J. M. Blaney, E. J. Martin, *Curr. Opin. Chem. Biol.* 1 (1997) 54.
- [147] E. J. Martin et al., *J. Med. Chem.* 38 (1995) 1431–1436.
- [148] W. H. Moos, G. D. Green, M. R. Pavia, *Annu. Rep. Med. Chem.* 28 (1993) 315.
- [149] D. E. Patterson et al., *J. Med. Chem.* 39 (1996) 3049–3059.
- [150] A. M. Ferguson, D. E. Patterson, C. D. Gary, T. L. Underiner, *J. Biomol. Screening* 1 (1996) 65.
- [151] S. D. Pickett, J. S. Mason, I. M. McLay, *J. Chem. Inf. Comput. Sci.* 36 (1996) 12.
- [152] H. Matter, *J. Med. Chem.* 40 (1997) 1219–1221.
- [153] P. Willett, *Trends Biotechnol.* 13 (1995) 516.
- [154] L. Weber, S. Wallbaum, C. Broger, K. Gubernator, *Angew. Chem. Int. Ed. Engl.* 34 (1995) 2280–2282; *Angew. Chem.* 107 (1995) 2452–2484.
- [155] A. Persidis, *Chemistry & Industries* 19 (1998) 782–784.
- [156] M. G. Edwards: 'Combinatorial Chemistry Alliances in the 1990s: Review of deal structure', in W. H. Moos, M. R. Pavia, A. D. Ellington, B. K. Kay (eds.): *Annual Reports in Combinatorial Chemistry and Molecular Diversity*, vol. 1, ESCOM, Leiden 1997, p. 314–320.



- [157] A front-end search engine of ReCap's RCA Database is available on the world wide web (<http://www.recap.com>).
- [158] Biovista industry review and company database: combinatorics 1998 edition, <http://www.biovista.com>.
- [159] D. V. Patel: 'Applications of combinatorial Technology to Drug Discovery', in W. H. Moos, M. R. Pavia, A. D. Ellington, B. K. Kay (eds.): *Annual Reports in Combinatorial Chemistry and Molecular Diversity*, vol. 1, ESCOM, Leiden 1997, p. 78–89.
- [160] K. Bozicevic: 'Patent Strategies in Molecular Diversity', in W. H. Moos, M. R. Pavia, A. D. Ellington, B. K. Kay (eds.): *Annual Reports in Combinatorial Chemistry and Molecular Diversity*, vol. 1, ESCOM, Leiden 1997, p. 298–313.

## 2 Survey of Solid-Phase Organic Reactions

*Susan E. Booth, C. Marijke Dreef-Tromp, Petro H. H. Hermkens,  
Jos A. P. A. de Man and Harry C. J. Ottenheijm*

### 2.1 Introduction

Although combinatorial chemistry is increasingly being extended to solution-phase synthesis approaches, the most significant work in this area has been performed using solid-phase strategies. As such, solid-phase chemistry has become a rapidly expanding area. More recently, this methodology has been adopted for the synthesis of small organic molecules and currently there is a growing number of classical organic reactions which have been successfully translated from solution onto the solid phase [1].

The development of a broad array of organic reactions on solid support will increase the scope of combinatorial chemistry. The synthetic challenge is the development of high-yielding and general reaction conditions. Moreover, a successful solid-phase synthesis comprises more than the synthetic route itself. It requires an appropriate choice of support, linker, reaction monitoring, purification and automation. For these aspects the reader is referred to corresponding chapters in this book. This chapter is meant to show that in the domain of Solid-Phase Organic Chemistry (SPOC) a start has been made, and that we probably have only glimpsed the tip of the iceberg.

Herein, an overview is given of some characteristic examples and highlights of reactions carried out on solid support. This overview does not include solid-phase peptide or other oligomer synthesis.

It is stressed that throughout this chapter only non-standard examples are given. As such, the tables are not comprehensive. The period covered is from 1992 until December 1997. This starting point is selected, as in 1992/1993 two key examples of solid-phase heterocyclic synthesis were published. These publications have provided a tremendous impetus to further studies. The groups of Ellman [2] and Hobbs DeWitt [3] published general and facile routes to benzodiazepines. However, these were not the first publications in this area. The techniques for solid-phase synthesis are based mainly on the pioneering work of Merrifield [4], and this was followed in the period 1970–1985 by some careful and elegant studies on solid-phase organic synthesis by Leznoff [5], Camps [6] and Frechet [7, 8]. It was Camps who, as early as 1974, published a synthesis of benzodiazepines using solid-phase chemistry.

## 2.2 Observed Trends

From the information presented the following trends can be observed and some potential areas for future developments are tentatively identified.

### 2.2.1 The Synthetic Repertoire

Although the synthetic repertoire of SPOC is growing rapidly, there are still opportunities for further development of reactions on solid phase, *e.g.*, Friedel–Crafts reactions. Those reactions that have been applied to solid phase in a reliable and robust manner, are listed below.

#### 2.2.1.1 Robust, Reliable Solid-Phase Reactions

These have been reported by several laboratories:

- Acylation: amide-bond formation, urea formation
- Alkylation: mainly N- and O-alkylation
- N-Arylation
- Bisaryl and cross-coupling (Heck, Suzuki, Stille)
- CC-condensation reactions (Aldol, Mannich, Claisen)
- Cyclo-additions (1,3 dipolar, [2 + 2], Diels–Alder)
- Deprotection (of N,O-functionalities)
- Halogenation of alcohols
- Heterocycle formation (Hantzsch condensation, Pictet–Spengler, Fischer synthesis of indoles, Pd-heteroannulation)
- Multiple component condensations (Ugi, Biginelli)
- Olefin formation (Wittig, Horner–Emmons, metathesis)
- Oxidation
- Pauson Khand
- Reduction (of nitro-groups)
- Reductive amination
- S<sub>N</sub>2 nucleophilic substitution (Mitsunobu, C-C, C-N, C-O formation)

Reactions carried out successfully on the solid phase, but which still warrant further study are highlighted below.

#### 2.2.1.2 Emerging Solid-Phase Reactions

Only a limited number of examples of these reactions have been published:

- Asymmetric reactions
- Aza–Wittig

- Baylis–Hillman
- Carbene, nitrene, radical reactions
- Cyclopropanation
- Long multi-step synthetic routes to construct non-oligomeric natural products and analogues
- Organometallic reactions (Grignard)
- Ozonolysis

### 2.2.2 Linkers and Cleavage Step

Cyclative cleavage strategies (*e.g.*, metathesis and heterocycle ring formation), traceless linkers (*e.g.*, Si, Ge linkers or *via* chemical reactions such as desulfurization, decarboxylation and cycloreversion) and cleavage steps designed to liberate an increasing diversity of functional groups are illustrated in Section 2.4.4.

For an overview of the linkers being used, the reader is referred to Chapter 5.

### 2.2.3 Reaction Monitoring

This remains a non-trivial issue and represents one of the practical aspects in which solid-phase chemistry frequently poses a greater challenge than the corresponding solution-phase chemistry. High-throughput NMR spectroscopy and liquid chromatography–mass spectrometry (LC-MS) are now common place among the armamentarium of the synthetic chemist.

### 2.2.4 Highlights

Recently, two elegant examples of multi-step solid-phase synthesis towards complex natural products have been reported [9, 10]. These developments demonstrate that the generation of natural product libraries using combinatorial chemistry is readily achievable. Other striking developments are the synthesis of ‘Multiple Core Structure Libraries’ (MCSL), featuring variations in the structure of the core as well as the substituents [11]. The use of polymer-bound reagents, either as scavengers or for ‘resin capture’, is also becoming common-place. These reagents enable the synthetic chemist to combine the best aspects of solution and solid-phase chemistry [12].

## 2.3 Conclusions

The rapidly emerging field of SPOC offers great opportunities and challenges for the organic chemist. Some of these challenges include: development of new SPOC reactions, linkers and resins; polymer-bound reagents; purification methods; and versatile methods for reaction monitoring. After lying dormant for 30 years, this powerful tool has now

awakened to become a readily accepted addition to medicinal chemistry and academic research, and is even finding its way into undergraduate teaching.

It was Merrifield who stated [13] in 1969 with regard to solid-phase organic chemistry: 'A gold mine awaits discovery by organic chemists'. This prompted Leznoff to add [5a] in 1978, 'Many gold nuggets have now been mined ... and some iron pyrites'. Twenty years later' one is inclined to conclude that we have only hit upon the first layer of ore [1a].

## 2.4 Reaction Tables

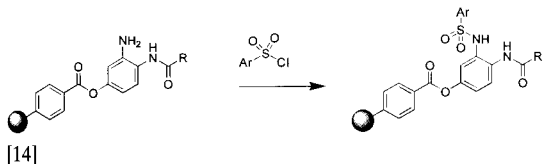
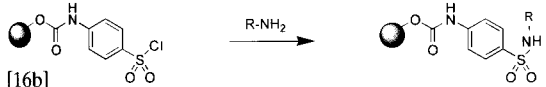
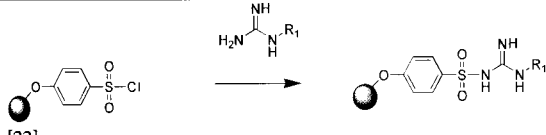
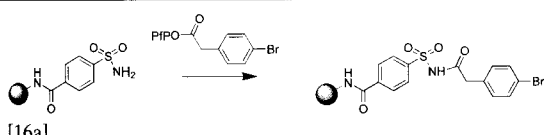
The reactions listed below are presented in a graphical abstract format, together with brief observations. The reference given in the reaction box refers to the graphical example. Related references are given in the final column. While we have attempted to make these tables as comprehensive as possible, some duplication was unavoidable. We have attempted, where possible to overcome this by cross-referencing. Two important developments are not addressed in these tables; Multiple Component reactions and Immobilization Strategies are discussed in Chapters 4 and 5, respectively.

- Amide bond formation and related reactions
- Aromatic substitution electrophilic and nucleophilic
- Cleavage
- Condensation
- Cyclo-addition
- Grignard and related reactions
- Heterocycle formation
- Michael addition
- Miscellaneous
- Olefin formation
- Oxidation
- Reduction
- Substitution nucleophilic and electrophilic

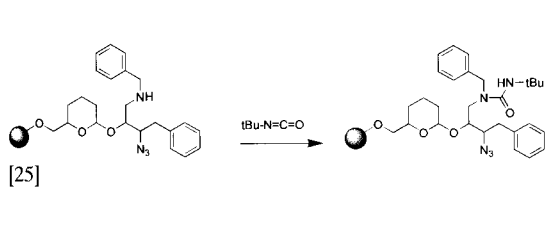
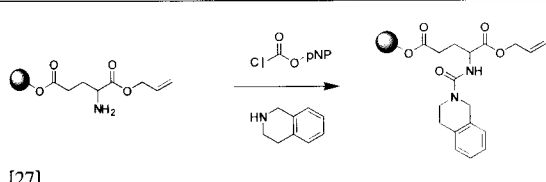
### 2.4.1 Substitution Nucleophilic and Electrophilic Type of Reaction: Amide Bond Formation and Related Reactions

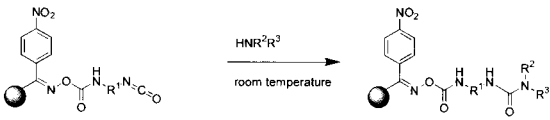
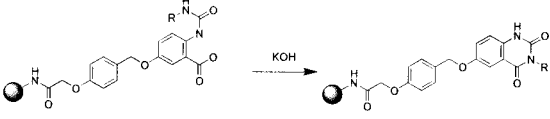
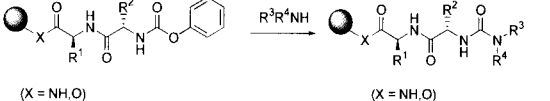
Amide bond formation on solid support is the basis of solid-phase peptide synthesis. Many different methods and reagents including a wide range of support-bound reagents, have been developed for these fundamental transformations. Among an ever-increasing number of publications highlighting amide bond formations and related reactions are procedures for the formation of ureas, thioureas, carbonates, urethanes, guanidines, sulphonamides and imides. These groups are important drug motifs and therefore justify further development.

## 2.4.1.1 Sulphonamide

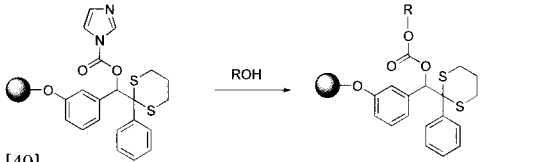
Some examples	General observations	Ref.
 <p>[14]</p>	<p>Resins/linkers used: Benzoic acid PS, TentaGel S OH, PEG-resin, Kaiser oxime and Wang.</p> <p>Sulphonylation of primary and secondary amines is possible.</p>	[14–24]
 <p>[16b]</p>	<p>Preparation of sulphonyl chlorides on resin is described for MBHA, MeO-PEG and Merrifield resin</p>	
 <p>[22]</p>	<p>The arenesulphonyl linker is suitable for producing guanidines (especially arginine). Both Boc and Fmoc protecting strategies are described.</p>	
 <p>[16a]</p>	<p>Activated acylsulphonamides (Safety Catch) can be cleaved with stoichiometric amounts of amines resulting in formation of corresponding amides.</p>	

## 2.4.1.2 (Thio)urea

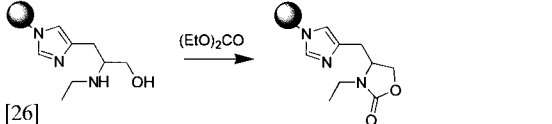
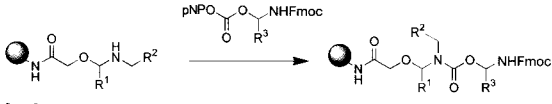
Some examples	General observations	Ref.
 <p>[25]</p>	<p>Both isocyanates and thioisocyanates are used on various resins to prepare (thio)ureas. Aromatic and aliphatic amines are suitable for condensation.</p>	[3] [14] [17]
 <p>[27]</p>	<p>An alternative procedure for urea formation starts from pNP-OCOCls and two amines. Repeating urea units afford unnatural biopolymers. TentaGel is often used as resin.</p>	[25–39]

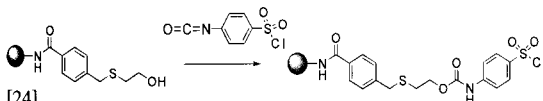
Some examples (Table 2.4.1.2 cont.)	General observations	Ref.
 <p>[32]</p>	<p>Selective urea formation on the Kaiser oxime resin without cleavage is only possible at room temperature. At elevated temperature cleavage is observed.</p> <p>Intramolecular formation of ureas is achieved <i>via</i> substitution reactions under basic conditions or by using carbodiimides. Library of 473,600 compounds is reported.</p> <p>Phenylcarbamate provides substituted ureas in 72–79% yield on TentaGel S AM resin.</p>	
 <p>[30]</p>		
 <p>(X = NH,O)</p> <p>[38]</p>		

## 2.4.1.3 Carbonate

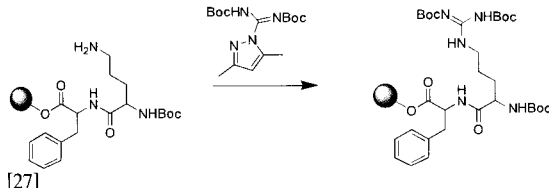
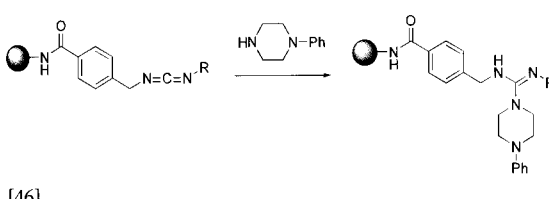
Some examples	General observations	Ref.
 <p>[40]</p>	<p>Addition of alcohols to solid-supported imidazocarbonates result in the generation of carbonates.</p> <p>Only a few examples are described.</p>	<p>[40]</p> <p>[41]</p>

## 2.4.1.4 Urethane

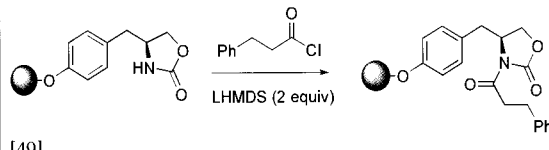
Some examples	General observations	Ref.
 <p>[26]</p>	<p>Addition of (EtO)<sub>2</sub>CO to amino alcohols yield cyclic substituted carbamates. Resin used: 2-Chlorotriphenyl methyl.</p> <p>A convenient method for the preparation of N-substituted urethanes involves the addition of pNP carbonates to resin bound amines.</p>	<p>[24]</p> <p>[26]</p> <p>[42–44]</p>
 <p>[42]</p>		

Some examples (Table 2.4.1.4 cont.)	General observations	Ref.
 <p>[24]</p>	<p>An alternative method is the condensation of resin bound isocyanates with alcohols. Resin used: MBHA.</p>	

## 2.4.1.5 Guanidine

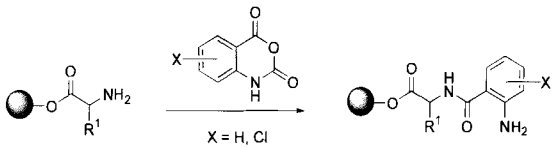
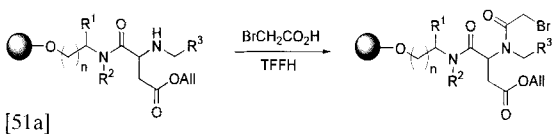
Some examples	General observations	Ref.
 <p>[27]</p>	<p>For direct guanidation of amines (for example lysine) many different reagents are used. This is an ideal method for the generation of diprotected guanidines.</p>	<p>[27] [39] [45–47]</p>
 <p>[46]</p>	<p>An alternative method is the nucleophilic addition of amines to carbodiimides (generated by an Aza Wittig reaction). Resin/linkers used: Rink or Wang. High yields reported.</p>	

## 2.4.1.6 Imide

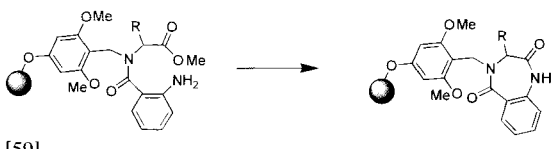
Some examples	General observations	Ref.
 <p>[49]</p>	<p>Acylation of amides and carbamates with acid chlorides to form imides can be monitored by FTIR and <sup>13</sup>C gel phase NMR.</p>	<p>[48] [49]</p>



## 2.4.1.7 Amide

Some examples	General observations	Ref.
 <p>[52]</p>	<p>Amides of different types are produced to form peptidomimetics and/or peptoids. Not only condensation reactions but also alkylation and substitution reactions are described.</p>	<p>[21] [50–57]</p>
 <p>[51a]</p>	<p>Resins/linkers used: Wang, Rink amide, Knorr TentaGel. Phosphoramidates were prepared on Pore glass, 3 split-pool libraries.</p>	

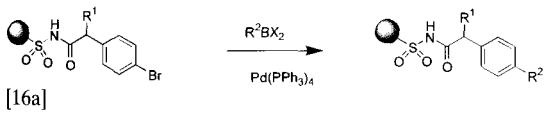
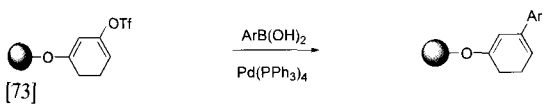
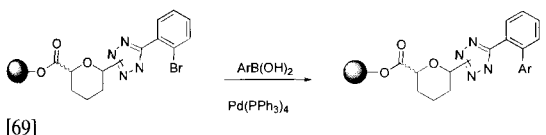
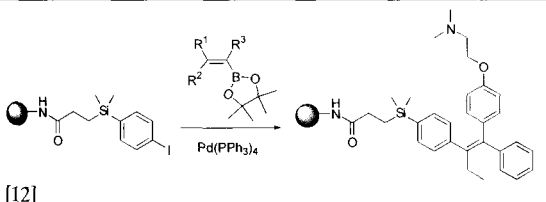
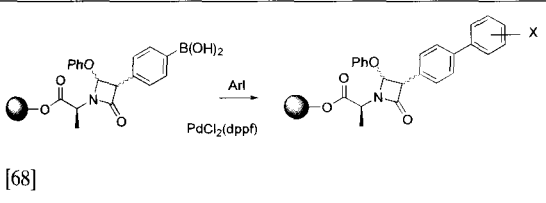
## 2.4.1.8 Lactam

Some examples	General observations	Ref.
 <p>[59]</p>	<p>Intramolecular cyclization is a common approach to lactams, e.g. benzodiazepines. Resins/linkers used: Rink amide and Wang.</p>	<p>[58] [59]</p>

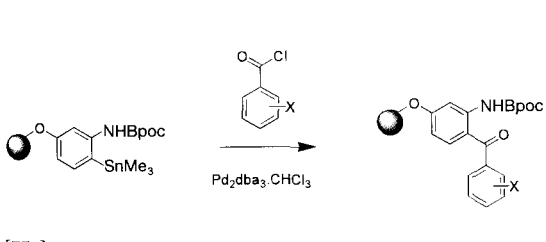
2.4.2 Type of Reaction: Aromatic Substitution;  
Electrophilic (Carbon-Carbon Bond Formation)

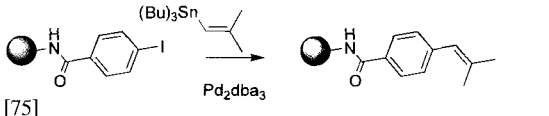
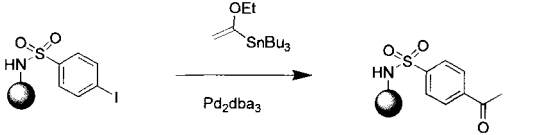
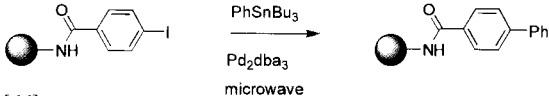
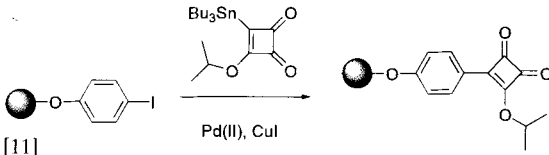
Since the renaissance of solid-phase organic chemistry in 1992, carbon-carbon bond formation reactions on solid support have probably been the best studied reactions. Many different facets of the Suzuki, Heck and Stille reactions have been evaluated. The influence of linkers, catalyst, solvents, microwave, polymer-bound aryl halides or polymer-bound arylboronic acids (or stannanes) have been studied in detail.

## 2.4.2.1 Suzuki

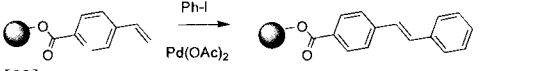
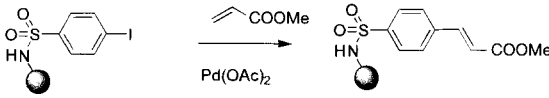
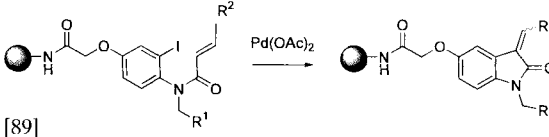
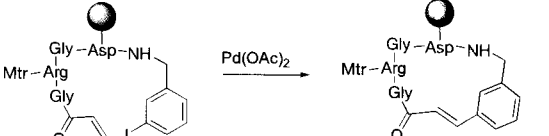
Some examples	General observations	Ref.
 <p>[16a]</p>	<p>Resins/linkers used: Safety Catch, Rink amide, Wang, Sasrin, THP, traceless.</p>	<p>[12] [16a] [60–73]</p>
 <p>[73]</p>	<p>In general, reactivity observed is I &gt; Br. Triflates have also been used successfully.</p>	
 <p>[69]</p>	<p><i>Ortho</i>-, <i>meta</i>- and <i>para</i>-substituted aryl halides react with both electron-rich and electron-poor boronic acids.</p>	
 <p>[12]</p>	<p>Both polymer-bound aryl halide and polymer-bound boronic acid give satisfactory results. Maximal diversity is obtained by <i>in situ</i> hydroboration of polymer-bound aryl iodides followed by addition of aryl iodides.</p>	
 <p>[68]</p>	<p>Reaction times are substantially decreased using microwave irradiation.</p> <p>Different Pd-catalysts have been evaluated. The optimal Pd catalyst varies depending on the boronic acid used. Pd<sub>2</sub>(dba)<sub>3</sub> shows excellent conversions with aryl- or heterobiarylboronic acids whereas PdCl<sub>2</sub>(dppf) works best for alkenylborane derivatives.</p>	

## 2.4.2.2 Stille

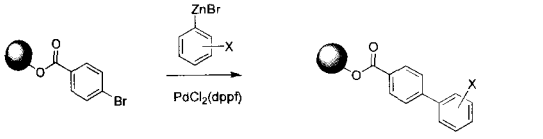
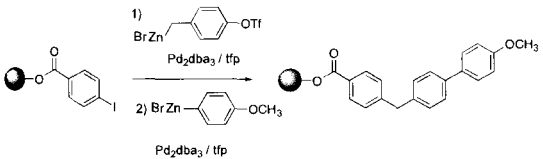
Some examples	General observations	Ref.
 <p>[77c]</p>	<p>Resins/linkers used: Rink, Wang, HMPA, traceless.</p> <p>The only example of a polymer-bound stannane is described in the coupling with acid chlorides. Good results are obtained with electron-rich and electron-poor acid chlorides.</p>	<p>[2] [11] [66] [74–80]</p>

Some examples (Table 2.4.2.2 cont.)	General observations	Ref.
 <p>[75]</p>	<p>Vinyl and aryl stannanes are coupled to polymer-bound aryl iodides. Hindered stannanes also show good results.</p> <p>Solution phase Stille couplings using polymer-bound or fluorous tin reagents have gained enormously in popularity.</p> <p>The use of microwave irradiation decreased reaction times.</p>	
 <p>[79]</p>		
 <p>[66]</p>		
 <p>[11]</p>		

## 2.4.2.3 Heck

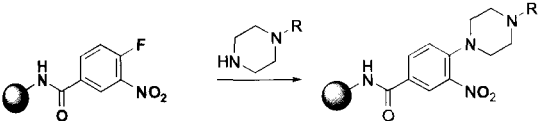
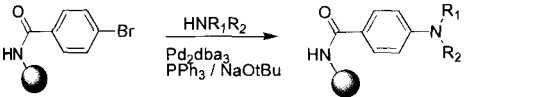
Some examples	General observations	Ref.
 <p>[82]</p>	<p>Resins/linkers used: Wang, Rink, NpSSM, THP, trityl.</p> <p>Heck reactions are described with either the alkene moiety or the aryl iodide attached to the solid support.</p> <p>Intramolecular Heck reactions have been used in the formation of:</p> <ul style="list-style-type: none"> <li>● indoles</li> <li>● oxindole</li> <li>● isoquinoline</li> <li>● benzofuran</li> </ul> <p>Macrocyclization by the Heck reaction has been shown to be a useful method in the preparation of new cyclic peptide structures.</p>	<p>[19] [65] [68] [79] [81-94]</p>
 <p>[79]</p>		
 <p>[89]</p>		
 <p>[92]</p>		

## 2.4.2.4 Other

Some examples	General observations	Ref.
 <p>[95]</p>	<p>Aryl zinc bromides undergo Pd-catalyzed coupling reactions to give polymer-bound aryl bromides. Aryl zinc reagents are easily prepared via transmetalation of lithium or magnesium organic reagents.</p> <p>Aryl and benzylic zinc reagents undergo efficient cross-coupling reactions with aryl iodides, bromides and triflates on solid phase. Multi-coupling reagents allow sequential formation of carbon-carbon bonds.</p>	<p>[95] [96]</p>
 <p>[96]</p>		

## 2.4.3 Type of Reaction: Aromatic Substitution; Nucleophilic (N-Arylation)

Two methods are used to introduce amines onto an aromatic system, *i.e.* substitution of halogens by either  $S_NAr$  or palladium-catalyzed amination.

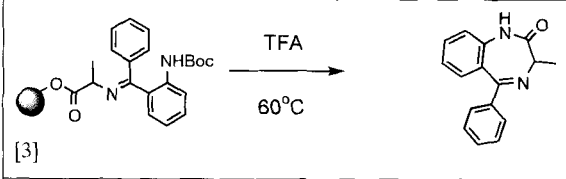
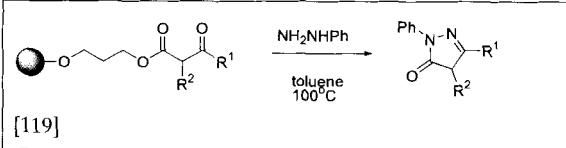
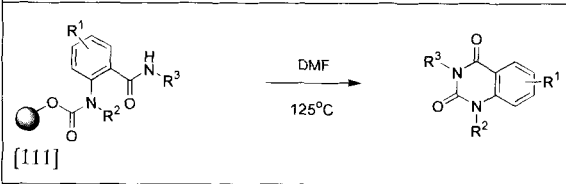
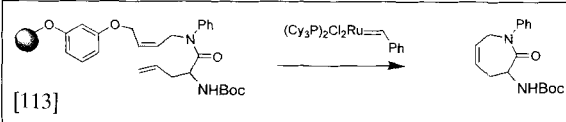
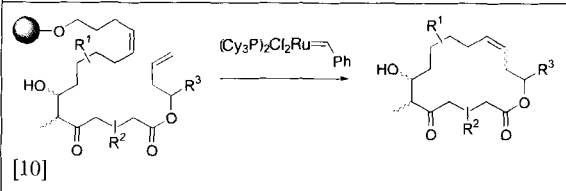
Some examples	General observations	Ref.
 <p>[97]</p>	<p><math>S_NAr</math> is only accomplished with aryl fluorides that are activated by the presence of <i>o</i>- or <i>p</i>-nitro groups.</p> <p>In Pd-catalyzed amination it is essential to add BINAP to suppress reduction of the bromide.</p>	<p>[26] [97–104]</p>
 <p>[105, 106]</p> <p>primary, secondary and aryl amines</p>		

## 2.4.4 Type of Reaction: Cleavage

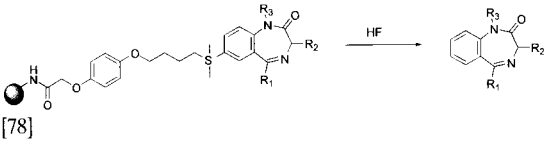

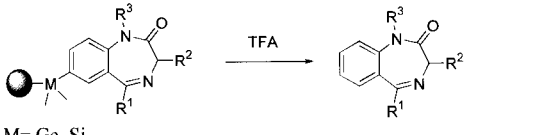

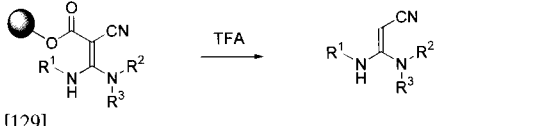
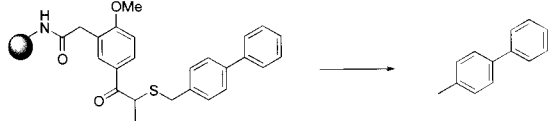
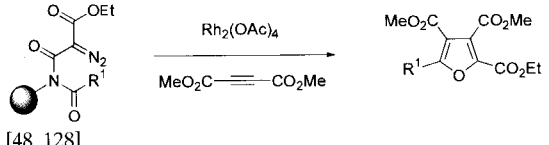
Linkers, cleavage and functional groups are condemned to each other. In the early days of SPOC, many of the linkers used originated from previous research on the solid-phase synthesis of oligomers, thus resulting in the cleavage of functional groups such as COOH (e.g. Wang and Sasrin resin) and CONH<sub>2</sub> (e.g. Rink amide resin). These functional groups are not always desirable in drug-like molecules. Of particular importance in the solid-phase synthesis of small molecule libraries has been the development of new modes of cleavage and new linkers. This has resulted in the development of methods and linkers which:

- allow cleavage without any trace of the linker (traceless linkers or cyclative cleavage)
- increase the diversity of functional groups to be cleaved (e.g. alcohols, amines, aldehydes, ketones, hydroxamic acids, amidine, guanidine).

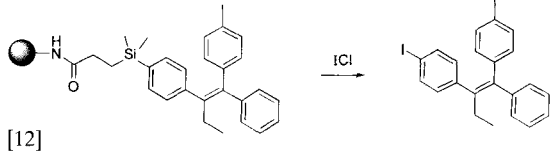

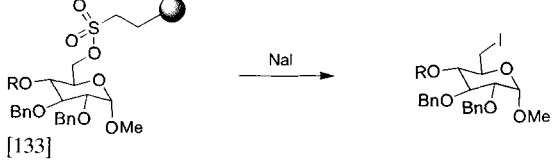
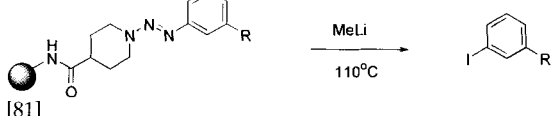
## 2.4.4.1 Cyclative Cleavage

Some examples	General observations	Ref.
 <p>[3]</p>	<p>Resins/linkers used: none, Wang, carbamate</p> <p>Heterocycle formation of 5-, 6- and 7-membered rings. Cleavage by cyclization enhances purity of the product, since incomplete products stay bound to resin.</p>	<p>[3] [9] [10] [20] [26] [29] [31] [101]</p>
 <p>[119]</p>	<p>Cyclization in solution after cleavage is also described as common methodology.</p>	<p>[107-122]</p>
 <p>[111]</p>	<p>Methods used for cyclization:</p> <ul style="list-style-type: none"> <li>● Ring Closure Metathesis (RCM)</li> <li>● Amidation (acidic or basic conditions)</li> </ul>	
 <p>[113]</p>		
 <p>[10]</p>		

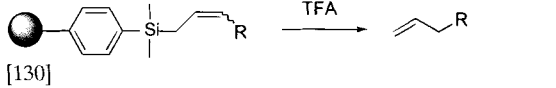
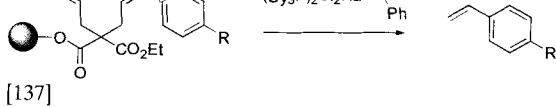
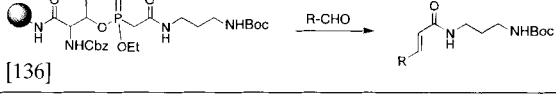
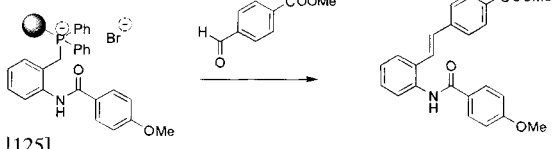
## 2.4.4.2 Functional Group: None (traceless)

Some examples	General observations	Ref.
 <p>[78]</p>	<p><i>Protodesilylation.</i> New linkers based on Si or Ge. The original linkers had to be cleaved by HF. Recently, more labile linkers have been developed that are cleaved with TFA or F<sup>-</sup>.</p>	<p>[12] [26] [48] [64] [67] [74] [78] [80]</p>
 <p>[64]</p>	<p>Substitution of Ge for Si results in more labile linkers. Electron-poor substrates are more difficult to cleave by protodesilylation.</p>	<p>[123–130]</p>
 <p>M = Ge, Si [80]</p>	<p>A drawback of many of today's Si-based traceless linkers is the cumbersome introduction of the first building block (immobilization). Usually the first group is introduced in solution.</p>	
 <p>[26]</p>	<p>Alternative traceless methods for cleavage are:</p> <ul style="list-style-type: none"> <li>● decarboxylation</li> <li>● desulphurization</li> <li>● cyclo-reversion</li> <li>● Reissert-based amide linkage</li> <li>● hydrolysis of phosphonium linker</li> </ul>	
 <p>[129]</p>		
 <p>[74, 123]</p>		
 <p>[48, 128]</p>		

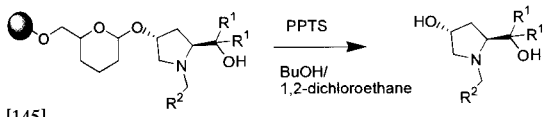
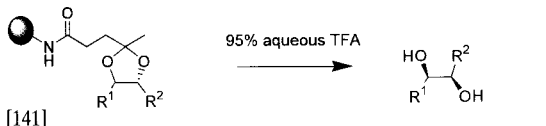

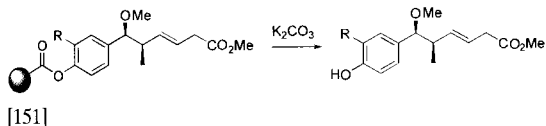
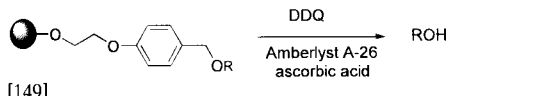
## 2.4.4.3 Functional Group: Halogens

Some examples	General observations	Ref.
 <p>[12]</p>	<p>Many of the new Si-based traceless linkers allow the introduction of halogen groups (halodesilylation).</p>	<p>[12] [81] [64] [131–135]</p>
 <p>[131]</p>	<p>Other methods for the introduction of halogens are:</p> <ul style="list-style-type: none"> <li>● iodolactonization</li> <li>● nucleophilic substitution (sulphonate ester linker)</li> <li>● halodediazonization; Sandmeyer-type reaction (triazene linker)</li> </ul>	
 <p>[133]</p>		
 <p>[81]</p>		

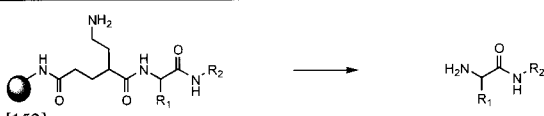
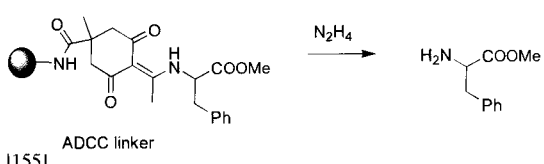
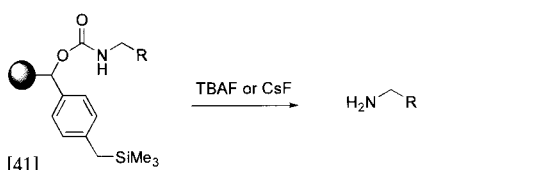
## 2.4.4.4 Functional Group: Alkenes

Some examples	General observations	Ref.
 <p>[130]</p>	<p>New designed linkers/methods are:</p>	<p>[125] [130] [136–138]</p>
 <p>[137]</p>	<ul style="list-style-type: none"> <li>● protodesilylation of allylsilyl linker</li> <li>● RCM of di-olefinic linker</li> <li>● phosphonate ester linker for Horner–Emmons based cleavage</li> </ul>	
 <p>[136]</p>	<ul style="list-style-type: none"> <li>● Phosphonium salt linker for Wittig-based cleavage. This linker is stable to a wide range of reaction conditions. Intramolecular Wittig reaction results in indole formation.</li> </ul>	
 <p>[125]</p>		

## 2.4.4.5 Functional Group: Alcohols, Phenols

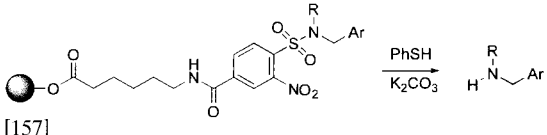
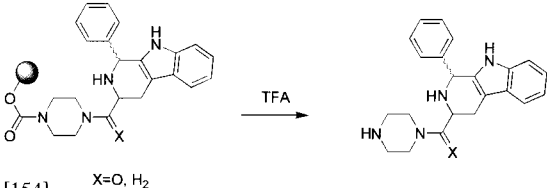
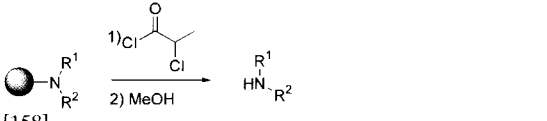
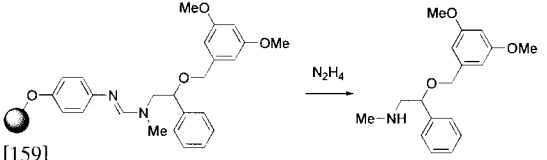
Some examples	General observations	Ref.
 <p>[145]</p>	<p>A traditional linker for cleavage of alcohols is the trityl linker.</p> <p>New linkers based on protecting groups for alcohols are:</p>	<p>[25] [26] [40] [41] [65] [74] [139–151]</p>
 <p>[141]</p>	<ul style="list-style-type: none"> <li>● THP linker for primary and secondary alcohols</li> <li>● ketal/acetal linkers</li> <li>● Si-based linkers</li> </ul>	
 <p>[142]</p>	<p>Alternative methods resulting in the cleavage of alcohols are:</p> <ul style="list-style-type: none"> <li>● reduction of ester-like compounds</li> <li>● saponification</li> <li>● oxidation of <i>p</i>-alkoxybenzyl linker with DDQ</li> </ul>	
 <p>[151]</p>		
 <p>[149]</p>		

## 2.4.4.6 Functional Group: Primary Amine

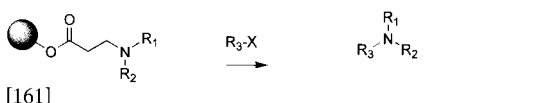
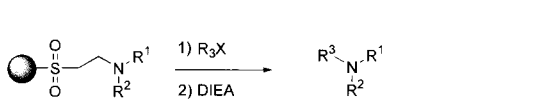
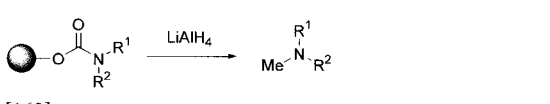
Some examples	General observations	Ref.
 <p>[152]</p>	<p>A traditional linker for cleavage of amines is the trityl linker.</p> <p>New linkers based on protecting groups for amines are:</p>	<p>[24] [41] [43] [70] [109] [152–156]</p>
 <p>[155] ADCC linker</p>	<ul style="list-style-type: none"> <li>● Carbamate linker prepared <i>via</i> phosgene (or derivative thereof), CDI or <i>p</i>-nitrophenol chloroformate</li> <li>● ADCC linker (stable to acidic and basic conditions)</li> </ul>	
 <p>[41]</p>	<p>An alternative cleavage method affording amines involved cyclative cleavage of the linker [152].</p>	



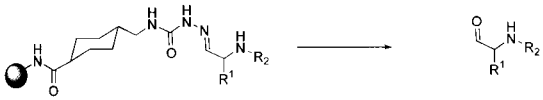



2.4.4.7 Functional Group: *sec*-Amine

Some examples	General observations	Ref.
 <p>[157]</p>	<p>New linkers based on protecting groups for <i>secondary</i> amines are:</p> <ul style="list-style-type: none"> <li>● Nosyl-like linkers</li> <li>● carbamate-based linkers</li> </ul>	<p>[154] [157–160]</p>
 <p>[154] X=O, H<sub>2</sub></p>	<p>Alternative cleavage methods for the formation of <i>secondary</i> amines are:</p> <ul style="list-style-type: none"> <li>● debenzylations</li> <li>● transamination of amidine linker</li> </ul>	
 <p>[158]</p>		
 <p>[159]</p>		

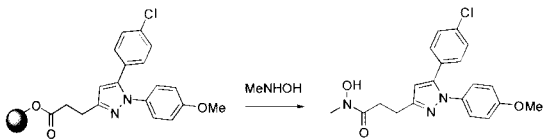
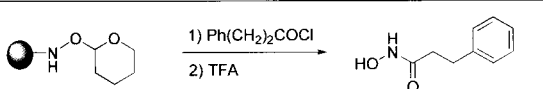
2.4.4.8 Functional Group: *tert*-Amine

Some examples	General observations	Ref.
 <p>[161]</p>	<p><i>tert</i>-Amines are liberated from the REM resin via Hoffmann elimination.</p> <ul style="list-style-type: none"> <li>● The alkylating reagent is limited to activated alkyl iodides and bromides.</li> </ul>	<p>[161–165]</p>
 <p>[165]</p>	<ul style="list-style-type: none"> <li>● REM resin is regenerated after cleavage.</li> </ul> <p>Vinyl sulphone is a more stable linker which allows a broader array of reaction conditions.</p>	
 <p>[162]</p>	<p>An alternative method to generate <i>tert</i>-amines is the reduction of carbamate-linked amines.</p>	

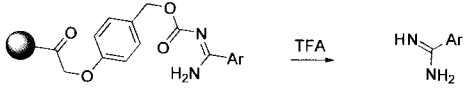
## 2.4.4.9 Functional Group: Aldehyde/Ketone

Some examples	General observations	Ref.
 <p>[166]</p>	<p>New linkers based on protecting groups for aldehydes/ketones are:</p> <ul style="list-style-type: none"> <li>● hydrazone-like linkers</li> <li>● enamine-based linkers</li> </ul>	<p>[26] [147] [166–174]</p>
 <p>[26]</p>	<p>Alternative methods resulting in the cleavage of aldehydes/ketones are:</p> <ul style="list-style-type: none"> <li>● selective reduction of esters and thioesters</li> <li>● reduction of Weinreb amide-based linker</li> <li>● reaction of Grignard reagents with Weinreb amide-based linkers</li> </ul>	
 <p>[147]</p>		
 <p>[174]</p>		

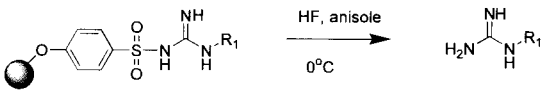
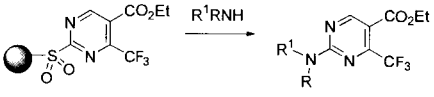
## 2.4.4.10 Functional Group: Hydroxamic Acid

Some examples	General observations	Ref.
 <p>[26]</p>	<p>Only one example of nucleophilic cleavage with hydroxylamine has been described. Recently, many different hydroxylamine-based linkers have been described based on attachment of hydroxylamine to trityl-, Wang-, Sasrin- or Rink-resin.</p>	<p>[26] [33] [175–177]</p>
 <p>[176]</p>		


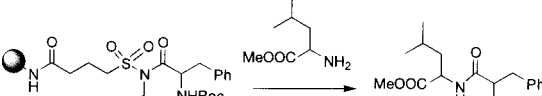
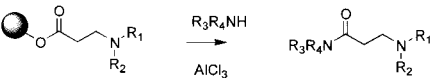
## 2.4.4.11 Functional Group: Amidine

Some examples	General observations	Ref.
 <p>[178]</p>	<p>One example using a carbamate linker has been described. The amidoxime formation is much slower on resin than in solution.</p> <p>The acid-labile linkers (Wang-type) gave higher yields than the photolabile linker.</p>	[178]

## 2.4.4.12 Functional Group: Guanidine

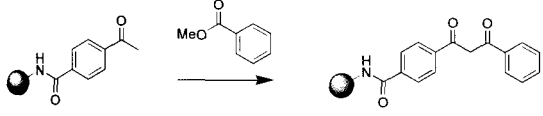
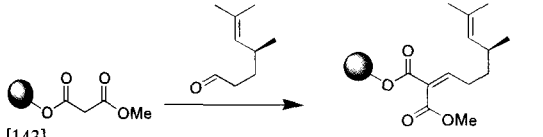
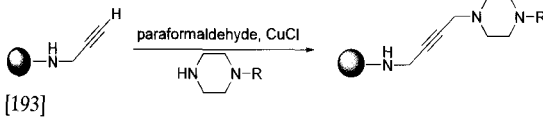
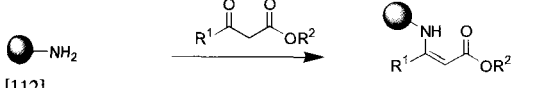
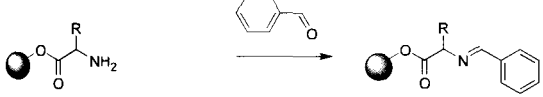
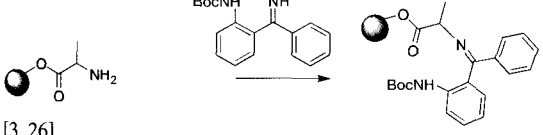
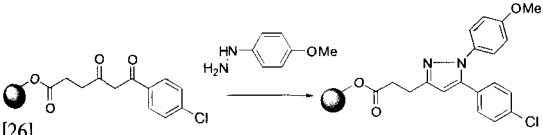
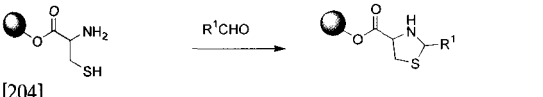
Some examples	General observations	Ref.
 <p>[22]</p>	<p>Based on arginine-protecting groups (Mtr-like), a new linker has been described that cleaves guanidine moieties. The disadvantage is that cleavage requires HF.</p> <p>An alternative method resulting in guanidine-like moieties is nucleophilic displacement of a sulphone group attached to a pyrimidine. The original sulphur linker is activated by oxidation (Safety Catch).</p>	[22] [179–180]
 <p>[180]</p>		

2.4.4.13 Functional Group: *sec* Amide/*tert* Amide/Sulfonamide

Some examples	General observations	Ref.
 <p>[184]</p>	<p>Three different possibilities have been described to cleave <i>sec/tert</i> amides/sulfonamides.</p> <ul style="list-style-type: none"> <li>● Known linkers (Wang, Sasrin) are modified for introduction of amines (reductive amination or alkylation). These can then be either acylated or sulfonylated.</li> <li>● Safety Catch linkers and Oxime resin.</li> <li>● Direct amidation of ester-linked moieties promoted by aluminium chloride results either in <i>sec</i> amides or <i>tert</i> amides.</li> </ul>	[16a] [52] [169] [181–186] [184] [185] [186]
 <p>[181]</p>		
 <p>[183]</p>		

## 2.4.5 Type of Reaction: Condensation

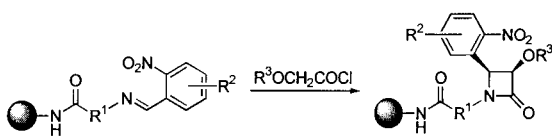
Condensation reactions are widely used in solid-phase organic chemistry. Transformations proceed cleanly and in high yield. Traditional Claisen, Knoevenagel, imine, Mannich, enamine and other condensation reactions are summarized in the next tables.

Some examples	General observations	Ref.
 <p>[187]</p>	<p><b>Claisen</b> Resins/linkers used: Merrifield, Rink amide and hydroxy methylene. Reaction unsuccessful if <math>\alpha</math>-hydrogens are present. Split-pool library prepared.</p>	<p>[26] [187–188]</p>
 <p>[143]</p>	<p><b>Knoevenagel</b> Resins/linkers used: Wang, Sasrin, Merrifield and hydroxy methylene. High yields reported.</p>	<p>[143] [153] [189–192]</p>
 <p>[193]</p>	<p><b>Mannich</b> Resins/linkers used: 2-chlorotriyl resin, CMPP (chloromethylphenyl pentyl polystyrene), Merrifield or Wang.</p>	<p>[148] [193]</p>
 <p>[112]</p>	<p><b>Enamine</b> A few examples are reported without yields. Resins/linkers used: Rink amide, PAL and Wang.</p>	<p>[112] [194]</p>
 <p>[197]</p>	<p><b>Imine-formation</b> Mostly imines are not isolated. Reaction times are slower if aldehyde component is tethered to solid-support. CH(OMe)<sub>3</sub> is commonly used as the dehydrating reagent.</p>	<p>[2] [3] [26] [59] [77] [78] [107] [195–197]</p>
 <p>[3, 26]</p>	<p>Resin/linkers used: Wang Rink amide, TentaGel, Sasrin, Sieber, Merrifield and multipins-aminomethylene.</p>	<p>[198–203]</p>
 <p>[26]</p>	<p><b>Pyrazole</b> Hydrazones can be isolated.</p>	<p>[26] [119] [187] [188]</p>
 <p>[204]</p>	<p><b>Thiazolidine and Thiazolidinone</b> Imine formation followed by cyclization.</p>	<p>[204–205]</p>

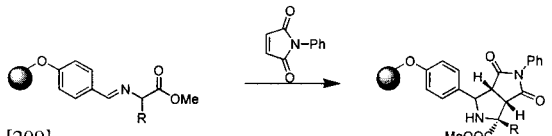
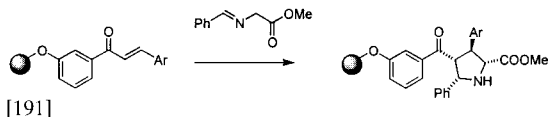
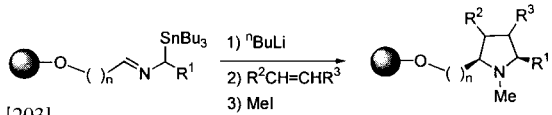
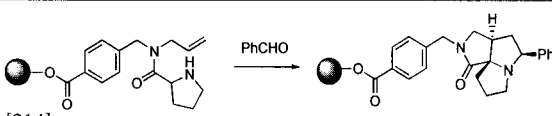
## 2.4.6 Type of Reaction: Cycloaddition

Cycloadditions are an important tool for the construction of cyclic systems of different ring-size. The reactions usually proceed smoothly and at most require moderate application of heat. Formation of 6-membered rings are realized *via* [4 + 2] cycloaddition [(hetero) Diels–Alder], 5-membered rings *via* the [3 + 2] cycloaddition [1,3-dipolar] and 4-membered rings *via* the [2 + 2] cycloaddition.

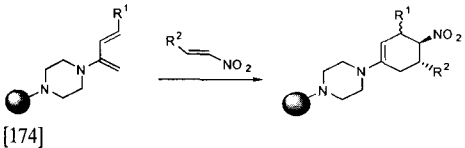
## 2.4.6.1 [2 + 2] Cycloaddition

Some examples	General observations	Ref.
 <p>[207]</p>	<p>Best known [2 + 2] cycloaddition is the formation of <math>\beta</math>-lactams.</p> <p>Sulphonyllactams are also reported.</p> <p>Resin used: TentaGel resin with photolabile linker, Sasrin and MBHA.</p>	<p>[202] [206] [207]</p>

## 2.4.6.2 [3 + 2] Cycloaddition

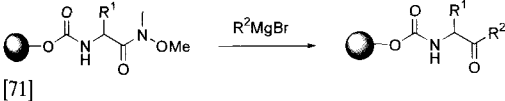
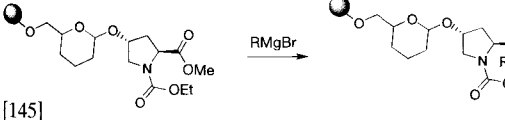

Some examples	General observations	Ref.
 <p>[209]</p>	<p>Resins/linkers used: PS carboxylic acid, Wang, Rink amide, Merrifield, TentaGel S AC and Sasrin.</p>	<p>[48] [131] [191] [197] [203]</p>
 <p>[191]</p>	<p>Many of the reactions have been monitored by FTIR. Several split-pool libraries reported.</p>	<p>[208–215]</p>
 <p>[203]</p>	<p>Pyrrolidines, isoxazoles and related small heterocycles prepared.</p>	
 <p>[214]</p>	<p>In general, regioisomers are formed.</p> <p>Moderate to good yields isolated.</p>	

## 2.4.6.3 [4+2] Cycloaddition

Some examples	General observations	Ref.
 <p>[174]</p>	(Hetero) Diels–Alder reactions. The first examples on solid support show comparable reactivity, regio- and stereo-selectivity. Resins used: Merrifield resin or PS carboxylic acid resin.	[174] [192] [216] [217]

## 2.4.7 Type of Reaction: Grignard and Related Reactions

Grignard (-like) reactions are emerging solid-phase reactions that have not yet been exhaustively studied. One is, therefore, tempted to conclude that the use of Grignard reactions in solid-phase chemistry is under-developed. A reason might be that the high reactivity of Grignard reagents coupled with the need for anhydrous conditions complicates with solid supported synthesis.

Some examples	General observations	Ref.
 <p>[71]</p>	Grignard (-like) reactions studied on solid support so far are: <ul style="list-style-type: none"> <li>● conversion of Weinreb amide to ketone</li> <li>● ester to ketone</li> <li>● ketones to alcohol</li> <li>● imine to amine</li> </ul> Resins/linkers used: ArgoGel, Wang, Merrifield, hydroxymethylene, chlorotrityl and Rink amide.	[41] [67, 71,73]
 <p>[145]</p>		[145] [218–221]
 <p>[219]</p>		

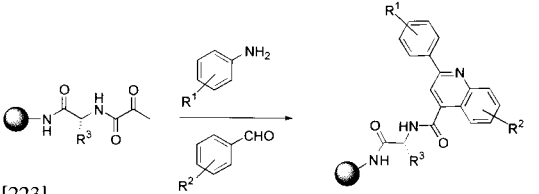
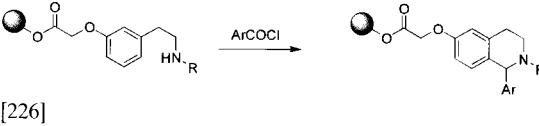
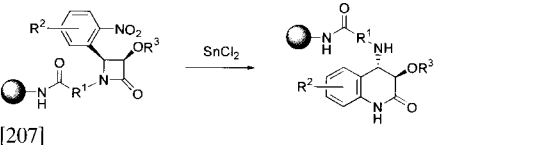
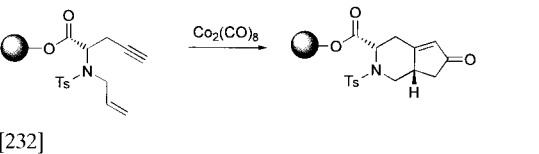
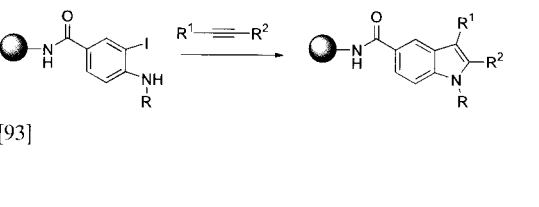
## 2.4.8 Type of Reaction: Heterocyclic Formation

Many of today's drugs are heterocyclic compounds. Heterocyclic compound libraries are therefore the focus of many pharmaceutical companies' drug discovery processes. Benzodiazepines were the first heterocyclic compounds prepared by solid-phase chemistry [2, 3, 6]. Since then, solid-phase routes to many heterocyclic structures have been published using well-established solution-phase protocols *e.g.* Pictet-Spengler reaction in the preparation of tetrahydro- $\beta$ -carboline.

## 2.4.8.1 Nitrogen-Containing Heterocycles

N-containing heterocycles prepared are:

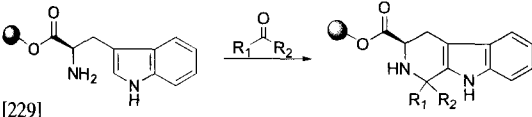
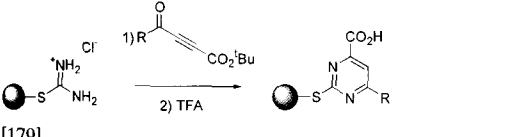
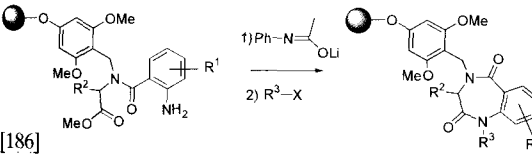
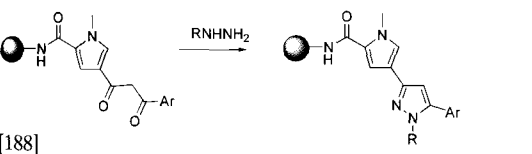
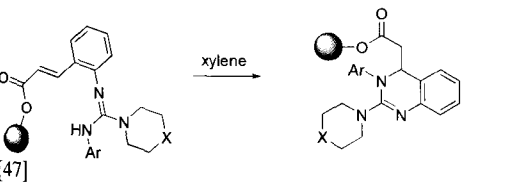
- 2-oxindole [88]
- dihydropiperidine [190]
- dihydroquinolinone [26, 207]
- dihydropyridine [221]
- pyrrole [222]
- quinoline [223]
- cyclopentapiperidinone [224]
- tetrahydroisoquinoline [225, 226]
- tetrahydrocyclopentapyridine [19, 227]
- indole [89, 90, 93, 94, 228, 229]

Some examples	General observations	Ref.
 <p>[223]</p>	<p><b>quinoline</b> Resins/linkers used: Rink amide.</p>	[223]
 <p>[226]</p>	<p><b>tetrahydroisoquinoline</b> <b>Bischler–Napieralski</b> Resins/linkers used: hydroxymethylene. High yields.</p>	[225] [226]
 <p>[207]</p>	<p><b>(dihydro)quinolinone</b> Resins/linkers used: MBHA resin. &gt;4000 Examples High overall yields.</p>	[26] [207]
 <p>[232]</p>	<p><b>tetrahydrocyclopenta pyridine (Pauson Khand);</b> Resins/linkers used: Wang resin. Co-catalyst Moderate to good yields.</p>	[19] [227]
 <p>[93]</p>	<p><b>indole</b> Via <b>Heck</b> reactions: Resins/ linkers used: Rink amide, Wang and Tentagel S-OH; Several Pd catalysts investi- gated. Several other methods reported.</p>	[89] [90] [93] [94] [228] [229]

## 2.4.8.2 Multiple Nitrogen-Containing Heterocycles

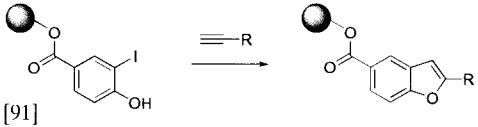
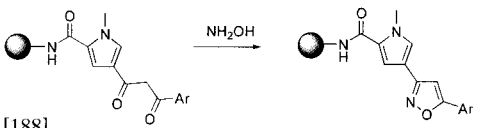
Multiple N-containing heterocycles prepared are:

- hydantoin [3, 26, 31, 38, 109]
- dihydroquinazoline [47]
- diazepine-2,5-dione [51]
- benzodiazepine-2-one [2,3,6,26,77,78]
- benzimidazole [99]
- tetrahydroquinoxalin-2-one [102]
- pyridobenzodiazepinone [126]
- pyrimidine [179]
- pyrazole [21, 26, 187, 188]
- triazole [194]
- tetrahydro- $\beta$ -carboline [154, 182, 230-232]
- imidazole [233]
- indazole [234]
- benzodiazepine-2,5-dione [186, 235]
- spiroindoline [236]
- triazine [237]
- quinazoline-2,4-dione [238]
- xanthine [239]


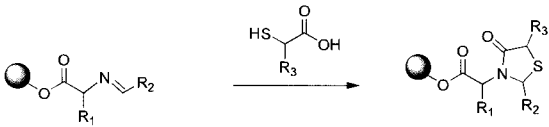
Some examples	General observations	Ref.
 <p>[229]</p>	<p><b>tetrahydro-<math>\beta</math>-carboline</b> Via <b>Pictet-Spengler</b> reaction Various resins used. High yields.</p>	<p>[154] [182] [230-232]</p>
 <p>[179]</p>	<p><b>pyrimidine</b></p>	<p>[179]</p>
 <p>[186]</p>	<p><b>benzodiazepine-2,5-dione</b> Merrifield resin; very high yields.</p>	<p>[186]</p>
 <p>[188]</p>	<p><b>pyrazole</b> Rink amide resin &gt; 5000 examples; split-pool synthesis. Other resins used: Tentagel S NH<sub>2</sub>; Hydroxymethylene</p>	<p>[21] [26] [187] [188]</p>
 <p>[47]</p>	<p><b>dihydroquinazoline</b> Wang resin. High overall yields.</p>	<p>[47]</p>



## 2.4.8.3 Oxygen-Containing Heterocycles

Some examples	General observations	Ref.
 <p>[91]</p>	<b>benzofuran</b> ; via Heck reaction Resins/linkers used: Rink amide, Tentagel S OH.	[26] [91] [93] [240]
 <p>[188]</p>	<b>benzopyran</b> : via Pictet-Spengler <b>isoxazole</b> : Rink amide resin 140 examples; split-pool synthesis.	[50] [187] [188]

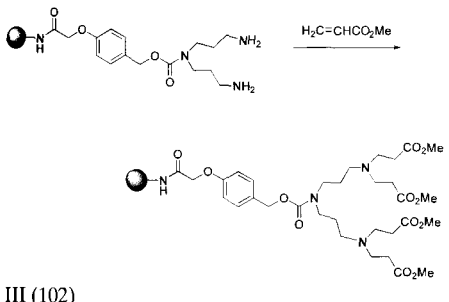
## 2.4.8.4 Sulphur-Containing Heterocycles

Some examples	General observations	Ref.
 <p>[56]</p>	<b>thiophene</b> : Resins/linkers used: Wang resin. Moderate to high yields.	[56] [241]
 <p>[200, 242]</p>	<b>4-thiazolidinone</b> , thiazole; thiazolidine; thiazanone; Tentagel or Sasrin resin; one-pot reaction; good yields	[200] [204] [205] [242]

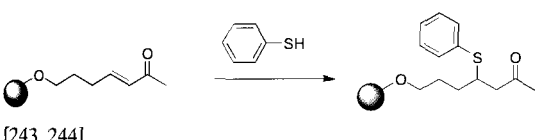
## 2.4.9 Type of Reaction: Michael Addition

The 1,4-Michael addition on solid phase was studied using various nucleophiles. Intramolecular and tandem Michael additions have also been investigated. Several representative examples are given. Until now, only one example has been reported where the nucleophile has been attached to the resin.

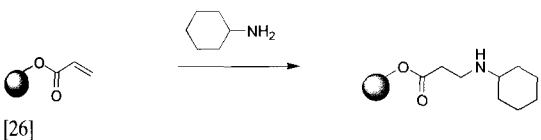
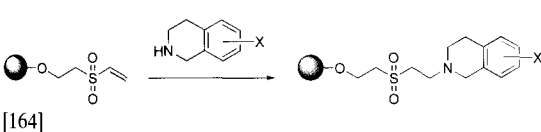
## 2.4.9.1 Immobilized Nucleophile

Some examples	General observations	Ref.
 <p>III (102)</p>	High-loading TentaGel; High yields	[246]

## 2.4.9.2 Thiol Addition

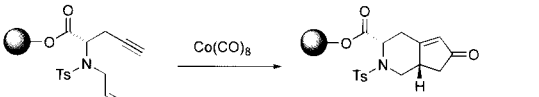
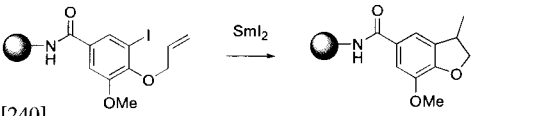
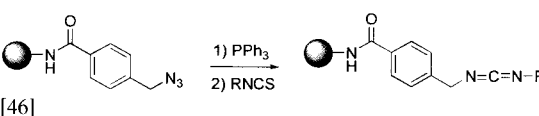
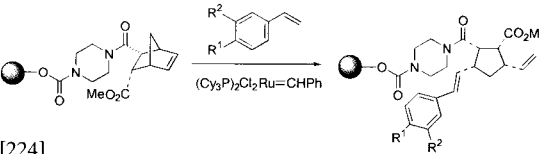
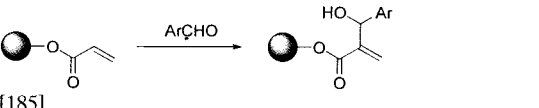
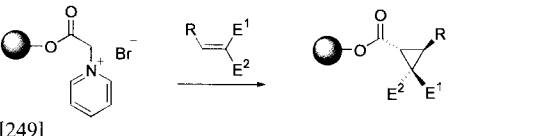
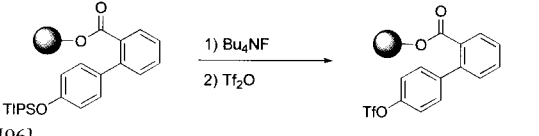
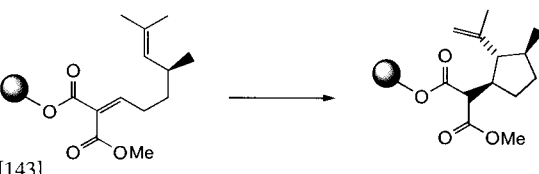
Some examples	General observations	Ref.
 <p>[243, 244]</p>	Resins/linkers used: tritylchloride resin; Rink resin reaction monitored by FT-IR	[89] [243] [244]

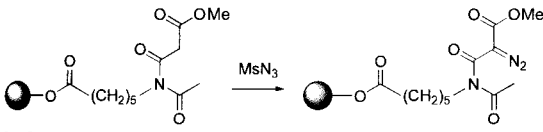

## 2.4.9.3 Amine Addition

Some examples	General observations	Ref.
 <p>[26]</p>	Resins/linkers used: Merrifield ;REM; Wang hydroxymethylene resin and Rink resin Primary and secondary amines are used	[26] [29] [161] [163] [164] [185] [245]
 <p>[164]</p>		

## 2.4.10 Type of Reaction: Miscellaneous

Some of the most interesting transformations are listed here that do not fit into the categorization outlined in this chapter.

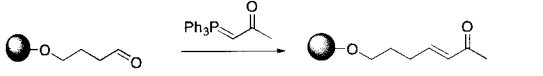
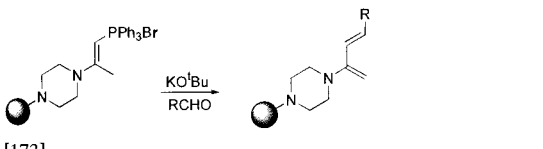
Some examples	General observations	Ref.
 <p>[232]</p>	<p><b>Pauson-Khand</b></p>	<p>[19] [227] [247]</p>
 <p>[240]</p>	<p><b>Radical Reaction</b> Radical cyclizations were mild, rapid and proceeded well at room temperature.</p>	<p>[240] [248]</p>
 <p>[46]</p>	<p><b>Aza-Wittig</b> The reaction proceeds <i>via</i> an <i>in situ</i> Staudinger reaction to generate the intermediate imino phosphorane.</p>	<p>[46] [47]</p>
 <p>[224]</p>	<p><b>Ring-Opening Metathesis</b> Regioselectivity is influenced by electronic and steric factors.</p>	<p>[224]</p>
 <p>[185]</p>	<p><b>Baylis-Hillman</b></p>	<p>[185]</p>
 <p>[249]</p>	<p><b>Cyclopropanation</b></p>	<p>[249]</p>
 <p>[96]</p>	<p><b>O-Triflation</b> Resin-bound triflates are powerful electrophiles of use in a number of cross-coupling reactions.</p>	<p>[96]</p>
 <p>[143]</p>	<p><b>Ene Rearrangement</b> Cyclopentane and cyclohexane derivatives have been prepared.</p>	<p>[143]</p>

Some examples (Table 2.4.10 cont.)	General observations	Ref.
 <p>[48]</p>	<p><b>Diazo transfer</b> Used as an intermediate in solid-phase rhodium(II)-catalyzed 1,3-dipolar cyclo-addition reactions.</p>	<p>[48] [128] [250]</p>
 <p>[124]</p>	<p><b>Electrophilic aromatic substitution</b></p>	<p>[124]</p>

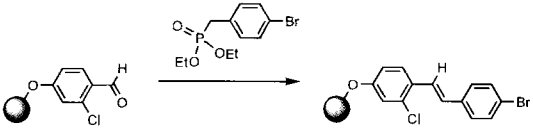
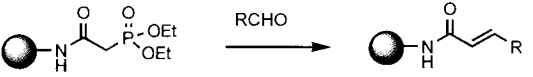
### 2.4.11 Type of Reaction: Olefin Formation

The classic ylid-based reactions used in the solution-phase synthesis of olefins have been applied successfully to the solid-phase. Cross-metathesis has been performed with a resin-bound and a dissolved olefin to provide more highly functionalized olefins.

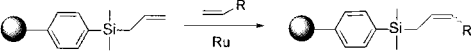
#### 2.4.11.1 Wittig

Some examples	General observations	Ref.
 <p>[244]</p>	<p>Attachment to the resin allows easy isolation of the olefinic products from the phosphorous containing by-products.</p>	<p>[9] [10] [121]</p>
 <p>[173]</p>	<p>The reaction works well when either the carbonyl or phosphine ylid is attached to the resin.</p> <p>Olefins built on the solid support are usually used as resin-bound building blocks for further derivatization.</p>	<p>[137] [173] [243] [244] [251]</p>

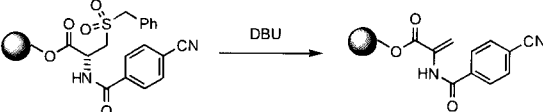
## 2.4.11.2 Horner–Emmons

Some examples	General observations	Ref.
 <p>[252]</p>	<p>High conversions to alkenes are accomplished under very mild conditions.</p>	<p>[136] [220] [251]</p>
 <p>[136]</p>	<p>Reaction progress can be followed by:</p> <ul style="list-style-type: none"> <li>● <math>^{31}\text{P}</math> NMR</li> <li>● ozonolysis and quantification of recovered aldehyde</li> </ul>	<p>[252] [253]</p>

## 2.4.11.3 Cross-Metathesis

Some examples	General observations	Ref.
 <p>[130]</p>	<p>Ring-closing metathesis has been employed to form nitrogen heterocycles.</p>	<p>[130] [221]</p>

2.4.11.4  $\beta$ -Elimination

Some examples	General observations	Ref.
 <p>[254]</p>	<p>Selenium compounds also undergo <math>\beta</math>-elimination after prior oxidation with <math>\text{NaIO}_4</math> [213].</p>	<p>[213] [254]</p>

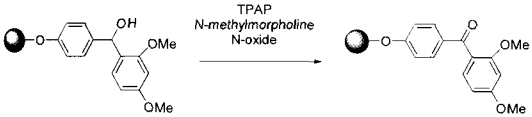
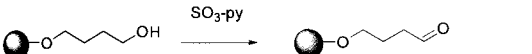
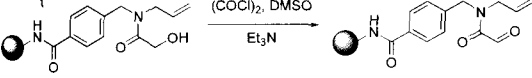
## 2.4.12 Type of Reaction: Oxidation

## 2.4.12.1 Alcohol to Aldehyde/Ketone


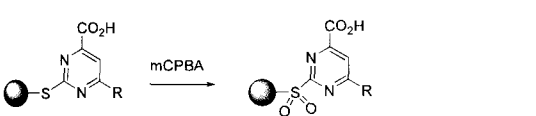
There are relatively few examples of oxidations on the solid support. As the conversions of aldehydes and ketones, however, lead to a wide range of molecules of interest for combinatorial chemistry, the on-resin oxidation of alcohols in particular is a useful syn-

thetic procedure. The oxidation of resin-bound 2-alkylthio-substituted pyrimidines to their corresponding 2-sulphonyl pyrimidines allows cleavage from the resin by displacement with various nucleophiles.


### 2.4.12.1 Alcohol to Aldehyde/Ketone

Some examples	General observations	Ref.
 <p>[255]</p>	<p>Primary and secondary alcohols can be oxidized to aldehydes and ketones using a variety of conditions:</p> <ul style="list-style-type: none"> <li>● SO<sub>3</sub>-py</li> <li>● NMO-TPAP</li> <li>● PDC</li> <li>● (COCl)<sub>2</sub>-DMSO-Et<sub>3</sub>N</li> </ul> <p>The transformations can be readily monitored by FT-IR.</p> <p>Oxidations are also used to prepare formyl and carboxylate resins.</p>	<p>[9, 10] [131] [159] [184] [199] [214] [243] [244] [251] [255] [256] [257]</p>
 <p>[256]</p>		
 <p>[214]</p>		

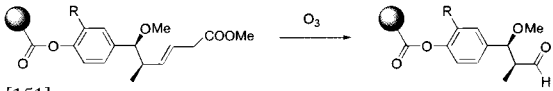
### 2.4.12.2 Sulfide to Sulfoxide/Sulfone

Some examples	General observations	Ref.
 <p>[26]</p>	<p>The oxidation to sulfoxide or sulfone can be controlled by the number of equivalents of mCPBA used.</p> <p>Selenium compounds have been oxidized with NaIO<sub>4</sub> in an oxidation/elimination sequence.</p>	<p>[26] [76] [165] [179] [180] [213] [258]</p>
 <p>[179]</p>		

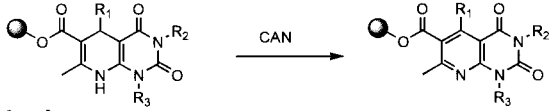
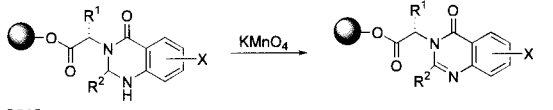
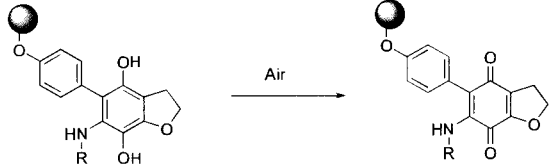
## 2.4.12.3 Epoxidation

Some examples	General observations	Ref.
 <p>[118]</p>	Nucleophilic opening of epoxides has been used to create new carbon-carbon or carbon-heteroatom bonds.	[118] [251] [259]

## 2.4.12.4 Ozonolysis

Some examples	General observations	Ref.
 <p>[151]</p>	Ozonolysis and quantification of recovered aldehyde can be used to monitor olefination reactions.	[151]

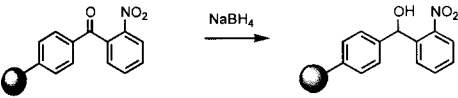
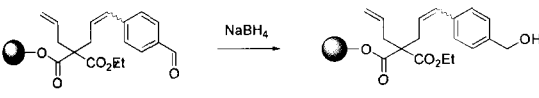
## 2.4.12.5 Other

Some examples	General observations	Ref.
 <p>[190]</p>	Aromatization	[11] [20] [52]
 <p>[52]</p>	iminal to amidine	[190] [234]
 <p>[11]</p>	1,4-phenol to 1,4-quinone	

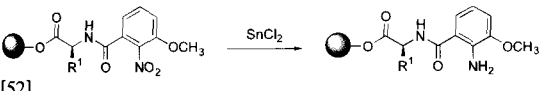
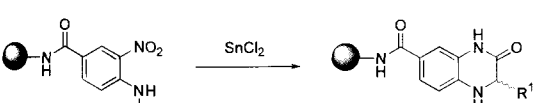
## 2.4.13 Type of Reaction: Reduction

The most common reduction performed on the solid phase is the reductive amination. The reaction has been performed with either resin-bound amine or aldehyde. Reduction of a number of other functional groups has also been applied to the solid phase, usually to generate handles for further diversification, *e.g.* nitro to amine, azide to amine.

## 2.4.13.1 Aldehyde/Ketone to Alcohol

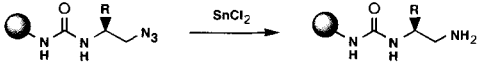
Some examples	General observations	Ref.
 <p>[260]</p>	Used to prepare photosensitive <i>o</i> -nitrobenzhydrylamino (NBHA) resin in the synthesis of fully protected C-terminal amidepeptides.	[26] [137] [260] [261]
 <p>[137]</p>		

## 2.4.13.2 Nitro to Aniline

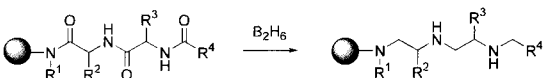
Some examples	General observations	Ref.
 <p>[52]</p>	Me <sub>2</sub> S.BH <sub>3</sub> has also been used successfully.	[14, 52] [99] [102] [114] [125]
 <p>[102]</p>		



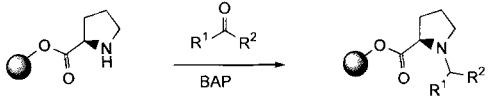
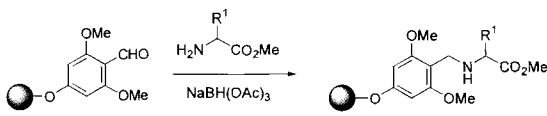
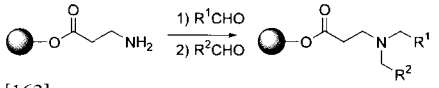
## 2.4.13.3 Azide to Amine

Some examples	General observations	Ref.
 <p>[264]</p>	<p>The azide serves as a 'protected' amine. Two reagents are successfully applied in the reduction of azide to the amine functionality, <i>i.e.</i>, SnCl<sub>2</sub> and PPh<sub>3</sub>.</p> <p>Reaction progress can easily be monitored by FT-IR.</p>	<p>[25] [126] [263] [264] [265]</p>

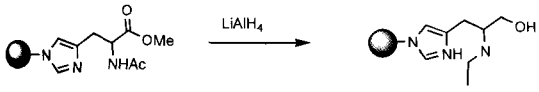
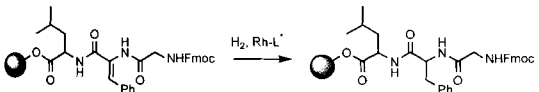
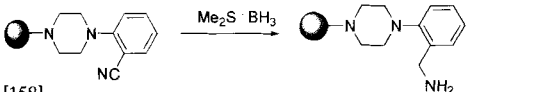
## 2.4.13.4 Amide to Amine

Some examples	General observations	Ref.
 <p>[34]</p>	<p><i>sec</i>- amides can be reduced to the corresponding <i>tert</i>- amines with Red-Al.</p> <p>'Libraries from libraries' have been produced by reduction of support bound-peptides to polyamines.</p>	<p>[34] [42] [145] [266]</p>

## 2.4.13.5 Reductive Alkylation/Amination

Some examples	General observations	Ref.
 <p>[273]</p>	<p>Reducing agents used:</p> <ul style="list-style-type: none"> <li>● LiBH<sub>4</sub></li> <li>● BH<sub>3</sub>-py (BAP)</li> <li>● NaCNBH<sub>3</sub></li> <li>● NaBH(OAc)<sub>3</sub></li> </ul>	<p>[26] [59] [65] [67] [107] [120]</p>
 <p>[186]</p>	<p>Reductive alkylation works well with aminoesters and a range of anilines.</p> <p>BAP has been found to be the best reducing agent for less stable iminium ion intermediates formed with <i>sec</i>- amines and aldehydes/ketones.</p>	<p>[161] [163] [186] [195] [196] [197] [198] [199] [219] [225]</p>
 <p>[163]</p>	<p>Partial racemization of amino acids is sometimes observed. <i>Bis</i>-alkylation is sometimes observed with aliphatic aldehydes.</p>	<p>[260] [267–273]</p>

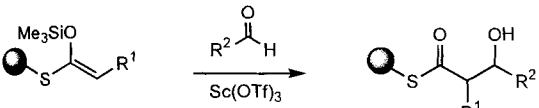
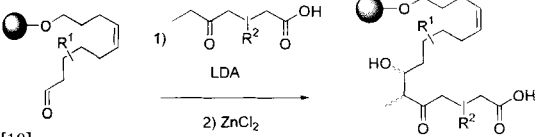
## 2.4.13.6 Other

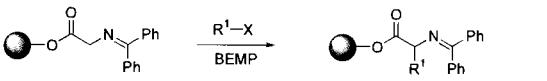
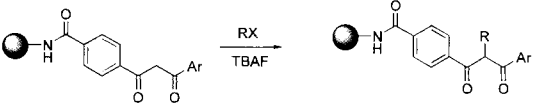
Some examples	General observations	Ref.	
 <p>[26]</p>	DIBAL has been used in the reductive cleavage of resin-bound esters to alcohols Alkenes and alkynes can both be reduced.	[26, 63] [67, 74] [123] [158] [180] [227] [251]	
 <p>[275]</p>			Other reductions include: <ul style="list-style-type: none"> <li>● nitrile to amine.</li> <li>● diazo to hydrazine</li> <li>● disulfide to thiol</li> <li>● epoxide to alcohol</li> </ul>
 <p>[158]</p>			

## 2.4.14 Type of Reaction: Substitution

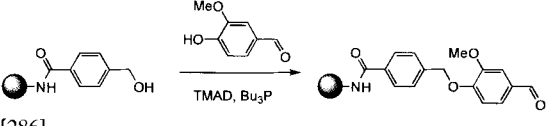
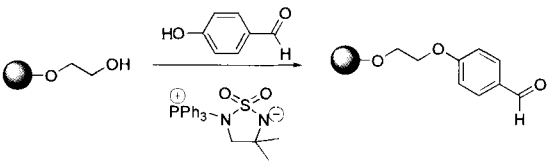
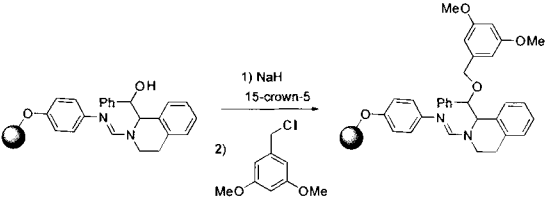
Many substitution and addition reactions have been carried out on the solid phase as these reactions are ideally suited for the introduction of commercially available building blocks such as amines and alkyl halides. The reaction conditions, however, are often incompatible with other functionalities in the molecule. An alternative reaction carried out under neutral conditions is the Mitsunobu reaction. A variety of Mitsunobu reactions have been carried out on the solid support, where work-up is much more straightforward than the analogous solution-phase reaction.

## 2.4.14.1 C-Alkylation (Aldol, Anion)

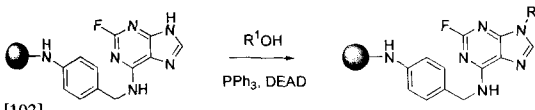
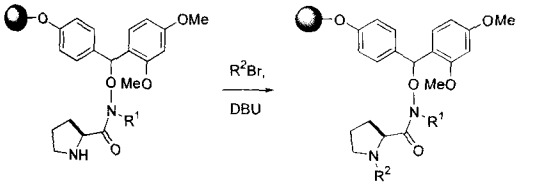
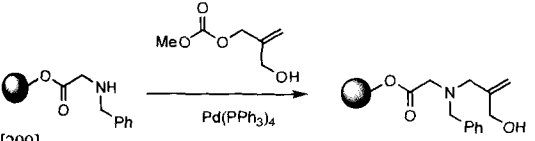
Some examples	General observations	Ref.
 <p>[147]</p>	Best results are observed using $\text{Sc}(\text{OTf})_3$ as a Lewis acid catalyst. Similar solution-phase reactions have been carried out using a resin-bound scandium catalyst.	[9, 10] [16a] [40, 49] [56, 71] [77b] [116] [119]
 <p>[10]</p>		

Some examples (Table 2.4.14.1 cont.)	General observations	Ref.
 <p>[282, 283]</p>	<p>Alkylation of the <math>\alpha</math>-carbon of benzophenone imines allows the on-resin preparation of unnatural amino acids for further transformation into unusual peptides and other classes of compounds. BEMP provides mild and selective alkylation conditions thus avoiding racemization.</p> <p>The use of TBAF as base avoids O-alkylation.</p>	<p>[146] [147] [148a] [150] [151] [160] [187] [188] [212] [237] [241] [277–283]</p>
 <p>[188]</p>		

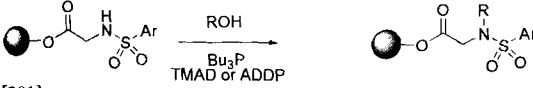
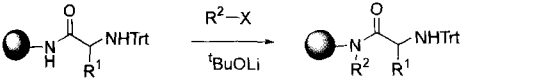
## 2.4.14.2 O-Alkylation (Mitsunobu, Anion)

Some examples	General observations	Ref.
 <p>[286]</p>	<p>Mitsunobu reactions can be performed with either the more or less acidic coupling partner attached to the solid support. Good results are seen with: DEAD, DIAD, ADDP or TMAD and either PPh<sub>3</sub> or PBu<sub>3</sub>.</p>	<p>[33, 61] [71, 74] [93] [127] [133] [137] [149] [159] [160] [164] [177] [180] [250] [259] [265] [271] [284–295]</p>
 <p>[271]</p>		
 <p>[160]</p>	<p>Organic bases such as DBU can also be used.</p>	

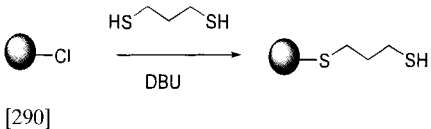
## 2.4.14.4 N-Alkylation (Mitsunobu, Anion)

Some examples	General observations	Ref.
 <p>[103]</p>	Sterically hindered <i>secondary</i> alcohols also work well. The mild reaction conditions of the Mitsunobu reaction allows alcohols containing base-sensitive groups to be incorporated.	[2, 25] [26, 58] [59, 31] [51a, 69] [70, 77] [84, 85] [88, 93]
 <p>[177]</p>	Sodium hydride and a number of other organic bases have been used successfully.	[94, 97] [102] [103] [113] [115] [118] [126] [127] [165]
 <p>[299]</p>	The Pd-catalyzed allylation reaction provides useful templates for combinatorial libraries.	[177] [196] [226] [239, 249] [263, 276] [296–300]

## 2.4.14.4 N-Alkylation (Sulfonamide, Amide)

Some examples	General observations	Ref.
 <p>[301]</p>	'Libraries of libraries' have been prepared by replacing amide hydrogens with different alkylating groups.	[2, 19] [26, 30] [34, 59] [70, 79] [121, 126]
 <p>[34]</p>	Successive amide alkylations allow selective choice of alkylating reagents at each step in the synthesis, in contrast to the more common permethylation of all amide sites.	[157] [176] [181] [227] [238] [269] [301, 302]

## 2.4.14.5 S-Alkylation

Some examples	General observations	Ref.
	Thiol alkylation on the solid-phase has been used to prepare $\beta$ -turn mimetics.	[10, 56] [70, 74] [118, 123] [138, 201] [276, 290] [303, 304]

## 2.5 References

- [1] a) Hermkens, P. H. H.; Ottenheijm, H. C. J.; Rees, D. C. *Tetrahedron* **1996**, *52*, 4527–4554. b) Hermkens, P. H. H.; Ottenheijm, H. C. J.; Rees, D. C. *Tetrahedron* **1997**, *53*, 5643–5678. c) Booth, S.; Hermkens, P. H. H.; Ottenheijm, H. C. J.; Rees, D. C. *Tetrahedron*, in press. d) Balkenhohl F.; von dem Bussche-Hunnefeld C.; Lansky A.; Zechel C., *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 2288–2337. d) Bunin, B. A. *The Combinatorial Index*, Academic Press, San Diego, 1998.
- [2] Bunin, B. A.; Ellman, J. A.; *J. Am. Chem. Soc.*, **1992**, *114*, 10997–10998.
- [3] Hobbs DeWitt, S.; Kiely, J. S.; Stankovic, C. J.; Schroeder, M. C.; Reynolds Cody D. M.; Pavia M. R., *Proc. Natl. Acad. Sci. USA*, **1993**, *90*, 6909–6913.
- [4] Merrifield, R. B. *J. Am. Chem. Soc.*, **1963**, *85*, 2149–2154.
- [5] Leznoff, C. C.; *Acc. Chem. Res.*, **1978**, *11*, 327–333 and all references cited therein. b) Worster, P. M.; Mc. Arthur, C. R.; Leznoff, C. C.; *Angew. Chem.*, **1979**, *91*, 255. c) Leznoff, C. C.; Yedidia, V.; *Can. J. Chem.*, **1980**, *58*, 287–290. d) Yedidia, V.; Leznoff, C. C.; *Can. J. Chem.*, **1980**, *58*, 1144–1150.
- [6] Camps, E.; Cartells, J.; Pi, J.; *Anales de Quimica*, **1974**, *70*, 848–849.
- [7] a) Frechet, J. M. J.; *Tetrahedron*, **1981**, *37*, 663–668. b) Farrall, M. J.; Frechet, J. M. J.; *J. Org. Chem.*, **1976**, *46*, 3877–3882.
- [8] Frechet, J. M. J.; Schuerch, C.; *J. Am. Chem. Soc.*, **1971**, *93*, 492–496.
- [9] Nicolaou, K. C.; Winssinger, N.; Pastor, J.; Ninkovic, S.; Sarabia, F.; He, Y.; Vourloumis, D.; Yang, Z.; Li, T.; Giannakakou, P.; Hamel, E. *Nature* **1997**, *387*, 268–272.
- [10] Nicolaou, K. C.; Vourloumis, D.; Li, T.; Pastor, J.; Winssinger, N.; He, Y.; Ninkovic, S.; Sarabia, F.; Vallberg, H.; Roschangar, F.; King, N. P.; Finlay, M. R. V.; Giannakakou, P.; Verdier-Pinard, P.; Hamel, E. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2097–2103.
- [11] Tempest, P. A.; Armstrong, R. W. *J. Am. Chem. Soc.* **1997**, *119*, 7607–7608.
- [12] Brown, S. D.; Armstrong, R. W. *J. Org. Chem.* **1997**, *62*, 7076–7077.
- [13] Merrifield, R. B., *Adv. Enzymol. Relat. Areas Mol. Biol.*, **1969**, *32*, 221.
- [14] Meyers, H. V.; Dilley, G. J.; Durgin, T. L.; Powers, T. S.; Winssinger, N. A.; Zhu, W.; Pavia, M. R.; *Molecular Diversity*, **1995**, *1*, 13–20.
- [15] Gennari, C.; Nestler, H. P.; Salom, B.; Still, W. C.; *Angew. Chem. Int. Ed. Engl.*, **1995**, *34*, 1763–1765.
- [16] a) Backes, J. A.; Ellman, J. A.; *J. Am. Chem. Soc.*, **1994**, *116*, 11171–11172. b) Han, H.; Wolfe, M. M.; Brenner, S.; Janda, K. D.; *Proc. Natl. Acad. Sci. USA*, **1995**, *92*, 6419–6423.
- [17] Mohan R.; Chou Y-L.; Morrissey M. M., *Tetrahedron Lett.*, **1996**, *37*, 3963–3966.
- [18] Gude M.; Piarulli U.; Potenza D.; Salom B.; Gennari C., *Tetrahedron Lett.*, **1996**, *37*, 8589–8592.
- [19] Bolton G. L., *Tetrahedron Lett.*, **1996**, *37*, 3433–3436.
- [20] de Bont D. B. A.; Moree W. J.; Liskamp R. M. J., *Bioorg. Med. Chem.*, **1996**, *4*, 667–672.
- [21] Watson, S. P.; Wilson, R. D.; Judd, D. B.; Richards, S. A. *Tetrahedron Lett.* **1997**, *38*, 9065–9068.
- [22] Zhong, H. M.; Greco, M. N.; Maryanoff, B. E. *J. Org. Chem.* **1997**, *62*, 9326–9330.
- [23] Raju, B.; Kogan, T. P. *Tetrahedron Letters* **1997**, *38*, 3373–3376.
- [24] García-Echeverría, C. *Tetrahedron Lett.* **1997**, *38*, 8933–8934.
- [25] Kick, E. K.; Ellman, J. A.; *J. Med. Chem.*, **1995**, *38*, 1427–1430.

- [26] Reynolds Cody, D. M.; Hobbs DeWitt, S. H.; Hodges, J. C.; Kiely, J. S.; Moos, W. M.; Pavia, M. R.; Roth, B. D.; Schroeder, M. C.; Stankovic, C. J.; Patent, US 5,324,483, **1994**.
- [27] Hutchins, S. M.; Chapman, K. T.; *Tetrahedron Lett.*, **1994**, *35*, 4055–4058.
- [28] Hutchins, S. M.; Chapman, K. T.; *Tetrahedron Lett.*, **1995**, *36*, 2583–2086.
- [29] Kolodziej S. A.; Hamper B. C.; *Tetrahedron Lett.*, **1996**, *37*, 5277–5280.
- [30] Buckman B. O.; Mohan R.; *Tetrahedron Lett.*, **1996**, *37*, 4439–4442.
- [31] Hanessian S.; Yang R-Y, *Tetrahedron Lett.*, **1996**, *37*, 5835–5838.
- [32] Scialdone M. A.; *Tetrahedron Lett.*, **1996**, *37*, 8141–8144.
- [33] Bauer, U.; Ho, W. B.; Koskinen, A. M. P. *Tetrahedron Lett.* **1997**, *38*, 7233–7236.
- [34] Nefzi, A.; Ostresh, J. M.; Meyer, J. P.; Houghten, R. A. *Tetrahedron Lett.* **1997**, *38*, 931–934.
- [35] Matthews, J.; Rivero, R. A. *J. Org. Chem.* **1997**, *62*, 6090–6092.
- [36] Smith, J.; Liras, J. L.; Schneider, S. E.; Anslyn, E. V. *J. Org. Chem.* **1996**, *61*, 8811–8818.
- [37] Wang, G. T.; Chen, Y.; Wang, S.; Sciotti, R.; Sowin, T. *Tetrahedron Lett.* **1997**, *38*, 1895–1898.
- [38] Xiao, X.; Ngu, K.; Chao, C.; Patel, D. V. *J. Org. Chem.* **1997**, *62*, 6968–6973.
- [39] Corbett, J. W.; Graciani, N. R.; Mousa, S. A.; Degrado, W. F. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1371–1376.
- [40] Routledge, A.; Abell, C.; Balasubramanian, S. *Tetrahedron Lett.* **1997**, *38*, 1227–1230.
- [41] Routledge, A.; Stock, H. T.; Flitsch, S. L.; Turner, N. J. *Tetrahedron Lett.* **1997**, *38*, 8287–8290.
- [42] Paikoff S. J.; Wilson T. E.; Cho C. Y.; Schultz P. G.; *Tetrahedron Lett.*, **1996**, *37*, 5653–5656.
- [43] Hernandez, A. S.; Hodges, J. C. *J. Org. Chem.* **1997**, *62*, 3153–3157.
- [44] Fitzpatrick, L. J.; Rivero, R. A. *Tetrahedron Lett.* **1997**, *38*, 7479–7482.
- [45] Robinson, S.; Roskamp, E. J. *Tetrahedron* **1997**, *53*, 6697–6705.
- [46] Drewry, D. H.; Gerritz, S. W.; Linn, J. A. *Tetrahedron Lett.* **1997**, *38*, 3377–3380.
- [47] Wang, F. J.; Hauske, J. R. *Tetrahedron Lett.* **1997**, *38*, 8651–8654.
- [48] Whitehouse, D. L.; Nelson, K. H.; Savinov, S. N.; Austin, D. J. *Tetrahedron Lett.* **1997**, *38*, 7139–7142.
- [49] Purandare, A. V.; Natarajan, S. *Tetrahedron Lett.* **1997**, *38*, 8777–8780.
- [50] Wang, F. J.; Hauske, J. R. *Tetrahedron Lett.* **1997**, *38*, 6529–6532.
- [51] a) Krchnák, V.; Weichsel, A. S. *Tetrahedron Lett.* **1997**, *38*, 7299–7302. b) Nefzi, A.; Ostresh, J. M.; Houghten, R. A. *Tetrahedron Lett.* **1997**, *38*, 4943–4946.
- [52] Mayer, J. P.; Lewis, G. S.; Curtis, M. J.; Zhang, J. W. *Tetrahedron Lett.* **1997**, *38*, 8445–8448.
- [53] Marti, R. E.; Yan, B.; Jarosinski, M. A. *J. Org. Chem.* **1997**, *62*, 5615–5618.
- [54] Nielsen, J.; Jensen, F. R. *Tetrahedron Lett.* **1997**, *38*, 2011–2014.
- [55] Xiao, X. Y.; Parandoosh, Z.; Nova, M. P. *J. Org. Chem.* **1997**, *62*, 6029–6033.
- [56] Stephensen, H.; Zaragoza, F. *J. Org. Chem.* **1997**, *62*, 6096–6097.
- [57] Fathi R.; Rudolph M. J.; Gentles R. G.; Patel R.; MacMillan E. W.; Reitman M. S.; Pelham D.; Cook A. F.; *J. Org. Chem.*, **1996**, *61*, 5600–5609.
- [58] Goff, D. A.; Zuckermann, R. N.; *J. Org. Chem.*, **1995**, *60*, 5744–5745.
- [59] Boojamra, C. G.; Burow, K. M.; Ellman, J. A.; *J. Org. Chem.*, **1995**, *60*, 5742–5743.
- [60] Frenette, R.; Friesen, R. W.; *Tetrahedron Lett.*, **1994**, *35*, 9177–9180.
- [61] Brown S. D.; Armstrong R. W.; *J. Am. Chem. Soc.*, **1996**, *118*, 6331–6332.
- [62] Guiles J. W.; Johnson S. G.; Murray M. V.; *J. Org. Chem.*, **1996**, *61*, 5169–5171.
- [63] Garigipati R. S.; Adams B.; Adams J. L.; Sarkar S. K.; *J. Org. Chem.*, **1996**, *61*, 2911–2914.
- [64] Han Y.; Walker S. D.; Young R. N.; *Tetrahedron Lett.*, **1996**, *37*, 2703–2706.
- [65] Koh J. S.; Ellman J. A.; *J. Org. Chem.*, **1996**, *61*, 4494–4495.
- [66] Larhed M.; Lindeberg G.; Hallberg A.; *Tetrahedron Lett.*, **1996**, *37*, 8219–8222.
- [67] Chenera B.; Finkelstein J. A.; Veber D. F.; *J. Am. Chem. Soc.*, **1995**, *117*, 11999–12000.
- [68] Ruhland, B.; Bombrun, A.; Gallop, M. A. *J. Org. Chem.* **1997**, *62*, 7820–7826.
- [69] Yoo, S.; Seo, J.; Yi, K.; Gong, Y. *Tetrahedron Lett.* **1997**, *38*, 1203–1206.
- [70] Garigipati, R. S. *Tetrahedron Lett.* **1997**, *38*, 6807–6810.
- [71] Porco, J. A.; Deegan, T.; Devonport, W.; Gooding, O. W.; Heisler, K.; Labadie, J. W.; Newcomb, B.; Nguyen, C.; v. Eikeren, P.; Wong, J.; Wright, P. *Molecular Diversity* **1996**, *2*, 197–206.
- [72] Piettre, S. R.; Baltzer, S. *Tetrahedron Lett.* **1997**, *38*, 1197–1200.
- [73] Fraley, M. E.; Rubino, R. S. *Tetrahedron Lett.* **1997**, *38*, 3365–3368.
- [74] Pavia, M. R.; Whitesides, G. M.; Hangauer, D. G.; Hediger, M. E.; Patent, WO 95/04277, **1995**.
- [75] Deshpande, M. S.; *Tetrahedron Lett.*, **1994**, *35*, 5613–5614.
- [76] Forman, F. W.; Sucholeiki, I.; *J. Org. Chem.*, **1995**, *60*, 523–528.

- [77] a) Bunin, B. A.; Plunkett, M. J.; Ellman, J. A.; *Proc. Natl. Acad. Sci. USA*, **1994**, 91, 4708–4712. b) Ellman, J. A.; Patent, US 5,288,514, 1994. c) Plunkett, M. J.; Ellman, J. A.; *J. Am. Chem. Soc.*, **1995**, 117, 3306–3307.
- [78] Plunkett, M. J.; Ellman, J. A.; *J. Org. Chem.*, **1995**, 60, 6006–6007.
- [79] Beaver K. A.; Siegmund A. C.; Spear K. L.; *Tetrahedron Lett.*, **1996**, 37, 1145–1148.
- [80] Plunkett, M. J.; Ellman, J. A. *J. Org. Chem.* **1997**, 62, 2885–2893.
- [81] Young, J. K.; Nelson, J. C.; Moore, J. S.; *J. Am. Chem. Soc.*, **1994**, 116, 10841–10842.
- [82] Yu, K. L.; Deshpande, M. S.; Vyas, D.; *Tetrahedron Lett.*, **1994**, 35, 8919–8922.
- [83] Hiroshige, M.; Hauske, J. R.; Zhou, P.; *Tetrahedron Lett.*, **1995**, 36, 4567–4570.
- [84] Goff, D. A.; Zuckermann, R. N.; *J. Org. Chem.*, **1995**, 60, 5748–5749.
- [85] Yun W.; Mohan R., *Tetrahedron Lett.*, **1996**, 37, 7189–7192.
- [86] Hiroshige M.; Hauske J. R.; Zhou P., *J. Am. Chem. Soc.*, **1995**, 117, 11590–11591.
- [87] Pop I. E.; Dhalluin C. F.; Deprez B. P.; Melnyk P. C.; Lippens G. M.; Tartar A. L., *Tetrahedron*, **1996**, 52, 12209–12222.
- [88] Collini, M. D.; Ellingboe, J. W. *Tetrahedron Lett.* **1997**, 38, 7963–7966.
- [89] Arumugam, V.; Routledge, A.; Abell, C.; Balasubramanian, S. *Tetrahedron Lett.* **1997**, 38, 6473–6476.
- [90] Fagnola, M. C.; Candiani, I.; Visentin, G.; Cabri, W.; Zarini, F.; Mongelli, N.; Bedeschi, A. *Tetrahedron Lett.* **1997**, 38, 2307–2310.
- [91] Fancelli, D.; Fagnola, M. C.; Severino, D.; Bedeschi, A. *Tetrahedron Lett.* **1997**, 38, 2311–2314.
- [92] Akaji, K.; Kiso, Y. *Tetrahedron Lett.* **1997**, 38, 5185–5188.
- [93] Zhang, H. C.; Maryanoff, B. E. *J. Org. Chem.* **1997**, 62, 1804–1809.
- [94] Zhang, H. C.; Brumfield, K. K.; Maryanoff, B. E. *Tetrahedron Letters* **1997**, 38, 2439–2442.
- [95] Marquais S.; Artl M., *Tetrahedron Lett.*, **1996**, 37, 5491–5494.
- [96] Rottländer, M.; Knochel, P. *Synlett* **1997**, 1084–1086.
- [97] Dankwardt, S. M.; Newman, S. R.; Krstenasky, J. L.; *Tetrahedron Lett.*, **1995**, 36, 4923–4926.
- [98] MacDonald A. A.; DeWitt S. H.; Hogan E. M.; Ramage R., *Tetrahedron Lett.*, **1996**, 37, 4815–4818.
- [99] Phillips G. B.; Wei G. P., *Tetrahedron Lett.*, **1996**, 37, 4887–4890.
- [100] Shapiro M. J.; Kumaravel G.; Petter R. C.; Beveridge R., *Tetrahedron Lett.*, **1996**, 37, 4671–4674.
- [101] Stankova M.; Lebl M., *Molecular Diversity*, **1996**, 2, 75–80.
- [102] Lee, J.; Murray, W. V.; Rivero, R. A. *J. Org. Chem.* **1997**, 62, 3874–3879.
- [103] Gray, N. S.; Kwon, S.; Schultz, P. G. *Tetrahedron Lett.* **1997**, 38, 1161–1164.
- [104] Nugiel, D. A.; Cornelius, L. A. M.; Corbett, J. W. *J. Org. Chem.* **1997**, 62, 201–203.
- [105] Ward Y. D.; Farina V., *Tetrahedron Lett.*, **1996**, 37, 6993–6996.
- [106] Willoughby C. A.; Chapman K. T., *Tetrahedron Lett.*, **1996**, 37, 7181–7184.
- [107] Gordon, D. W.; Steele, J.; *Bioorg. Med. Chem. Lett.*, **1995**, 5, 47–50.
- [108] Kowalski J.; Lipton M. A., *Tetrahedron Lett.*, **1996**, 37, 5839–5840.
- [109] Dressman B. A.; Spangle L. A.; Kaldor S. W., *Tetrahedron Lett.*, **1996**, 37, 937–940.
- [110] Gouilleux L.; Fehrentz J-A.; Winternitz F.; Martinez J., *Tetrahedron Lett.*, **1996**, 37, 7031–7034
- [111] Smith A. L.; Thompson C. G.; Leeson P. D., *Bioorg. Med. Chem. Lett.*, **1996**, 6, 1483–1486.
- [112] Gordeev M. F.; Patel D. V.; Gordon E. M., *J. Org. Chem.*, **1996**, 61, 924–928.
- [113] van Maarseveen J. H.; den Hartog J. A. J.; Engelen V.; Finner E.; Visser G.; Kruse C. G., *Tetrahedron Lett.*, **1996**, 37, 8249–8252.
- [114] Mayer J. P.; Zhang J.; Biergarde K.; Lenz D. M.; Gaudino J. J., *Tetrahedron Lett.*, **1996**, 37, 8081–8084.
- [115] Virgilio A. A.; Schürer S. C.; Ellman J. A., *Tetrahedron Lett.*, **1996**, 37, 6961–6964.
- [116] Tietze L. F.; Steinmetz A., *Synlett.*, **1996**, 66–67.
- [117] Szardenings, A. K.; Burkoth, T. S.; Lu, H. H.; Tien, D. W.; Campbell, D. A. *Tetrahedron* **1997**, 53, 6573–6593.
- [118] Le Hetet, C.; David, M.; Carreaux, F.; Carboni, B.; Sauleau, A. *Tetrahedron Lett.* **1997**, 38, 5153–5156.
- [119] Tietze, L. F.; Steinmetz, A.; Balkenhohl, F. *Bioorg. Med. Chem. Lett.* **1997**, 7, 1303–1306.
- [120] Kim, S. W.; Ahn, S. Y.; Koh, J. S.; Lee, J. H.; Ro, S.; Cho, H. Y. *Tetrahedron Lett.* **1997**, 38, 4603–4606.
- [121] Piscopio, A. D.; Miller, J. F.; Koch, K. *Tetrahedron Lett.* **1997**, 38, 7143–7146.
- [122] Pernerstorfer, J.; Schuster, M.; Blechert, S. *Chem. Commun.* **1997**, 1949–1950.

- [123] Sucholeiki, I.; *Tetrahedron Lett.*, **1994**, *35*, 7307–7310.
- [124] Boehm T. L.; Showalter H. D. H.; *J. Org. Chem.*, **1996**, *61*, 6498–6499.
- [125] Hughes I.; *Tetrahedron Lett.*, **1996**, *37*, 7595–7598.
- [126] Woolard, F. X.; Paetsch, J.; Ellman, J. A. *J. Org. Chem.* **1997**, *62*, 6102–6103.
- [127] Newlander, K. A.; Chenera, B.; Veber, D. F.; Yim, N. C. F.; Moore, M. L. *J. Org. Chem.* **1997**, *62*, 6726–6732.
- [128] Gowravaram, M. R.; Gallop, M. A. *Tetrahedron Lett.* **1997**, *38*, 6973–6976.
- [129] Zaragoza, F. *Tetrahedron Lett.* **1997**, *38*, 7291–7294.
- [130] Schuster, M.; Lucas, N.; Blechert, S. *Chem. Commun.* **1997**, 823–824.
- [131] a) Beebe, X.; Schore, N. E.; Kurth, M. J.; *J. Am. Chem. Soc.*, **1992**, *114*, 10061–10062. b) Beebe, X.; Schore, N. E.; Kurth, M. J.; *J. Org. Chem.*, **1995**, *60*, 4196–4203.
- [132] Moon, H. S.; Schore, N. E.; Kurth, M. J.; *J. Org. Chem.*, **1992**, *57*, 6088–6089.
- [133] Hunt J. A.; Roush W. R.; *J. Am. Chem. Soc.*, **1996**, *118*, 9998–9999.
- [134] Nelson J. C.; Young J. K.; Moore J. S.; *J. Org. Chem.*, **1996**, *61*, 8160–8168.
- [135] Kawai K.; Onta H.; Channing M. A.; Kubodera A.; Eckelman W. C.; *Appl. Radiat. Isot.*, **1996**, *47*, 37–44.
- [136] Johnson C. R.; Zhang B.; *Tetrahedron Lett.*, **1995**, *36*, 9253–9256.
- [137] Peters, J. U.; Blechert, S. *Synlett* **1997**, 348–350.
- [138] Halm, C.; Everts, J.; Kurth, M. J. *Tetrahedron Lett.* **1997**, *38*, 7709–7712.
- [139] Thompson, L. A.; Ellman, J. A.; *Tetrahedron Lett.*, **1994**, *35*, 9333–9336.
- [140] Kurth, M. J.; Randall, L. A.; Chen, C.; Melander, C.; Miller, R. B.; McAllister, K.; Reitz, G.; Kang, R.; Nakatsu, T.; Green, C.; *J. Org. Chem.*, **1994**, *59*, 5862–5864.
- [141] Wang, G. T.; Li, S.; Wideburg, N.; Krafft, G. A.; Kempf, D. J.; *J. Med. Chem.*, **1995**, *38*, 2995–3002.
- [142] Ley, S. V.; Mynett, D. M.; Koot, W.-J.; *Synlett.*, **1995**, 1017–1020.
- [143] Tietze L. F.; Steinmetz A. *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 651–652.
- [144] Wess G.; Bock K.; Kleine H.; Kurz M.; Guba W.; Hemmerle H.; Lopez-Calle E.; Baringhaus K.-H.; Glombik H.; *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 2222–2224.
- [145] Liu G.; Ellman J. A.; *J. Org. Chem.* **1995**, *60*, 7712–7713.
- [146] Kobayashi S.; Hachiya I.; Suzuki S.; Moriwaki M.; *Tetrahedron Lett.*, **1996**, *37*, 2809–2812.
- [147] Kobayashi S.; Hachiya I.; Maaru Y.; *Tetrahedron Lett.*, **1996**, *37*, 5569–5572.
- [148] a) Kobayashi S.; Moriwaki M.; Akiyama R.; Suzuki I.; Hachiya I.; *Tetrahedron Lett.*, **1996**, *37*, 7783–7786. b) Kobayashi, S.; Moriwaki, M. *Tetrahedron Lett.* **1997**, *38*, 4251–4254.
- [149] Deegan, T. L.; Gooding, O. W.; Baudart, S.; Porco, J. A. *Tetrahedron Lett.* **1997**, *38*, 4973–4976.
- [150] Burgess, K.; Lim, D. *Chem. Commun.* **1997**, 785–786.
- [151] Panek, J. S.; Zhu, B. *J. Am. Chem. Soc.* **1997**, *119*, 12022–12023.
- [152] Armstrong, R. W.; Patent, **1995**, WO 95/02566. .
- [153] Zaragoza F.; *Tetrahedron Lett.*, **1995**, *36*, 8677–8680.
- [154] Kaljuste K.; Linden A.; *Tetrahedron Lett.*, **1995**, *36*, 9211–9214.
- [155] Bannwarth W.; Huebscher J.; Barner R.; *Bioorg. Med. Chem. Lett.*, **1996**, *6*, 1525–1528.
- [156] Kaljuste K.; Unden A.; *Tetrahedron Lett.*, **1996**, *37*, 3031–3034.
- [157] Kay, C.; Murray, P. J.; Sandow, L.; Holmes, A. B. *Tetrahedron Lett.* **1997**, *38*, 6941–6944.
- [158] Conti, P.; Demont, D.; Cals, J.; Ottenheijm, H. C. J.; Leysen, D. *Tetrahedron Lett.* **1997**, *38*, 2915–2918.
- [159] Furth, P. S.; Reitman, M. S.; Cook, A. F. *Tetrahedron Lett.* **1997**, *38*, 5403–5406.
- [160] Furth, P. S.; Reitman, M. S.; Gentles, R.; Cook, A. F. *Tetrahedron Lett.* **1997**, *38*, 6643–6646.
- [161] Morphy J. R.; Rankovic Z.; Rees D. C.; *Tetrahedron Lett.*, **1996**, *37*, 3209–3212.
- [162] Ho, C. Y.; Kukla, M. J. *Tetrahedron Lett.* **1997**, *38*, 2799–2802.
- [163] Brown, A. R.; Rees, D. C.; Rankovic, Z.; Morphy, J. R. *J. Am. Chem. Soc.* **1997**, *119*, 3288–3295.
- [164] Heinonen, P.; Lönnberg, H. *Tetrahedron Lett.* **1997**, *38*, 8569–8572.
- [165] Kroll, F. E. K.; Morphy, R.; Rees, D.; Gani, D. *Tetrahedron Lett.* **1997**, *38*, 8573–8576.
- [166] Murphy, A. M.; Dagnino, R.; Vallar, P. L.; Trippe, A. J.; Sherman, S. L.; Lumpkin, R. H.; Tamura, S. Y.; Webb, T. R.; *J. Am. Chem. Soc.*, **1992**, *114*, 3156–3157.
- [167] Fehrentz, J. A.; Paris, M.; Heitz, A.; Velek, J.; Liu, C.-F.; Winternitz, F.; Martinez, J.; *Tetrahedron Lett.*, **1995**, *36*, 7871–7874.
- [168] Dinh, T. Q.; Armstrong R. W.; *Tetrahedron Lett.*, **1996**, *37*, 1161–1164.
- [169] Vlattas, I.; Dellureficio, J.; Dunn, R.; Sytwu, I. I.; Stanton, J. *Tetrahedron Lett.* **1997**, *38*, 7321–7324.
- [170] Ede, N. J.; Bray, A. M. *Tetrahedron Lett.* **1997**, *38*, 7119–7122.



- [171] Pothion, C.; Paris, M.; Heitz, A.; Rocheblave, L.; Rouch, F.; Fehrentz, J. A.; Martinez, J. *Tetrahedron Lett.* **1997**, *38*, 7749–7752.
- [172] Fehrentz, J. A.; Paris, M.; Heitz, A.; Velek, J.; Winternitz, F.; Martinez, J. *J. Org. Chem.* **1997**, *62*, 6792–6796.
- [173] Hird, N. W.; Irie, K.; Nagai, K. *Tetrahedron Lett.* **1997**, *38*, 7111–7114.
- [174] Crawshaw, M.; Hird, N. W.; Irie, K.; Nagai, K. *Tetrahedron Lett.* **1997**, *38*, 7115–7118.
- [175] Richter, L. S.; Desai, M. C. *Tetrahedron Lett.* **1997**, *38*, 321–322.
- [176] Ngu, K.; Patel, D. V. *J. Org. Chem.* **1997**, *62*, 7088–7089.
- [177] Mellor, S. L.; Chan, W. C. *Chem. Commun.* **1997**, 2005–2006.
- [178] Roussel, P.; Bradley, M.; Matthews, I.; Kane, P. *Tetrahedron Lett.* **1997**, *38*, 4861–4864.
- [179] Obrecht, D.; Abrecht, C.; Grieder, A.; Villalgordo, J. M. *Helvetica Chimica Acta* **1997**, *80*, 65–72.
- [180] Gayo, L. M.; Suto, M. J. *Tetrahedron Lett.* **1997**, *38*, 211–214.
- [181] Backes B. J.; Virgilio A. A.; Ellman J. A., *J. Am. Chem. Soc.*, **1996**, *118*, 3055–3056.
- [182] Yang L.; Guo L., *Tetrahedron Lett.*, **1996**, *37*, 5041–5044.
- [183] Barn D. R.; Morphy R. J.; Rees D. C., *Tetrahedron Lett.*, **1996**, *37*, 3213–3216.
- [184] Fivush, A. M.; Willson, T. M. *Tetrahedron Lett.* **1997**, *38*, 7151–7154.
- [185] Prien, O.; Röfling, K.; Thiel, M.; Künzer, H. *Synlett* **1997**, 325–326.
- [186] Boojamra, C. G.; Burow, K. M.; Thompson, L. A.; Ellman, J. A. *J. Org. Chem.* **1997**, *62*, 1240–1256.
- [187] Marzinzik A. L.; Felder E. R., *Tetrahedron Lett.*, **1996**, *37*, 1003–1006
- [188] Marzinzik, A. L.; Felder, E. R. *Molecules* **1997**, *2*, 17–30.
- [189] Schuster, M.; Wang, P.; Paulson, J. C.; Wong, C-H.; *J. Am. Chem. Soc.*, **1994**, *116*, 1135–1136.
- [190] Gordeev M. F.; Patel D. V.; Wu J.; Gordon E. M., *Tetrahedron Lett.*, **1996**, *37*, 4643–4646.
- [191] Hollinshead S. P., *Tetrahedron Lett.*, **1996**, *37*, 9157–9160.
- [192] Tietze L. F.; Hippe T.; Steinmetz, A., *Synlett*, **1996**, 1043–1044.
- [193] Youngman, M. A.; Dax, S. L. *Tetrahedron Lett.* **1997**, *38*, 6347–6350.
- [194] Zaragoza F.; Peterson S. V., *Tetrahedron*, **1996**, *52*, 10823–10826.
- [195] Ho, P. T.; Chang, D.; Zhong, J. W. X.; Musso, G. F.; *Peptide Research*, **1993**, *6*, 10–12.
- [196] Green, J.; *J. Org. Chem.*, **1995**, *60*, 4287–4290.
- [197] Murphy, M. M.; Schullek, J. R.; Gordon, E. M.; Gallop, M. A.; *J. Am. Chem. Soc.*, **1995**, *117*, 7029–7030.
- [198] Look, G. C.; Murphy, M. M.; Campell, D. A.; Gallop, M. A.; *Tetrahedron Lett.*, **1995**, *36*, 2937–2940.
- [199] Bray, A. M.; Chiefari, D. S.; Valerio, R. M.; Maeji, N. J.; *Tetrahedron Lett.*, **1995**, *36*, 5081–5084.
- [200] Look G. C.; Schullek J. R.; Holmes C. P.; Chinn J. P.; Gordon E. M.; Gallop M. A., *Bioorg. Med. Chem. Lett.*, **1996**, *6*, 707–712
- [201] Boyd E. A.; Chan W. C.; Loh V. M., *Tetrahedron Lett.*, **1996**, *37*, 1647–1650.
- [202] Ruhland B.; Bhandari A.; Gordon E. M.; Gallop M. A., *J. Am. Chem. Soc.*, **1996**, *118*, 253–254.
- [203] Pearson, W. H.; Clark, R. B. *Tetrahedron Lett.* **1997**, *38*, 7669–7672.
- [204] Patek, M.; Drake, B.; Lebl, M.; *Tetrahedron Lett.*, **1995**, *36*, 2227–2230.
- [205] Look, G. C.; Holmes, C. P.; Chinn, J. P.; Galop, M. A.; *J. Org. Chem.*, **1994**, *59*, 7588–7590.
- [206] Gordeev, M. F.; Gordon, E. M.; Patel, D. V. *J. Org. Chem.* **1997**, *62*, 8177–8181.
- [207] Pei, Y. Z.; Houghten, R. A.; Kiely, J. S. *Tetrahedron Lett.* **1997**, *38*, 3349–3352.
- [208] Pei, Y.; Moos, W. M.; *Tetrahedron Lett.*, **1994**, *35*, 5825–5828.
- [209] Hamper B. C.; Dukeshere D. R.; South M. S., *Tetrahedron Lett.*, **1996**, *37*, 3671–3674.
- [210] Bicknell A. J.; Hird N. W., *Bioorg. Med. Chem. Lett.*, **1996**, *6*, 2441–2444.
- [211] Mjalli A. M. M.; Sarshar S.; Baiga T. G., *Tetrahedron Lett.*, **1996**, *37*, 2943–2946.
- [212] Lorsbach, B. A.; Miller, R. B.; Kurth, M. J. *J. Org. Chem.* **1996**, *61*, 8716–8717.
- [213] Kurth, M. J.; Randall, L. A. A.; Takenouchi, K. *J. Org. Chem.* **1996**, *61*, 8755–8761.
- [214] Marx, M. A.; Grillot, A. L.; Louer, C. T.; Beaver, K. A.; Bartlett, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 6153–6167.
- [215] Kantorowski, E. J.; Kurth, M. J. *J. Org. Chem.* **1997**, *62*, 6797–6803.
- [216] Ritter, H.; Sperber, R.; *Macromolecules*, **1994**, *27*, 5919–5920.
- [217] Panek J. S.; Zhu B., *Tetrahedron Lett.*, **1996**, *37*, 8151–8154.
- [218] Wallace, O. B. *Tetrahedron Lett.* **1997**, *38*, 4939–4942.
- [219] Katritzky, A. R.; Xie, L. H.; Zhang, G. F.; Griffith, M.; Watson, K.; Kiely, J. S. *Tetrahedron Lett.* **1997**, *38*, 7011–7014.
- [220] Wipf, P.; Henninger, T. C. *J. Org. Chem.* **1997**, *62*, 1586–1587.
- [221] Schuster M.; Pernerstorfer J.; Blechert S., *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 1971–1980.

- [222] Strocker A. M.; Keating T. A.; Tempest P. A.; Armstrong R. W., *Tetrahedron Lett.*, **1996**, 37, 1149–1152.
- [223] Gopalsamy, A.; Pallai, P. V. *Tetrahedron Lett.* **1997**, 38, 907–910.
- [224] Cuny, G. D.; Cao, J. R.; Hauske, J. R. *Tetrahedron Lett.* **1997**, 38, 5237–5240.
- [225] Meutermsans, W. D. F.; Alewood, P. F.; *Tetrahedron Lett.*, **1995**, 36, 7709–7712.
- [226] Rölfling K.; Thiel M.; Künzer H., *SynLett.*, **1996**, 1036–1038.
- [227] Bolton, G. L.; Hodges, J. C.; Rubin, J. R. *Tetrahedron* **1997**, 53, 6611–6634.
- [228] Hutchins S. M.; Chapman K. T., *Tetrahedron Lett.*, **1996**, 37, 4869–4872.
- [229] Kim R. M.; Manna M.; Hutchins S. M.; Griffin P. R.; Yates N. A.; Bernick A. M.; Chapman K. T., *Proc. Natl. Acad. Sci. (USA)*, **1996**, 93, 10012–10017.
- [230] Mayer J. P.; Bankaitis-Davis D.; Zhang J.; Beaton G.; Bjergarde K.; Andersen C. M.; Goodman B. A.; Herrera C. J., *Tetrahedron Lett.*, **1996**, 37, 5633–5636.
- [231] Mohan R.; Chou Y-L.; Morrissey M. M., *Tetrahedron Lett.*, **1996**, 37, 3963–3966.
- [232] Hutchins S. M.; Chapman K. T., *Tetrahedron Lett.*, **1996**, 37, 4865–4868.
- [233] Zhang C.; Moran E. J.; Woiwode T. F.; Short K. M.; Mjalli A. M. M., *Tetrahedron Lett.*, **1996**, 37, 751–754.
- [234] Yan B.; Gstach H., *Tetrahedron Lett.*, **1996**, 37, 8325–8328.
- [235] Bhalay, G.; Blaney, P.; Palmer, V. H.; Baxter, A. D. *Tetrahedron Lett.* **1997**, 38, 8375–8378.
- [236] Cheng, Y.; Chapman, K. T. *Tetrahedron Lett.* **1997**, 38, 1497–1500.
- [237] Deans, R.; Rotello, V. M. *J. Org. Chem.* **1997**, 62, 4528–4529.
- [238] Gordeev, M. F.; Hui, H. C.; Gordon, E. M.; Patel, D. V. *Tetrahedron Lett.* **1997**, 38, 1729–1732.
- [239] Heizmann, G.; Eberle, A. N. *Molecular Diversity* **1997**, 2, 171–174.
- [240] Du, X. H.; Armstrong, R. W. *J. Org. Chem.* **1997**, 62, 5678–5679.
- [241] Zaragoza F., *Tetrahedron Lett.*, **1996**, 37, 6213–6216.
- [242] Holmes C. P.; Chinn J. P.; Gordon E. M.; Gallop M. A., *J. Org. Chem.*, **1995**, 60, 7328–7333.
- [243] Chen, C.; Ahlberg Randall, L. A.; Miller, R. B.; Jones, A. D.; Kurth, M. J.; *J. Am. Chem. Soc.*, **1994**, 116, 2661–2662.
- [244] Chen, C. X.; Randall, L. A. A.; Miller, R. B.; Jones, A. D.; Kurth, M. J. *Tetrahedron* **1997**, 53, 6595–6609.
- [245] Goff D. A.; Zuckermann R. N., *Tetrahedron Lett.*, **1996**, 37, 6247–6250.
- [246] Swali, V.; Wells, N. J.; Langley, G. J.; Bradley, M. *J. Org. Chem.* **1997**, 62, 4902–4903.
- [247] Spitzer, J. L.; Kurth, M. J.; Schore, N. E.; Najdis, D. *Tetrahedron* **1997**, 53, 6791–6808.
- [248] Sibi, M. P.; Chandramouli, S. V. *Tetrahedron Lett.* **1997**, 38, 8929–8932.
- [249] Vo, N. H.; Eyermann, C. J.; Hodge, C. N. *Tetrahedron Lett.* **1997**, 38, 7951–7954.
- [250] Zaragoza F.; Petersen V., *Tetrahedron Lett.*, **1996**, 37, 5999–6002.
- [251] Rotella D. P., *J. Am. Chem. Soc.*, **1996**, 118, 12246–12247.
- [252] Williard, R.; Jammalamadaka, V.; Zava, D.; Benz, C. C.; Hunt, C. A.; Kusher, P. J.; Scanlan, T. S.; *Chemistry and Biology*, **1995**, 2, 45–51.
- [253] Vagner, J.; Krchnak, V.; Lebl, M.; Barany, G. *Collection of Czechoslovak Chemical Communications* **1996**, 61, 1697–1702.
- [254] Gosselin F.; DiRenzo M.; Ellis T. H.; Lubell W. D., *J. Org. Chem.*, **1996**, 61, 7980–7981.
- [255] Yan, B.; Sun, Q.; Wareing, J. R.; Jewell, C. F. *J. Org. Chem.* **1996**, 61, 8765–8770.
- [256] Hoekstra, W. J.; Greco, M. N.; Yabut, S. C.; Hulshizer, B. L.; Maryanoff, B. E. *Tetrahedron Lett.* **1997**, 38, 2629–2632.
- [257] Yan, B.; Li, W. *J. Org. Chem.* **1997**, 62, 9354–9357.
- [258] Mata, E. G. *Tetrahedron Lett.* **1997**, 38, 6335–6338.
- [259] Danishefsky, S. J.; McClure, K. F.; Randolph, J. T.; Ruggeri, R. B.; *Science*, **1993**, 260, 1307–1309.
- [260] Ajayaghosh, A.; Rajasekharan Pillai, V. N.; *Tetrahedron Lett.*, **1995**, 36, 777–780.
- [261] Katritzky, A. R.; Toader, D.; Watson, K.; Kiely, J. S. *Tetrahedron Lett.* **1997**, 38, 7849–7850.
- [262] Kiselyov, A. S.; Armstrong, R. W. *Tetrahedron Lett.* **1997**, 38, 6163–6166.
- [263] Tortolani D. R.; Biller S. A., *Tetrahedron Lett.*, **1996**, 37, 5687–5690.
- [264] Kim J-M.; Bi Y.; Paikoff S. J.; Schultz P. G., *Tetrahedron Lett.*, **1996**, 37, 5305–5308.
- [265] Liang R.; Yan L.; Loebach J.; Ge M.; Uozumi Y.; Sekanina K.; Horan N.; Gildersleeve J.; Thompson C.; Smith A.; Biswas K.; Still W. C.; Kahne D., *Science*, **1996**, 274, 1520–1522.
- [266] Brown, P. J.; Hurley, K. P.; Stuart, L. W.; Willson, T. M. *Synthesis* **1997**, 778–782.
- [267] Kracnak V.; Weichsel A. S.; Cabel D.; Flegelova Z.; Lebl M., *Molecular Diversity*, **1995**, 1, 149–164.
- [268] Szardenings A. K.; Burkoth T. S.; Look G. C.; Campbell D. A., *J. Org. Chem.* **1996**, 61, 6720–6722.

- [269] Swayze, E. E. *Tetrahedron Lett.* **1997**, 38, 8643–8646.
- [270] Brown, E. G.; Nuss, J. M. *Tetrahedron Lett.* **1997**, 38, 8457–8460.
- [271] Swayze, E. E. *Tetrahedron Lett.* **1997**, 38, 8465–8468.
- [272] Devraj, R.; Cushman, M. J. *Org. Chem.* **1996**, 61, 9368–9373.
- [273] Khan N. M.; Arumugam V.; Balasubramanian S., *Tetrahedron Lett.*, **1996**, 37, 4819–4822.
- [274] Semenov, A. N.; Gordeev, K. Y.; *Int. J. Peptide Protein Res.*, **1995**, 45, 303–304.
- [275] Ojima, I.; Tsai, C-Y.; Zhang, Z.; *Tetrahedron Lett.*, **1994**, 35, 5785–5788.
- [276] Virgilio, A. A.; Ellman, J. A.; *J. Am. Chem. Soc.*, **1994**, 116, 11 580–11 581.
- [277] Moon, H-S.; Schore, N. E.; Kurth, M. J.; *Tetrahedron Lett.*, **1994**, 35, 8915–8918.
- [278] Reggelin M.; Brenig V., *Tetrahedron Lett.*, **1996**, 37, 6851–6852.
- [279] Ruhland T.; Künzer H., *Tetrahedron Lett.*, **1996**, 37, 2757–2760.
- [280] O'Donnell M. J.; Zhou C.; Scott W. L., *J. Am. Chem. Soc.*, **1996**, 118, 6070–6071.
- [281] Allin S. M.; Shuttlesworth S. J., *Tetrahedron Lett.*, **1996**, 37, 8023–8026.
- [282] Griffith, D. L.; O' Donnell, M. J.; Pottorf, R. S.; Scott, W. L.; Porco, J. A. *Tetrahedron Lett.* **1997**, 38, 8821–8824.
- [283] O'Donnell, M. J.; Lugar, C. W.; Pottorf, R. S.; Zhou, C. Y.; Scott, W. L.; Cwi, C. L. *Tetrahedron Lett.* **1997**, 38, 7163–7166.
- [284] Douglas, S. P.; Whitfield, D. M.; Krepinsky, J. J.; *J. Am. Chem. Soc.*, **1995**, 117, 2116–2117.
- [285] Richter, L. S.; Gadek, T. R.; *Tetrahedron Lett.*, **1994**, 35, 4705–4706.
- [286] Rano, T. A.; Chapman, K. T.; *Tetrahedron Lett.*, **1995**, 36, 3789–3792.
- [287] Krchnak, V.; Flegelova, Z.; Weichsel, A. S.; Lebl, M.; *Tetrahedron Lett.*, **1995**, 36, 6193–6196.
- [288] Campbell, D. A.; Bermak, J. C.; Burkoth, T. S.; Patel, D. V.; *J. Am. Chem. Soc.*, **1995**, 117, 5381–5382.
- [289] Pavia M. R.; Cohen M. P.; Dilley G. J.; Dubuc G. R.; Durgin T. L.; Forman F. W.; Hediger M. E.; Milot G.; Powers T. S.; Sucholeiki I.; Zhou S.; Hangauer D. G., *Bioorg. Med. Chem.*, **1996**, 4, 659–666.
- [290] Rademann J.; Schmidt R. R., *Tetrahedron Lett.*, **1996**, 37, 3989–3990.
- [291] Wang Z. -G.; Douglas S. P.; Krepinsky J. J., *Tetrahedron Lett.*, **1996**, 37, 6985–6988.
- [292] Ito Y.; Kanie O.; Ogawa T., *Angew. Chem. Int. Ed. Engl.*, **1996**, 35, 2510–2512.
- [293] Valerio R. M.; Bray A. M.; Patsiouras H., *Tetrahedron Lett.*, **1996**, 37, 3019–3022.
- [294] Krchnak V.; Weichsel A. S.; Issakova O.; Lam K. S.; Lebl M., *Molecular Diversity*, **1995**, 1, 177–182.
- [295] Rodebaugh, R.; Fraser-Reid, B.; Geysen, H. M. *Tetrahedron Lett.* **1997**, 38, 7653–7656.
- [296] Zuckermann, R. N.; Martin, E. J.; Spellmeyer, D. C.; Stauber, G. B.; Shoemaker, K. R.; Kerr, J. M.; Figliozzi, G. M.; Goff, D. A.; Siani, M. A.; Simon, R. J.; Banville, S. C.; Brown, E. G.; Wang, L.; Richter, L. S.; Moos, W.H.; *J. Med. Chem.*, **1994**, 37, 2678–2685.
- [297] Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H.; *J. Am. Chem. Soc.*, **1992**, 114, 10646–10649.
- [298] Ostresh, J. M.; Husar, G. M.; Blondelle, S. E.; Dorner, B.; Weber, P. A.; Houghten, R. A.; *Proc. Natl. Acad. Sci. USA*, **1994**, 91, 11 138–11 142.
- [299] Flegelova Z.; Patek M., *J. Org. Chem.*, **1996**, 61, 6735–6738.
- [300] Ngu, K.; Patel, D. V. *Tetrahedron Letters* **1997**, 38, 973–976.
- [301] Dankwardt, S. M.; Smith, D. B.; Porco, J. A.; Nguyen, C. H. *Synlett* **1997**, 854–855.
- [302] Yang, L.; Chiu, K. *Tetrahedron Lett.* **1997**, 38, 7307–7310.
- [303] Cao X.; Mjalli A. M. M., *Tetrahedron Lett.*, **1996**, 37, 6073–6076.
- [304] Zhang C.; Mjalli A. M. M., *Tetrahedron Lett.*, **1996**, 37, 5457–5460.

## 3 Solution-Phase Combinatorial Chemistry

*Volkhard Austel*

### 3.1 Comparison with Solid-Phase Combinatorial Synthesis

Solution-phase combinatorial chemistry so far has played a considerably lesser role than its solid-phase counterpart. This is probably due to the main problem of solution-phase combinatorial synthesis, *i.e.*, to obtain pure products. In solid-phase synthesis, components such as auxiliary reagents and unreacted starting materials can be easily separated from the desired products by simple washing procedures since both reside in different phases. In solution-phase synthesis, all components occur in the same phase so that purification becomes a much more demanding task. With respect to side products derived from the resin-bound reaction component, the purification problems are, however, the same in both solution- and solid-phase synthesis.

Presently, solution-phase combinatorial synthesis seems to attract growing interest which may arise from a number of advantages offered by this technique:

- In solution-phase synthesis, all reactive groups of the starting materials are available for structural modifications, whereas in solid-phase one such group needs to be sacrificed for the attachment to the solid support. This reduces the potential size and the achievable diversity of the libraries. Moreover, in particular with respect to the application in medicinal chemistry, the attachment groups are often highly polar (*e.g.*, carboxyl, amino, aminocarbonyl) and may therefore obscure interaction potentials of other groups in the molecule or even reduce binding to the target to a non-detectable level. Even though linkers that do not leave functional groups behind have been developed, their application often requires a number of additional synthetic steps. In cases where the functional group that is produced on cleavage from the resin is required for a specific purpose, *e.g.*, for binding to a biological target, the sacrifice of this group for attachment is not a disadvantage.
- Normally, combinatorial reaction paths are primarily worked out in solution and only then adapted to solid phase. Therefore, developing a solid-phase synthesis may often require more time.
- Solution-phase synthesis is not subject to constraints arising from limitations of the thermal or chemical stability of the resin or the linker.
- Solution-phase synthesis is shorter by one to two steps, depending on whether or not a separate cleavage step is involved in solid-phase synthesis. This does not apply if the solid phase fulfils the role of a protective group.

- Unlike on solid phase, syntheses in solution can be more conveniently done in convergent fashion.
- Reactions on solid phase often need considerably more time than in solution. Therefore, sluggish reactions and less active reagents are frequently unsuitable for solid-phase chemistry. The reactions can be speeded up by using large excesses of reagents or by heating, even though the latter measure finds its limits by the thermal stability of the solid support. There are, however, also reactions that proceed reasonably well only on solid support, *e.g.*, various intramolecular reactions.
- Reactions that involve insoluble components, *e.g.*, hydrogenations with the conventional insoluble catalysts, are confined to solution phase.
- Solution-phase reactions can be followed conveniently by simple means such as TLC, conventional NMR or UV, whereas solid-phase synthesis requires either prior cleavage from the support or special equipment, *e.g.* ATR (attenuated total reflection)-IR or MAS (magic angle spinning)-NMR. This difference affects primarily the development of new reactions but may also be relevant for the actual library synthesis, *e.g.*, checking completion of a reaction or the quality of intermediates.
- In general, reaction volumes in relation to the amount of product are smaller in solution- than in solid-phase synthesis. This facilitates the synthesis of common intermediates in larger quantities.
- Finally, the need for using excess reagents and for thorough washing in solid-phase chemistry produces considerable waste which needs to be disposed of. With further miniaturization of biological testing and chemical synthesis, this factor may become less important.

Combinatorial synthesis is only feasible if the same reaction conditions can be applied to a broad spectrum of reactants, even if they are structurally 'diverse' and may therefore differ to a larger extent in their reactivity. Moreover, irrespective of whether combinatorial chemistry is done on solid support or in solution, it is desirable to have reactions come to completion with respect to at least one of the components. With solid-phase synthesis this can mostly be achieved by using a large excess of those reactants that are added in solution. Applying such a procedure to solution synthesis is often limited by difficulties encountered in removing the surplus reagent(s).

Even though there are principally no constraints on solution-phase synthesis as far as the nature of the reagents (highly reactive and aggressive chemicals do not cause problems) or the reaction conditions (particularly high temperatures) are concerned the purification issue imposes limits on the type of reactions that are suitable. Ideally, such reactions should produce only negligible amounts of side products, do not need auxiliary reagents, and come to completion with equimolar amounts of the components. In practice, the constraints are not quite as stringent. Thus, excess reagents that are volatile (*e.g.*, hydrogen in catalytic reductions) or auxiliaries that are insoluble under the reaction conditions (*e.g.*, Pd catalysts for hydrogenations or potassium carbonate as an auxiliary base) are easily removed. As will be discussed later, more and more techniques which lend themselves for automated purification of products synthesized in solution are being developed, and these widen the spectrum of reactions that can be used in solution-phase combinatorial chemistry.

Generally, one needs to take into consideration that the significance of purity may vary with the purpose for which a combinatorial library is produced. If the primary aim is to search for new lead compounds in high-throughput screening (HTS), the demands on purity may not be too high, as suggested by the success of screening extracts from natural sources such as plants or fermentation broths. In particular, the outcome of binding assays should be comparatively little affected by impurities. With functional assays, there is a risk that impurities cancel the effect of the desired synthesis product, but this risk may be tolerable. Moreover, combinatorial libraries synthesized for screening purposes are normally composed of a large number of more or less similar compounds. Therefore, it is to be expected that if active compounds are present, they should fit into some structure–activity pattern which ought to be recognizable, even if some of the biological data are obscured by impurities.

The situation may be different for lead optimization where more reliable biological data are required, *e.g.*, for deriving more detailed structure–activity relationships. At that stage, the libraries mostly comprise fewer compounds so that erroneous biological data have a much higher impact.

Nevertheless, the achievable purity will remain an important criterion for whether or not a reaction is suitable for solution-phase combinatorial synthesis. The extent to which this criterion can be fulfilled depends on the uniformity of the reaction and on the efficiency with which the products can be purified. Therefore, this chapter will focus on reactions that have been more or less successfully applied to solution-phase combinatorial chemistry (including parallel synthesis), and particularly on the purification techniques that have been used.

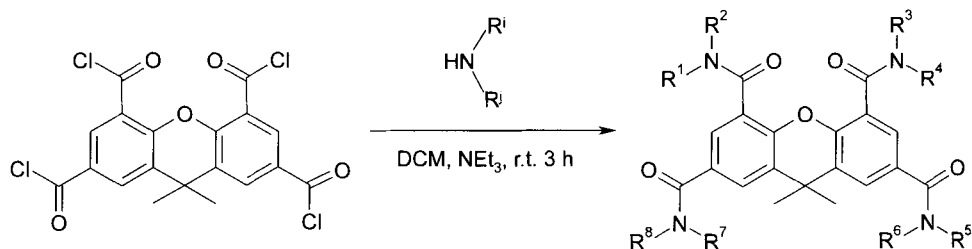
As with solid-phase combinatorial synthesis, libraries can be prepared either as mixtures or in the form of single compounds. Nowadays, the latter strategy seems to be preferred.

## 3.2 Synthesis of Mixtures

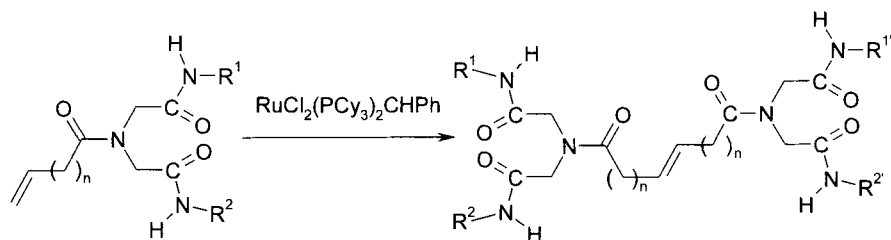
From the point of view of synthetic effort, preparation of combinatorial mixtures is by far the most economical approach. It can be done with ordinary laboratory equipment and does not take more time than the synthesis of any one of the individual components of the library. This simplicity, however, has its price: firstly, the more components a mixture contains the more difficult it becomes to follow the reaction analytically and to determine the actual composition of the reaction product. Secondly, if hits are found in a biological assay, deconvolution is required. In most cases this is done via resynthesis either of the individual components or of subsets of the mixture. If the composition of the initial mixtures was carefully planned it may be possible to identify the active component(s) by simply comparing the composition of the active mixtures with those of the inactive ones. Corresponding procedures have been reported in the literature (*e.g.*, the techniques of ‘indexed’ [1, 2] and ‘orthogonal’ [3] chemical libraries have been used in solution-phase synthesis). However, the biological effect of a mixture may also be due to a combined action of several weakly active members, with the result that deconvolution does not identify a significantly active compound. Finally, the problem of impurities multiplies with the complexity of the mixtures.

Synthesizing mixtures is a very simple way of producing unsymmetrically modified derivatives from scaffolds that contain several equally reactive groups of one type. This strategy obviates the use of protective groups which would be required if the compounds were to be synthesized individually. A typical example is the synthesis of a mixture of xanthene tetracarboxamides from the corresponding tetrachloride and 19 derivatives of  $\alpha$ -amino acids (Scheme 3.1) [4]. Further examples for such a strategy can be found in references [4–6, 92, 94]. The synthesis of further scaffolds with four carboxylic acids including partially protected species has been reported [97]. They comprise biphenyl 2,2',6,6'-, biphenyl 2,2',4,4'-, and 4,4'-bipyridyl-2,2',6,6'-tetracarboxylic acids.

Similarly, mixtures of dicarboxylic acid diamides that also contain asymmetric species were prepared by metathesis from olefinic mono-carboxamides (Scheme 3.2) [7, 95–119, 123].



**Scheme 3.1.** Preparation of a combinatorial library of xanthene tetracarboxamides as a mixture.



**Scheme 3.2.** Preparation of mixtures of unsymmetrical dicarboxylic acid diamides by olefine metathesis.

### 3.3 Reactions Applied to Solution-Phase Combinatorial Chemistry

As has been mentioned above, there are principally no constraints in terms of reaction conditions in solution-phase combinatorial chemistry. However, the necessity to obtain sufficiently pure products may limit the types of reactions that are suitable for this purpose. These limitations may decrease with the advent of automated purification techniques, particularly those based on chromatography coupled with analytical tools such as mass spectrometry.

## 3.3.1 Acylation of Alcohols and Amines

Acylation of alcohols and amines is in many cases a smooth unidirectional reaction that gives high yields and does not require large excesses of starting materials. These features made it one of the most frequently applied reactions in solution-phase combinatorial chemistry. Table 3.1 provides an overview of the structural fields to which amide and ester formation has been applied. The table also indicates additional functional groups (groups that are already mentioned in earlier columns of this and all following tables are not specified in this column), particularly those that can exert stronger intermolecular interactions. This may give a rough estimate of the diversity that was achieved by the corresponding libraries. Groups that exert weak or predominantly lipophilic interactions such as alkyl, aryl, ether, thioether, trifluoromethyl or halogenes are not mentioned explicitly.

**Table 3.1.** Acylations.

Acids	Nucleophile	Activation	Presence of other functional groups	Reference
alkyl, aryl, heteroaryl carboxylic acids	1-alkyl, 1-aryl piperazines primary, secondary aliphatic amines <sup>a)</sup>	acid chlorides	<i>tert</i> -amino	[8]
cubane tetracarboxylic acid, xanthene tetracarboxylic acid benzene triacetic acid	$\alpha$ -amino acid esters <sup>b)</sup>	acid chlorides	protected carboxy, hydroxy, amino, imidazolyl	[4]
not specified, including monoalkyl-carbonic acid	1-acyl piperazines, aliphatic, aromatic amines	acid chlorides acids/CDI	aromatic amino	[9]
aliphatic, aromatic carboxylic acids	1-nitrophenyl piperazines	acid chlorides	methoxycarbonyl, nitro	[10]
monoalkyl-carbonic acids	indolyl-5-amines	acid chloride	indolyl (unprot.), <i>tert</i> -amino	[18]
aliphatic, aromatic carboxylic acids	1-alkyl-piperazines	acid chlorides	aminothiazolyl	[19]
aliphatic, aromatic carboxylic acids	<i>N</i> -methyl-phenethylamines 4-acetylamino-4-phenyl-piperidines	acid chlorides		[19]
aliphatic, aromatic, monoalkyl-carbonic acids	aromatic, primary, secondary aliphatic amines	acid chlorides		[20]
aliphatic and aromatic carboxylic acids	<i>N</i> -methyl-benzylamine	acid chlorides	nitro	[80]



**Table 3.1.** Continued

Acids	Nucleophile	Activation	Presence of other functional groups	Reference
3-nitro-5-trifluoromethyl benzoic acid	ammonia, primary aliphatic amines, aliphatic alcohols	acid chloride	diethylphosphono	[55]
benzoic acids	primary aliphatic and aromatic amines, aliphatic alcohols	acid chlorides	cyano, furyl	[60]
1-methyl-3- <i>tert</i> -butyl-pyrazol-5-carboxylic acid	3-amino-benzophenones, variously subst. anilines	acid chloride	Si(Me) <sub>3</sub> , ketone, ester, acylamino, aminocarbonyl, ureido, nitro, 2,3-diketoidolyl, β-ketoester and amide, sulfonyl, isothiazolyl	[17]
2-chloro-4-trifluoromethyl-pyrimidine-5-carboxylic acid	3-amino-5-trifluoromethyl-benzoic amides and esters	acid chloride <sup>c)</sup>	diethylphosphono	[55]
unspecified acids, carbamic acid	unspecified	acid chlorides		[52]
alkyl, aryl, heteroaryl, amino carboxylic acids	piperidines, piperazines	acid chlorides	keto, <i>tert</i> -amino, furyl	[100]
aliphatic and aromatic acids	4- <i>secondary</i> amino-1-benzylpiperidines	acid chlorides	<i>N</i> -Me-pyrrolyl, furyl, thienyl	[91]
alkyl, aryl, alkoxy carboxylic acids	5-amino-anthranilic acids <sup>d)</sup>	acid chlorides	C-trimethylsilyl, sulfonylamino	[102]
aromatic and heteroaromatic acids	unspecified amines	acid chlorides		[93]
α-subst. acetic acids	anthranilic acid esters	acid chlorides	cyano, nitro, ester, pyridyl	[106]
aliphatic and aromatic acids	3β-aminomethyl-3α-androstane-17-ones	acid chlorides		[111]
aliphatic and aromatic acids	3-aminoimidazo[1,2-a]pyridines and -pyrazines	acid chlorides	<i>tert</i> -anilino, ester	[116]
benzoic acids	2-hydroxy-acetophenones	acid chlorides (in pyridine)	nitro	[126]
aliphatic, aromatic, heterocyclic acids	2-amino-thiophene-3,5-dicarboxylic acid derivatives	acid chlorides		[149]
aliphatic, aromatic carboxylic acids	1-diphenylpropyl-4-aminopiperidine	acid/ DIC/ HOBt	hydroxy, sulfonyl, acylamino, ureido	[11]

**Table 3.1.** Continued

Acids	Nucleophile	Activation	Presence of other functional groups	Reference
aliphatic, aromatic carboxylic acids	hexahydro-isoindole-5,6-dicarboxylic-monoamide	acid/ EDC	indolyl (unprotected)	[13, 15, 118]
pyrazole-3-carboxylic acids	primary amines	EDC/ HOAt		[156]
acetic acids	$\alpha$ -(alkylamino) alkanolic acids	acid/ HOBt, DCC	cyano, diethyl-phosphono, nitro	[26]
alkyl and aryl carboxylic acids	5-amino-benzimidazoles	EEDQ	pyridyl	[50]
$\omega$ -unsaturated aliphatic, aromatic, heterocyclic carboxylic acids	iminodiacetic diamides	acid/ PyBOP or PyBrOP	hydroxy, ester, acyl-amino and -guanidino	[7, 13, 14, 95]
pyrimidine-5-carboxylic acid	<i>N</i> -methyl-benzyl-amines	acid/ HOBt, EDC		[54]
$\alpha$ -ureido-alkanoic acids	not specified	acid/ EDC	thiazolyl	[16]
Boc-iminodiacetic acid monoamides	primary, secondary aliphatic, primary aromatic amines <sup>b)</sup>	acid/ PyBOP	hydroxy, ester, acylamino	[7, 13, 14, 21, 95, 123, 132]
unsaturated, aromatic, cycloaliphatic dicarboxylic acids	iminodiacetic diamides <sup>b)</sup>	diacid/PyBrOP	hydroxy, acylamino, ester	[21, 123]
tricarboxylic acids	iminodiacetic diamides <sup>b)</sup>	triacid/PyBrOP		[132]
<i>N</i> -Boc-hexahydro-isoindole-5,6-dicarboxylic acid monoamide	primary, secondary aliphatic amines	acid/ EDC	cyano	[13, 15]
diphenylmethane- and biphenyl-tetracarboxylic acid	$\alpha$ -amino acid esters <sup>b)</sup>	acid/TBTU		[94]
aromatic, heteroaromatic carboxylic acids	3-amino-benzophenones	anhydrides	NH- <i>t</i> -Bu, SiMe <sub>3</sub>	[17]
aliphatic dicarboxylic acids	aromatic amines	cyclic anhydrides	phthalimido, pyridyl	[12]
Boc-iminodiacetic acid	primary aliphatic and aromatic amines <sup>b)</sup>	cyclic anhydride	hydroxy, ester, acylamino	[7, 13, 14, 95, 123, 132]
<i>N</i> -Boc-hexahydro-isoindole-5,6-dicarboxylic acids	primary aliphatic amines <sup>b)</sup>	cyclic anhydride		[13, 15]

**Table 3.1.** Continued

Acids	Nucleophile	Activation	Presence of other functional groups	Reference
Boc-phenyl alanine	aliphatic amines	4-nitrophenyl-ester	hydroxy, ester	[11]
aryl-glyoxylic acids	primary, secondary alkylene-diamines	methyl ester	secondary amino	[24]
unspecified carboxylic acids	1,1,1-trialkyl-hydrazinium	acyl chloride, ester, oxazolone <sup>e)</sup>		[43]
$\alpha$ -acyl-amino acids	amino acids, peptides	oxazolones		[44]
protected amino acids	unspecified aliphatic amines			[54]

a) Alcohols were also included (formation of esters).

b) Mixtures were synthesized.

c) Sonication.

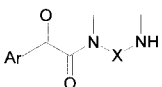
d) 1) Acylation of the 2-amino group (5-amino Boc protected); 2) acylation of the 5-amino group.

e) Product: acyl aminimides ( $R_3N^+-N^--COR'$ ).

**Table 3.2.** Synthesis of ureas and thioureas.

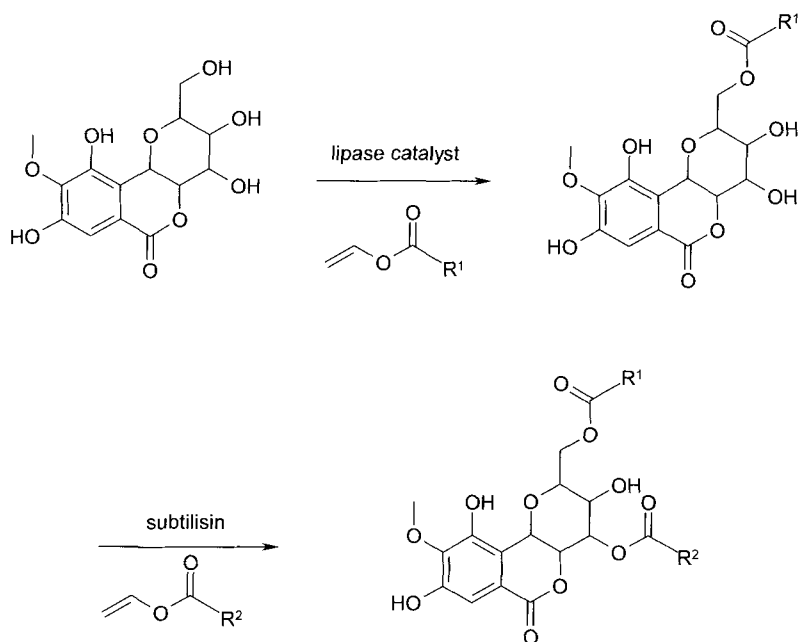
$R-N=C=X$	Amine	Presence of other functional groups	Reference
X = O R = aliphatic	aromatic, primary, secondary aliphatic amines		[20]
X = O, R = aryl	$\alpha$ -(benzyloxyamino)-acid esters		[22]
X = O R = aliphatic, aromatic	<i>N</i> -methyl-phenethylamines 4-acetylamino-4-phenyl-piperidines 4-aralkyl piperazines	amino-thiazolyl	[19]
X = O R = aliphatic, aromatic	<i>N</i> -methyl-benzylamine	nitro, ester	[80]
X = O R = aliphatic, aromatic	anthranilic acids		[83]
X = O R = aliphatic	5-amino-anthranilic acids <sup>a)</sup>	C-trimethylsilyl, sulfonylamino, amino-carbonyl	[102]
2,2,5,7-tetra-isocyanato-10,10-dimethyl-xanthene	methyl esters of $\alpha$ -amino acids	hydroxy, indolyl (un-protected), imidazolyl, acylamino	[5]

Table 3.2. Continued

R-N=C=X	Amine	Presence of other functional groups	Reference
X=O R not specified	unspecified primary and secondary amines		[23]
X=O R not specified	1-acyl-piperazines, aromatic amines		[9]
X=O derived from $\alpha$ -amino acids	2-amino-thiazoles		[16]
X=O R derived from $\alpha$ -amino acids	secondary aralkylamines	indolyl (unprotected), aromatic hydroxy, <i>tert</i> -amino	[31]
X=O R = aryl	3-aminoimidazo[1,2-a]-pyridines and -pyrazines	<i>tert</i> -anilino, ester	[116]
X=S	primary amines		[54]
X=S		benzothienyl	[142]
aliphatic and aromatic isothiocyanates	<i>N</i> -alkyl- $\beta$ -amino acid esters		[56]
X=O, S	primary and secondary amines	aromatic hydroxy, ester, indolyl (unprotected)	[52]
X=O, S R = aliphatic, aromatic	5-aminoindoles	indolyl (unprotected), <i>tert</i> -amino, ester, ketone	[18]
2 different $\alpha$ -amino acids and 1,1'-carbonylbisbenzotriazole		aminocarbonyl, protected imidazole, protected guanidino, ester	[137]

a) 1) acylation of the 2-amino group (5-amino Boc protected), 2) acylation of the 5-amino group

Producing combinatorial libraries via regioselective enzymatic acylation of alcohols has recently been reported [90]. The authors produced a library of bis-acyl derivatives of bergenin through successive application of a lipase catalyst and subtilisin (Scheme 3.3). The acylating agents were either vinyl or trifluoroethyl esters and were used in large excess. Unreacted esters were removed either during evaporation or subsequently by extraction with hexane.



**Scheme 3.3.** Enzymatic combinatorial bisacylation of bergenin.

### 3.3.2 Sulfonation of Amines

Combinatorial sequences that comprise amine acylation can frequently be expanded to sulfonamide formation using sulfonyl chlorides (unspecified sulfochlorides [9, 52]; aliphatic and aromatic sulfochlorides [19]; aromatic sulfochlorides [80, 102]). Sulfonation of aliphatic diamines with protected  $\beta$ -aminoalkane sulfonyl chlorides was carried out in the presence of polymer-bound 4-dialkylaminopyridine [130] (see also Table 3.11).

### 3.3.3 Formation of Ureas, Thioureas and Carbamates

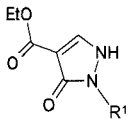
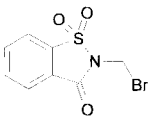
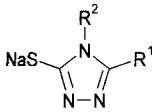
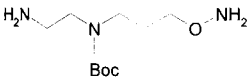
In some cases mentioned in Table 3.1, isocyanates and isothiocyanates have also been applied, leading to ureas and thioureas. These and additional examples can be found in Table 3.2. Alternatively, ureas have been prepared by the reaction of amines with *sec*-amino-carbonylchlorides (see [18], and corresponding entries in Table 3.1).

### 3.3.4 Alkylation and Addition Reactions

These types of reactions are highly interesting for combinatorial syntheses, as the starting materials – particularly amines, alcohols, phenols or alkyl halides – are commercially

available in great numbers or can be readily prepared via condensation reactions (e.g.,  $\alpha,\beta$ -unsaturated carbonyl compounds). Table 3.3 summarizes applications to solution-phase combinatorial synthesis.

**Table 3.3.** Alkylation and addition reactions.

Nucleophile	Electrophile	Presence of other functional groups	Reference
phenolic and heteroaromatic hydroxy	alkyl bromides	ester, pyridyl, quinolyl, pyrimidyl, pyrazolyl, isoxazolyl, triazolopyrimidyl	[28]
	alkyl bromides		[156]
alkyl and aryl carboxylates		protected amino, substituted ureido	[147]
	alkyl bromide		[54]
	benzyl bromides		[78]
polyazacyclophanes	benzyl bromides, bromoacetic acid derivatives <sup>21)</sup>	pyridyl, acylamino, cyano nitro, ester, aminocarbonyl	[6a,b, 92, 108, 140] <sup>19)</sup>
aliphatic polyamines	benzyl bromides, bromoacetic acid derivatives	pyridyl, acylamino, nitrile nitro, ester, aminocarbonyl	[6c,d]
<i>N</i> -mono substituted or protected piperazines	benzyl bromides, bromoacetamides, $\alpha$ -bromo ketones <sup>21)</sup>	pyridyl, benzothiazolyl	[86, 121]
4-aryl and heteroaryl piperazines, piperidines	benzyl bromides and chlorides, $\alpha$ -chloro ketones and acetanilides	hydroxy	[100]
4-substituted piperazines and piperidines, other secondary alicyclic amines	aralkyl mesylate	acylamino	[19]
	aralkyl iodide	aminothiazolyl	
secondary amines	alkyl bromides, iodides, and mesylates	aminocarbonyl	[52]

**Table 3.3.** Continued

Nucleophile	Electrophile	Presence of other functional groups	Reference
1,1-dialkyl-hydrazides	alkyl halide, epoxide <sup>b)</sup>		[43]
1-aryl-4-ethoxycarbonyl-5-hydroxy-pyrazole (O-alkylation)	benzyl bromides dialkylamino and pyridyl alkanols <sup>c)</sup>	sulfonyl  aromatic amino, acetamino	[25]
monoacyl-piperazines	epoxides		[9]
primary aliphatic amines	epoxides <sup>d)</sup>	hydroxy, aminocarbonyl, indolyl (unprotected)	[27]
primary aliphatic and aromatic, secondary aliphatic amines	epoxides	ester, <i>tert</i> -amino, hydroxy, sulfhydryl, pyridyl, quinolyl, tetrazolyl, thiadiazolyl, benzothiazolyl	[32]
primary and secondary aliphatic amines, ammonia	epoxides		[129]
primary aliphatic amines	spiro-3(R)-[oxirane-2',5 $\alpha$ -androstan]-17-one		[111]
monoacyl-alkylenediamines	2-aryl-4-ethoxymethylene-oxazole-5-ones	nitro, pyridyl, furyl	[24, 142]
secondary amines, thiols	2-vinyl oxazolones		[44]
2,6-diaminopurines <sup>e)</sup>	benzylchloride  2,3-difluorobenzyl alcohol <sup>f)</sup>		[88]

a) Mixtures were synthesized.

b) Product: acyl aminimides ( $R_3N^+ - N^- - COR'$ ).

c) Mitsunobu reaction.

d) Trimethylsilylation of the amine prevents double alkylation.

e) 9-alkylation.

f) In addition formation of guanidino derivatives with bis-Boc-S-methylisothiurea.

g) the ethyl ester was saponified after the alkylation

### 3.3.5 Reductive Amination

An alternative synthetic approach to amine alkylation, reductive amination, has also been used in solution-phase combinatorial chemistry (Table 3.4). Parallel syntheses implying alkylation of amines via acylation and subsequent reduction with diborane has also been reported [29, 86].

**Table 3.4.** Reductive amination.

Carbonyl component	Amino component	Presence of other functional groups	Reference
aromatic aldehydes	$\alpha$ -amino acid esters	N-methyl-pyrrolyl, furyl, pyridyl, quinolyl, indolyl (unprotected), imidazolyl, ester, nitro, dimethylamino	[30]
aromatic aldehydes	primary aliphatic	aromatic hydroxy, indolyl (unprotected)	[31]
aromatic aldehydes	$\beta$ -amino acid esters	furyl	[56]
aromatic and aliphatic aldehydes	methyl anthranilate		[106]
aldehydes	secondary amines		[52]
aldehydes, ketones	primary aliphatic amines	aromatic hydroxy, amino-carbonyl, aminosulfonyl, pyridyl, indolyl (unprotected)	[52]
aliphatic ketones	primary ethanolamines	aromatic hydroxy, amino-carbonyl, aminosulfonyl, cyano, indolyl (unprotected)	[89]
1-acyl-4-piperidone	unspecified amines		[9]

### 3.3.6 Arylation of Amines

Arylation of amines in combinatorial synthesis has been achieved by nucleophilic aromatic substitution or by metal-catalyzed amine/halogen exchange (Table 3.5).

### 3.3.7 C-C Bond Formation via Condensation Reactions

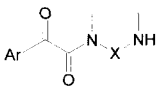
This paragraph refers to reactions between carbonyl and C-H acidic groups in library synthesis (Table 3.6).

### 3.3.8 Pd-Catalyzed C-C Bond Formation

Examples where such reactions have been applied to solution-phase combinatorial chemistry are listed in Table 3.7.

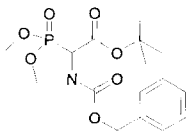
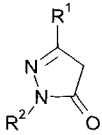
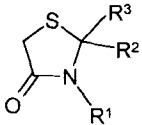


**Table 3.5.** Arylation of amines and alcohols.

Aromatic system	Nucleophile	Presence of other functional groups	Reference
fluoro-nitro-benzenes	Boc-piperazines	nitro, ester	[10]
2-chloro-6-trifluoromethyl-pyridine and pyrimidine-5-carboxylic acid derivatives	polyaza macrocycles <sup>a)</sup>	cyano, pyridyl	[92]
2,4-dichloro-triazines	anilines (first step)  primary, secondary aliphatic amines (second step)		[29]
2,3,6-trichlorotriazine	1) anilines  2, 3) primary secondary aliphatic and cyclic amines, dipeptide amides	aliphatic and aromatic hydroxy, carbohydrate moieties, <i>tert</i> -amino, nitro, keto, aminocarbonyl, aminimides	[101]
2,3,6-trichlorotriazine	1, 2) N-protected ethanolamine  3) $\alpha$ -amino acid methyl esters		[133]
triazines			[142]
chloro-pyrimidines, -triazines, -quinazolines, -quinoxalines	primary and secondary alkyl- amines, alkanols, phenols		[98]
2,6-dichloropurine		hydroxy	[88]
	primary and secondary aliphatic amines, hydrazine (2- and 6-Cl)		
2-amino-6-chloropurine			[88]
2-chloro-benzoic acids (cat. CuBr)	ammonia, primary aliphatic amines, anilines	nitro, pyridyl, pyrazolyl	[33]

a) Mixtures were synthesized.

**Table 3.6.** Condensation reactions.

Carbonyl component	C-H acidic component	Presence of other functional groups	Reference
nitro benzaldehydes	2-phosphorylmethyl-thiazoles	pyridyl	[12]
3-aryl propionic aldehydes	(triphenyl phosphoranylidene)acetic acid ester, cyanoacetic acid ester	pyridyl, quinolyl, pyrazinyl, furyl	[45]
3-formyl-thiophene and 3-formyl-(protected)pyrrole		furyl	[16]
aldehydes			[54]
aromatic and hetero-aromatic aldehydes	2-indolinones	ester, carboxy, aliphatic secondary and <i>tert</i> -amino, aromatic primary and <i>tert</i> -amino, acyl amino, aliphatic and aromatic hydroxy, nitro, sulfonyl, aminosulfonyl, pyrrolyl (substituted and unsubstituted NH), imidazolyl, pyridyl, indolyl (substituted and unsubstituted NH)	[96]
aldehydes	acetophenones	furyl, thienyl, pyridyl, pyrrolyl, <i>tert</i> -amino	[99, 107]
benzaldehyde			[105]

**Table 3.7.** Pd-catalyzed C-C bond formation.

First component	Second component	Product	Presence of other functional groups	Reference
aryl halides, aryl triflates <sup>a)</sup>	methyl acrylate, acrylonitrile, styrene, open chain and cyclic enol ethers, allyl derivatives, phenyl boronic acid, tributylstannyl-benzene	corresponding aryl-ethenes and biphenyls	trimethylsilyl, dimethylamino	[36]
phenyl boronic acid (4-substituted)	bromo-nitro-pyridines	corresponding phenylpyridine	dialkyl-piperazinyll,	[19]
bis(boryl)alkenes	alkyl, vinyl, aryl halides	mono(boryl)alkenes		[34]
2-alkoxy-5-stannyl-benzaldehydes	3-acetoxy-5-silyloxy-iodobenzene	corresponding biphenyls		[35]

a) Under microwave irradiation.

**Table 3.8.** Hydrogenations and reductions.

Process	Reducing agent	Presence of other functional groups	Reference
aromatic nitro to amino	Pd(PPh <sub>3</sub> ) <sub>4</sub> /H <sub>2</sub>	pyridyl, 4-benzyl-piperazino <sup>a)</sup>	[19]
aromatic nitro to amino	hydrazine/ Raney-Ni	pyridyl, benzimidazolyl	[50]
aromatic nitro to amino	Fe/ AcOH	ketone, ester, aromatic hydroxy, dialkylamino, sulfonyl	[17]
aromatic nitro to amino	H <sub>2</sub> / Pd-C	diethylphosphono	[55]
cleavage of benzyloxy-carbonyl amides	H <sub>2</sub> / Pd-C		[53]
cleavage of benzyloxy-carbonyl amides	H <sub>2</sub> / kat.		[54]
<i>N</i> -benzyloxycarbonyl-2,3-dehydro- $\alpha$ -amino acid- <i>t</i> -Bu-ester	H <sub>2</sub> / [Rh(cod)Me-DuPhos]BF <sub>4</sub> <sup>b)</sup>	thienyl, <i>N</i> -Ts-pyrrolyl	[16]
debenzylation of <i>O</i> -benzyl hydroxamic acids	H <sub>2</sub> / Pd-C	<i>N</i> -benzyl-piperazino, cyano	[86]

a) According to the reference, the benzyl group was unaffected.

b) Chiral catalyst (97.5–99.3% ee achieved); cod = cyclo-octadienyl, Me-DuPhos = 1,2-Bis(2,5-dimethylphospholano)benzene.

### 3.3.9 Hydrogenations and Reductions

Hydrogenations and reductions other than reductive aminations that occurred in combinatorial synthesis paths are shown in Table 3.8.

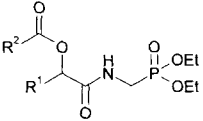
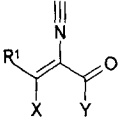
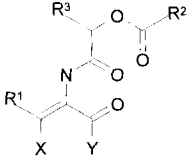
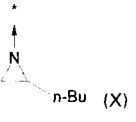
### 3.3.10 Multicomponent Reactions

Multicomponent reactions offer the possibility to introduce structural variations at more than two positions of a basic scaffold in a single step (for recent reviews, see [37]). In many cases at least some classes of components are commercially available in great numbers and reasonable structural diversity, *e.g.*, in the case of the Ugi reaction, primary amines, aldehydes, and carboxylic acids. For these reasons, multicomponent reactions provide, at least in principle, one of the most economical tools for synthesizing large libraries.

However, it must be borne in mind that multicomponent syntheses imply a number of reaction steps. Each of them may not only form the desired product, but also by-products to a larger or lesser extent. In addition, every one of the reagents may contribute its own impurities to the resulting reaction mixture. Both are expected to aggravate the purification problems. Therefore, much of the time that could principally be saved by using multicomponent reactions may have to be spent on finding appropriate reaction conditions and eventually on purification of the products (and perhaps even the starting materials).

Some examples for the use of multicomponent reactions in combinatorial chemistry are given in Table 3.9. Multicomponent reactions have also been employed in the synthesis of rings (see Section 3.3.11).

**Table 3.9.** Multicomponent reactions.<sup>a)</sup>

Components	Product	Presence of other functional groups	Reference
aliphatic, aromatic, $\alpha,\beta$ -unsaturated aldehydes aliphatic, aromatic, $\alpha,\beta$ -unsaturated carboxylic acids diethylphosphonomethyl-isonitrile		dimethylamino	[40]
commercially available 'diverse' aldehydes, carboxylic acids and isonitriles aliphatic, aromatic primary amines	amides of 2-acyl-amino-alkanoic acids	sulfonyl, hydroxy, amidino, guanidino, 2,4-diaminopyrimidyl	[41]
 $R^2\text{-COOH}, R^3\text{-CHO}$		 (S)-Me-glycidyl ( $R^3$ )	[125]

**Table 3.9.** Continued

Components	Product	Presence of other functional groups	Reference
Boc- $\alpha$ -amino acids, isonitriles, aldehydes, prim amines	Boc- $\alpha$ -aminoacyl- $\alpha$ -amino acid amides		[57]
aliphatic alcohols, aliphatic primary amines, isonitriles, aliphatic, aromatic and heterocyclic aldehydes, CO <sub>2</sub> (Ugi 5-component condensation) <sup>b)</sup>	<i>N</i> -alkoxycarbonyl- $\alpha$ -amino acid amides	pyridyl	[79]
mono <i>t</i> -butyl esters of succinic acids, aliphatic and aromatic aldehydes, ammonia or 2,4-dimethoxybenzyl-amine, aliphatic isonitriles	succinyl- $\alpha$ -amino acid amides	phthalimido	[117]
araliphatic methyl ketones araliphatic amines aliphatic carboxylic acids aliphatic isonitriles	$\alpha$ -methyl- $\alpha$ -acyl-aminoacyl-amides	protected amino, hydroxy	[120]
aliphatic, aromatic and hetero-aromatic aldehydes aryl and aralkyl amines silyl-enol ethers, alkenes	$\beta$ -amino acids and -ketones, quinolines <sup>c), d)</sup>	furyl, thienyl	[104]

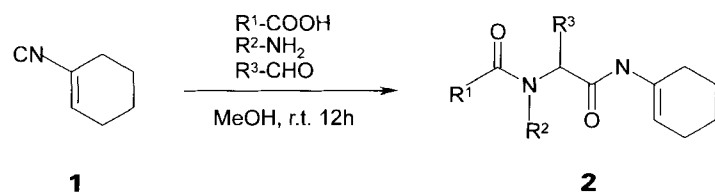
a) For an overview of various applications of Ugi multicomponent reactions to combinatorial chemistry, see [38].

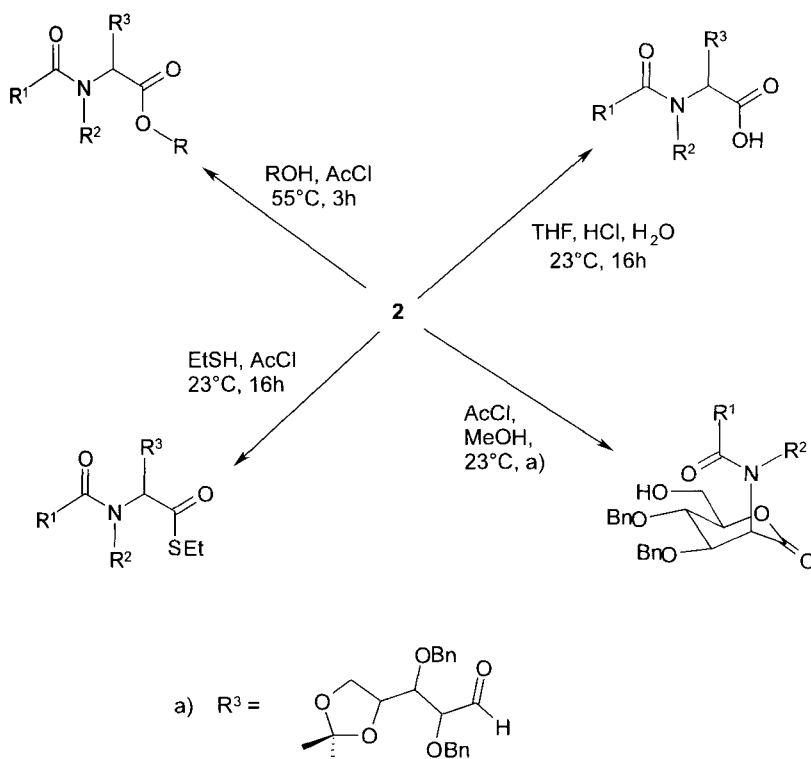
b) Also CS<sub>2</sub> and COS instead of CO<sub>2</sub> have been investigated and found to give partly different products [79].

c) Lanthanide(OTf)<sub>3</sub> as catalyst.

d) Purification by silica gel chromatography.

A very versatile application of the Ugi reaction has been reported by Keating and Armstrong [42]. They use cyclohexenyl isonitrile (**1**) as a 'universal isocyanide' in an Ugi four-component condensation. The other components are varied. The reaction products, 2-acylamino-alkanoic acid amides of generic structure **2**, can undergo a number of further reactions leading in parts to new types of structures and increased diversity (Scheme 3.4). If a resin with free alcoholic groups is used the acylaminoacyl moiety can be bound to the solid phase before it is subjected to further modifications (resin capture).





**Scheme 3.4.** Reactions of compounds 2.

### 3.3.11 Cyclization Reactions

Ring formation is a very important reaction in synthetic chemistry. This is particularly true for heterocyclic structures since they play a central role in medicinal chemistry, a field where combinatorial chemistry nowadays finds its most important application. Table 3.10 provides an overview over the achievements that have been made in combinatorial synthesis of carbo- and heterocyclic rings.

### 3.3.12 Miscellaneous Reactions

Other reactions that have been tried in solution-phase combinatorial chemistry comprise:

- Claisen rearrangement of allyloxy-pyrazoles [25].
- Formation of an aldehyde via lithiation of a bromo-substituted aromatic ring (3-bromo-thiophene or 3-bromo-1-tosyl-pyrrole) and reaction with DMF [16].

**Table 3.10.** Ring formation.

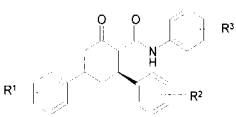
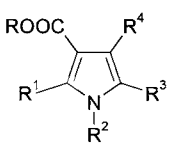
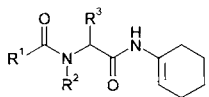
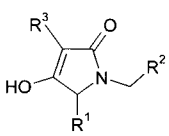
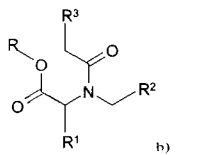
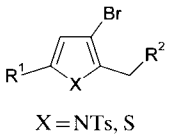
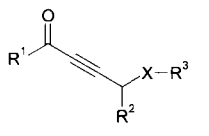
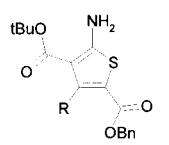
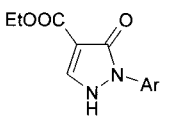
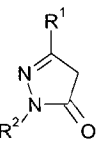
Ring system	Starting material(s)	Presence of other functional groups	Reference
	1,3-diaryl-1-propenones; acetoacetanilides	furyl, thienyl, pyridyl, pyrrolyl, <i>tert</i> -amino	[99]
$\beta$ -lactams	$\beta$ -amino acids, aliphatic and aromatic aldehydes, aliphatic isonitriles	ester, aminocarbonyl, hydroxy, benzotriazolyl	[58]
	 acetylene mono and dicarboxylic acid esters <sup>a)</sup>		[42]
	 b)	cyano, diethylphosphono, nitro	[26]
 X = NTs, S	 XR <sup>3</sup> = NHTs, STr	furyl	[16]
	$\beta$ -keto-carboxylic acid benzyl esters, <i>t</i> Bu-cyano- acetate, sulfur		[149]
	ethoxymethylene-diethyl- malonate, arylhydrazines	nitro	[25]
	$\beta$ -keto esters, hydrazines		[54]

Table 3.10. Continued.

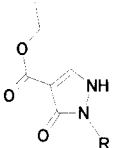
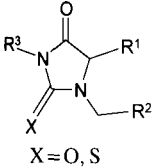
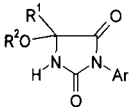
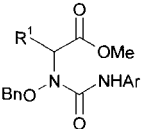
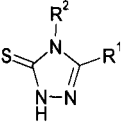
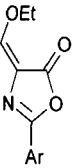
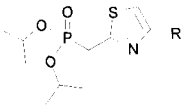
Ring system	Starting material(s)	Presence of other functional groups	Reference
	R-NHNH <sub>2</sub> Diethyl-ethoxymethylene malonate		[156]
1,3,5-triaryl-4,5-dihydro-pyrazoles	1,3-diaryl-1-propenones; arylhydrazines	furyl, thienyl, pyridyl, pyrrolyl, <i>tert</i> -amino	[99]
 X = O, S	$\alpha$ -(subst. amino) acid alkyl esters, iso(thio)cyanates	ester, nitro, dimethylamino, indolyl (unprotected), imidazolyl, pyridyl, quinolyl, thienyl, furyl, pyrrolyl	[30]
		hydroxy	[22]
	alkanols, allyl alcohol, glycol		
	acyl thiosemicarbazides		[54]
3,5-diaryl-isoxazolines	1,3-diaryl-1-propenones; hydroxylamine	furyl, thienyl, pyridyl, pyrrolyl, <i>tert</i> -amino	[99]
	$\alpha$ -aroylamino acids, acetic anhydride, triethyl orthoformate		[24]
	diisopropylphosphoryl-thioacetamide, alkyl and aryl halomethyl ketones	pyridyl	[12]



Table 3.10. Continued.

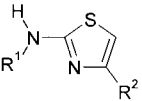
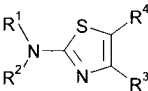
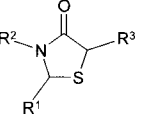
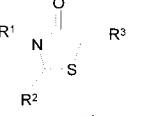
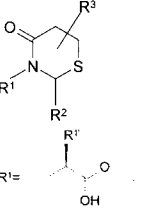
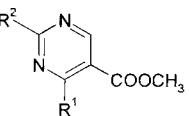
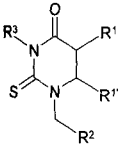
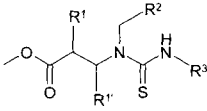
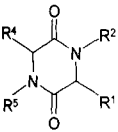
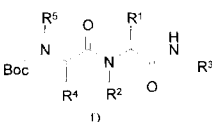
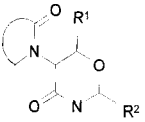
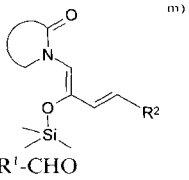
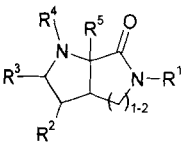
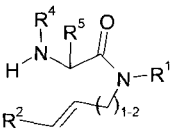
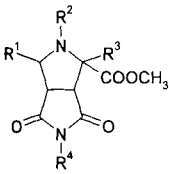
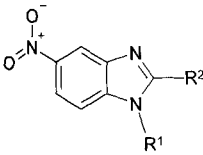
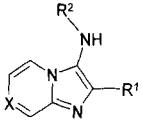
Ring system	Starting material(s)	Presence of other functional groups	Reference
	thioureas, bromomethyl ketones		[16]
amino- and imino-thiazoles	thioureas, $\alpha$ -halo ketones		[54]
Thiazoles	thiocarboxamides, $\alpha$ -halo ketones		[54]
	thioureas, $\alpha$ -bromo ketones	<i>tert</i> -amino, carboxy, pyridyl	[51]
	aliphatic and aromatic amines, $\alpha$ - and $\beta$ -amino acids; aromatic and heteroaromatic aldehydes 2-mercapto acids	hydroxy, ester, protected amino, furyl, thienyl	[109]
	$\alpha$ - and $\beta$ -amino acids aromatic aldehydes 2- or 3-mercaptoalkanoic acids	pyridyl	[39]
			
1,2,4-oxadiazoles	carboxylic acids, hydroxamic acid amides <sup>c)</sup>		[54]
pyridines	aromatic and heteroaromatic aldehydes, $\beta$ -keto esters or $\beta$ -diketones, ammonium nitrate (5:1 mixt. with bentonite) <sup>d)</sup>	hydroxy, acetamino, cyano, nitro, pyridyl, thienyl, ferrocenyl	[81]
2,4-diaryl-5-cyano-6-methyl-pyridines	1,3-diaryl-1-propenones; 3-amino-crotonitrile	furyl, thienyl, pyridyl, pyrrolyl, <i>tert</i> -amino	[99]
	$\beta$ -keto esters, amidines <sup>c)</sup>		[54]

Table 3.10. Continued.

Ring system	Starting material(s)	Presence of other functional groups	Reference
		ester, furyl	[56]
			[57, 135] <sup>ii)</sup>
		nitro, pyridyl, furyl, thienyl	[124]
		furyl, 2-keto-indolyl	[48]
	aldehydes, $\alpha$ -amino acid ester, maleinimides		[54]
	2-amino-4-nitroanilines, carboxylic acids	<i>tert</i> -amino, pyridyl, silyl ether	[50]
	2-amino-pyridine or -pyrazine	<i>tert</i> -anilino, ester	[116]
	aliphatic, aromatic isonitriles		
	aliphatic, aromatic aldehydes <sup>k)</sup>		

**Table 3.10.** Continued.

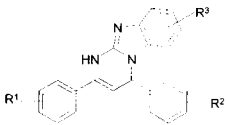
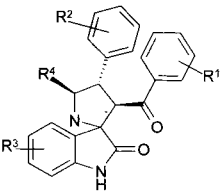
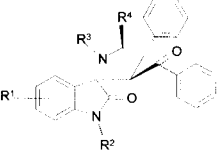
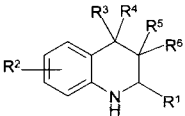
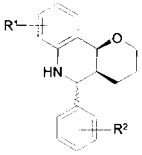
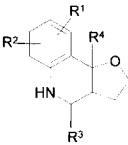
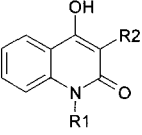
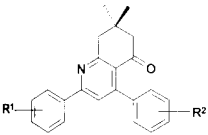
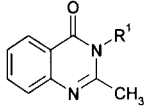
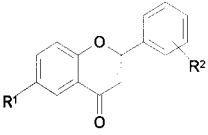
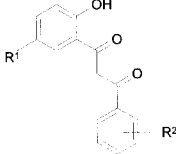
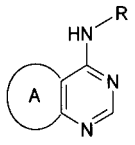
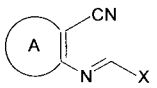
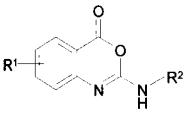
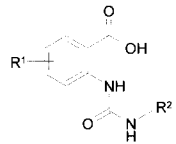
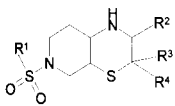
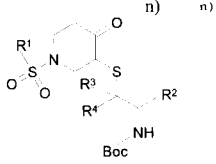
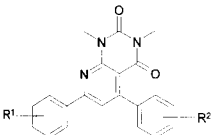
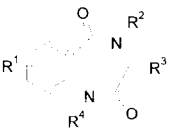
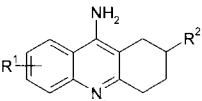
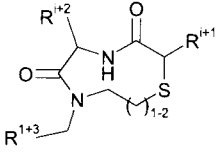
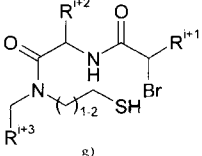
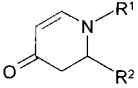
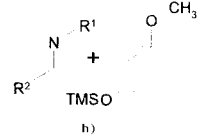
Ring system	Starting material(s)	Presence of other functional groups	Reference
	1,3-diaryl-1-propenones; 2-amino-benzimidazoles	furyl, thienyl, pyridyl, pyrrolyl, <i>tert</i> -amino	[99]
	1,3-diaryl-1-propenones; $\alpha$ -amino acids; isatines	furyl, thienyl, pyridyl, pyrrolyl, <i>tert</i> -amino, aminocarbonyl, hydroxy, sulfinyl, sulfonyl, indolyl (unprotected)	[99]
	1,3-diaryl-1-propenones; $\alpha$ -amino acids; isatines	nitro, hydroxy, sulfinyl, sulfonyl, aminocarbonyl, indolyl (unprotected), <i>tert</i> -amino, pyrrolyl ( <i>N</i> -methyl), thienyl	[107]
	aliphatic, aromatic, heterocyclic aldehydes; anilines; dihydrofuran, cycloolefins <sup>(i),j)</sup>	furyl, thienyl,	[104]
	aromatic, heteroaromatic aldehydes anilines dihydropyran	nitro, cyano, ester, furyl, thienyl, pyridyl	[112]
	aromatic aldehydes anilines dihydrofuranes		[115]
	( $\alpha$ -substituted acetyl) anthranilic acid esters	cyano, nitro, ester, pyridyl	[106]

Table 3.10. Continued.

Ring system	Starting material(s)	Presence of other functional groups	Reference
	1,3-diaryl-1-propenones; 3-amino-5,5-dimethyl-cyclohexenone	furyl, thienyl, pyridyl, pyrrolyl, <i>tert</i> -amino	[99]
	<i>N</i> -acetyl-anthranilic acids, anilines		[49, 103]
		nitro	[126]
	 X = OCH <sub>3</sub> , N(CH <sub>3</sub> ) <sub>2</sub> A = <i>e.g.</i> benzo, oxazolo, imidazo primary aliphatic amines		[114]
			[83]
6-acylamino-benzoxazinones	2,5-bis-acylamino-benzoic acids	sulfonylamino, aminocarbonyl	[102]
			[134]
	1,3-diaryl-1-propenones; 6-amino-1,3-dimethyluracil	furyl, thienyl, pyridyl, pyrrolyl, <i>tert</i> -amino	[99]

**Table 3.10.** Continued.

Ring system	Starting material(s)	Presence of other functional groups	Reference
	aliphatic and aromatic aldehydes, primary amines, cyclohexenylisonitrile, anthranilic acid (also N-alkyl and aryl) <sup>p)</sup>	ester, nitro, pyridyl	[42, 47, 151]
	2-cyanoanilines, cyclohexanones	hydroxy, ester, carboxy, dialkylamino	[1]
		protected indolyl, acylamino	[46]
		thienyl, furyl	[91]
5,10,15,20-tetraphenylporphyrin	pyrrole, 4-substituted benzaldehydes <sup>l)</sup>	acetoxy, ester, hydroxy, cyano	[122]

a) Propiolic acid ester also reacted, yields varied between 63% and 9%.

b)  $R^3 = \text{CN}$ ,  $2\text{-NO}_2\text{C}_6\text{H}_4$ ,  $\text{P}(\text{O})(\text{OEt})_2$ ,  $\text{SC}_6\text{H}_5$ .

c) O-acyl-hydroxamic acid amide is formed as an intermediate and subsequently cyclized.

d) Microwave-assisted.

e) Via dimethylaminomethylene  $\beta$ -keto esters.

f) The derivative with  $R^3 = \text{cyclohexenyl}$  (derived from cyclohexenyl isonitrile) was particularly suitable for ring closure.

g) Synthesized on solid support.

h) Imines from aldehydes and primary amines in the presence of trimethyl orthoformate, Diels–Alder reaction in the presence of  $\text{Yb}(\text{OTf})_3$ .

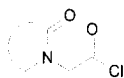
i) Lanthanide( $\text{OTf}$ )<sub>3</sub> as catalyst.

j) Purification by silica gel chromatography.

k)  $\text{Sc}(\text{OTf})_3$  catalysis.

l) Synthesis of mixtures.

m) From trimethylsilyl-aldimines and

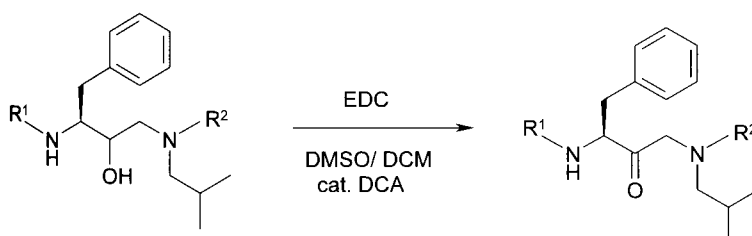


n) See Table 3.11.

o) One-pot conversion of the open-chain Ugi-type product with 1)  $\text{AcCl}/\text{MeOH}$ , 2) diethylamine or *N,N*-(diisopropyl)aminomethyl polystyrene.

p) Better yields are obtained, if *N*-Boc-anthranilic acids are used [151].

- Addition of Grignard or organo-lithium compounds to aldehydes [20].
- Moffat oxidation of amino alcohols with EDC (Scheme 3.5) [20].
- Epoxidation of olefins with dimethyldioxirane [129].
- Oxidation of thioethers to sulfones with dimethyldioxirane [134].
- Synthesis of diimines from 1,2-diketones and anilines as potential ligands for metal catalysts [59].
- Synthesis of 3,3-diphenyl-2-indolones as mixtures from isatins and benzenes in the presence of triflic acid [113].
- Synthesis of dibenzoylmethanes by rearrangement of 2-benzoyloxyacetophenones [126].



**Scheme 3.5.** Moffat oxidation of amino alcohols.

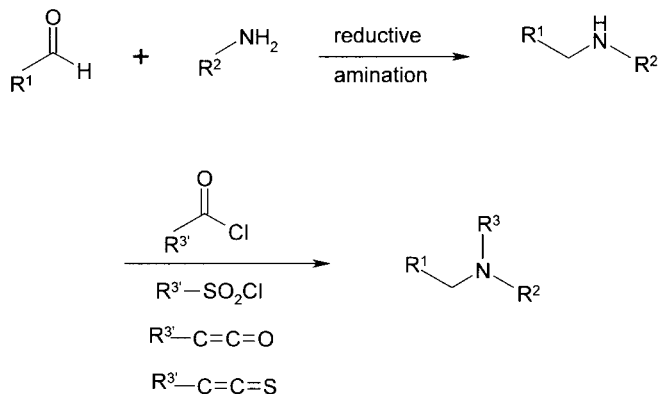
### 3.3.13 Reaction Sequences

One-step reactions have been used in combinatorial chemistry. However, if larger numbers of compounds are required, larger sets of starting materials of the same type have to be used. Even though these may often be commercially available in sufficient numbers, the resulting libraries are inadvertently less than optimal with respect to diversity. Therefore, the reactions listed above often form part of a reaction sequence that allows more than two positions in a target structure to be varied combinatorially.

Particularly in solution phase, it is necessary to minimize the number of synthesis steps in which the number of products is not multiplied. Such 'unproductive' steps potentially contribute to impurities, but increase diversity at best moderately. Most commonly, such steps refer to protection and deprotection or activation of functional groups. Principally, also cyclizations would belong into this category. Here, however, the spatial arrangement of the groups that potentially interact with targets may change dramatically. This in turn increases diversity and thereby the amount of obtainable structure–property information considerably, even though the number of products does not differ from that of the starting materials.

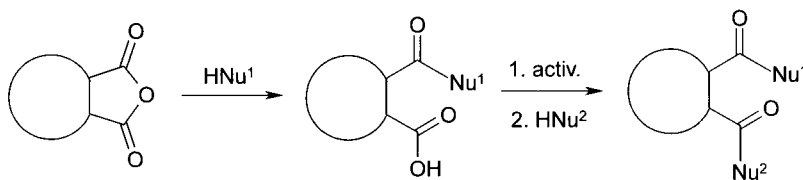
Considering efficiency, reaction sequences are most valuable in which structural variations can be introduced in such a way that the products can be used immediately in the next step. This condition is, of course, ideally fulfilled by multicomponent reactions (see above), but there are also other ways of approaching this aim.

One of the most common examples of such a sequence is reductive amination of aldehydes or ketones, followed by acylation of the newly formed secondary amine (Scheme 3.6, see [26, 30, 31]). Such a sequence was also applied in the synthesis of carbohydrate-based libraries [110]. The carbonyl compounds were formed by reacting protected 1-thiol-carbohydrates with an  $\alpha$ -halo or  $\alpha,\beta$ -unsaturated ketone.



**Scheme 3.6.** Combinatorial sequence: reductive amination/acylation.

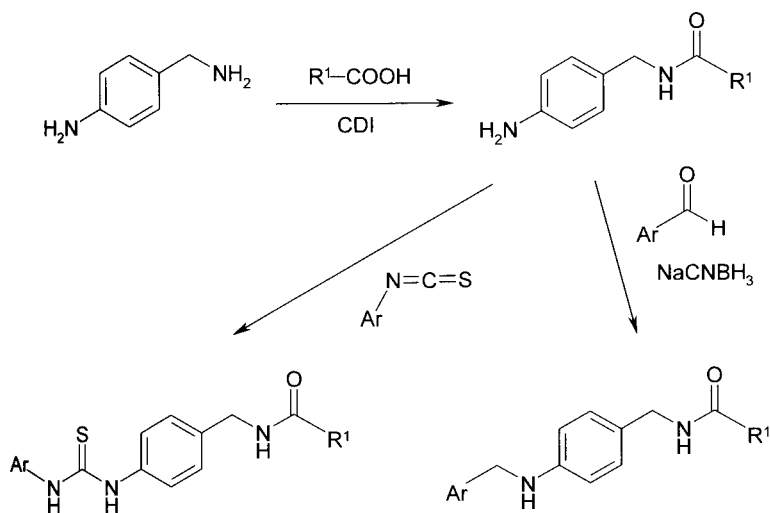
Another sequence of that type starts from cyclic anhydrides of dicarboxylic acids which on reaction with a set of nucleophiles, *e.g.*, amines, open under liberation of a free carboxylate. These products can then be coupled with a further set of nucleophiles (Scheme 3.7, see [13–15, 21]).



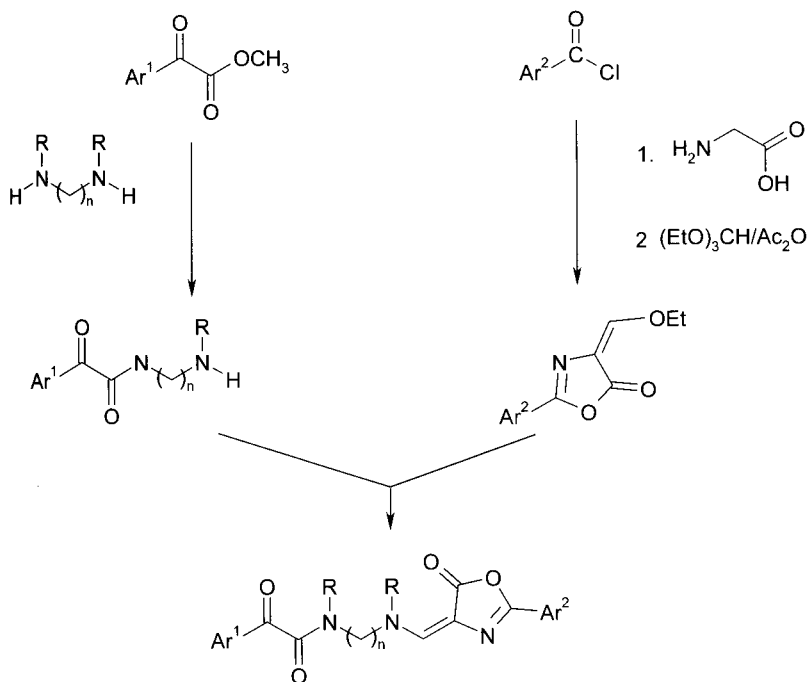
**Scheme 3.7.** Combinatorial sequence: double amination of cyclic anhydrides.

Sequences that allow combinatorial variations to be introduced in immediate succession may also be verified by starting from compounds with functional groups that can be transformed independently. As an example, 4-amino-benzylamine was first selectively acylated at the aliphatic amino group. Subsequently the aromatic amino group was converted to a thioureido or a secondary amino group (Scheme 3.8, [9]).

It has been pointed out previously that solution-phase combinatorial synthesis is open to convergent strategies. A corresponding example is shown in Scheme 3.9 [24].



**Scheme 3.8.** Combinatorial sequence: successive acylation of diamines.



**Scheme 3.9.** Example of a convergent combinatorial synthesis in solution.



### 3.4 Purification

It has already been mentioned that the scope of reactions that lend themselves to solution-phase combinatorial chemistry is dependent on the possibility of purifying the products. A number of techniques have been used for this purpose, though their significance often depends on the size of the library and on how easily they can be automated. This applies particularly to most chromatographic methods. One of these that can be easily run in parallel fashion is flash chromatography, and this has therefore also been applied in library synthesis (*e.g.* [47, 111]) even of mixtures (*e.g.* [6]). Also simply shaking with sorbents (*e.g.*, alumina/silica gel [115]) may be sufficient to purify the products. Extraction procedures have been followed widely in solution-phase combinatorial chemistry, and these and some other techniques will now be reviewed briefly.

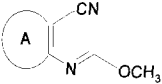
#### 3.4.1 Solid-Phase-Bound Reagents

Auxiliary reagents, or products derived from them (*e.g.*, tertiary amines or coupling reagents) are one of the most common sources of impurities. Adding such reagents bound to solid support provides a very efficient way of avoiding such impurities (Table 3.11).

**Table 3.11.** Solid-phase-bound reagents.

Reaction	Reagent	Presence of other functional groups	Reference
acylation of amines and alcohols with acid chlorides	basic ion exchange resin (Amberlite® IRA-68) <sup>a)</sup>	cyano, furyl	[60]
acylation of amines and alcohols with acid chlorides	basic ion exchange resin (Amberlyst 21) <sup>a, b)</sup>		[93]
acylation of aromatic amines	polymer-supported morpholine	<i>tert</i> -aromatic amine, ester	[116]
alkylation of aromatic hydroxyl	anion exchange resin (Amberlite® IRA-900) <sup>c)</sup>	ester, pyridyl, quinolyl, pyrazolyl, pyrimidyl, isoxazolyl, triazolopyrimidyl, quinoxalyl	[28]
alkylation of Boc-mercaptoamines with <i>N</i> -sulfonyl-3-bromo-4-piperidone	dimethylamino resin (Amberlyst A21)		[134]
intramolecular alkylation of a thiol	solid support-bound tetramethyl guanidine		[46]
intramolecular Dieckmann condensation	anion exchange resin (Amberlyst-A26 (OH <sup>-</sup> ))	cyano, diethylphosphono, nitro	[24]
intramolecular Claisen-type condensation	anion exchange resin (Amberlyst-A26 (OH <sup>-</sup> )) <sup>c)</sup>	cyano, nitro, ester, pyridyl	[106]

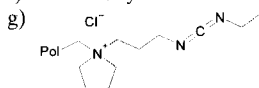
Table 3.11. Continued.

Reaction	Reagent	Presence of other functional groups	Reference
addition of primary amines to isocyanates	acidic ion exchange resin (Amberlite® IR-120 or Amberlyst 15) <sup>d)</sup>		[60]
formation of 	sulfonic acid resin		[114]
from the amine and orthoformate			
reductive amination of aldehydes	Amberlite® IRA-400 borohydride resin	indolyl (unprotected)	[31, 52]
reductive amination of 1-benzyl-2-phenyl piperidine-4-one	Amberlite® IRA-400 borohydride resin	<i>N</i> -Me-pyrrolyl, furyl, thienyl	[91]
reductive amination of <i>N</i> -sulfonated 3-aminoethylthio-4-piperidones forming piperidino-thiomorpholines	polymer-supported cyanoborohydride		[134]
benzoylation of sulfonamides	solid support-bound EDC	ester, nitro	[136]
formation of symmetric anhydrides of pyrazole carboxylic acids	solid support-bound EDC	ester	[17]
formation of 2-amino-benz-oxazinones from 2-ureido-benzoic acids	solid support-bound EDC		[83]
formation of benzoxazinones from <i>N</i> -acyl-anthranilic acids	solid support-bound EDC	sulfonylamino, aminocarbonyl	[102]
amide formation from carboxylic acids and amines	solid support-bound EDC <sup>g)</sup>		[153]
oxidation of primary alcohols to aldehydes	polymer (Amberlyst A-26)-supported perruthenate	pyridyl, <i>tert</i> -amino, aromatic primary amino, epoxy, nitro	[127–129, 131]
Wittig olefination with aliphatic and aromatic aldehydes and ketones	diphenylphosphine derivatized polystyrene		[129]
aryl ethers by Mitsunobu reaction from phenols and alcohols	diphenylphosphine derivatized polystyrene	ketone, cyano	[146]

**Table 3.11.** Continued.

Reaction	Reagent	Presence of other functional groups	Reference
Mukaiyama aldol reaction of aliphatic trimethylsilyl enol ethers with aromatic aldehydes	trimethylsilyl nafion	nitro	[131]
sulfonation of amines	polymer-bound <i>N</i> <sup>1</sup> -sulfonated DMAP <sup>0</sup>	nitro	[127]
3-bromination of <i>N</i> -sulfonyl-piperidine-4-one	polymer-bound pyridinium bromide perbromide		[134]

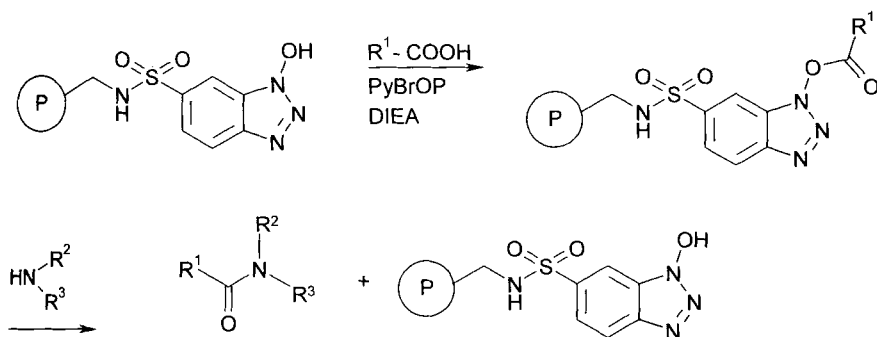
- a) A slight excess of acid chloride was used and quenched with a small amount of water after the reaction.  
 b) With sonication.  
 c) Amberlite® IRA-900 (Cl<sup>-</sup>) was first converted to the ArO<sup>-</sup>-form.  
 d) A slight excess of amine was used.  
 e) Impurities could be removed by filtration as the hydroxyquinolones remained tightly bound to the resin (liberation with TFA/MeOH).  
 f) Or sulfonyl chloride in the presence of polymer-bound DMAP [134].



It must be borne in mind, however, that like solid-phase chemistry generally, reactions involving solid support-bound auxiliaries – particularly catalysts – may lead to results differing from those obtained with their solution-phase counterparts. A detailed discussion of the factors that are relevant for the course of reactions on solid phase has recently been published [84].

Even more simply, one can resort to solid auxiliaries, *e.g.*, potassium or cesium carbonate in alkylation reactions [6a–c, 25, 88, 147], nucleophilic aromatic substitution [10], or Wittig–Horner reactions [12]. Such auxiliaries can subsequently be removed by filtration or aqueous work-up whereby the pure products may precipitate (*e.g.* [33]) or are otherwise obtained by extraction (see Section 3.4.3).

In a related strategy, one of the reagents is activated via the link to a solid support. In this way, it is possible to avoid the use of activating reagents in the synthesis of the final products. Such a procedure has been elucidated for the solution-phase acylation of amines. The acyl group was activated via a support-bound hydroxybenztriazolyl group (Scheme 3.10) [61]. A number of aromatic heteroaromatic and aliphatic carboxylic acids as well as aliphatic aromatic and heteroaromatic amines have been investigated. The yield of activated acids was mostly over 80 %, whereas the yields of amide ranged between 17 and 91 %. Solid support-bound *N*-hydroxy-succinimide has been used for the same purpose (primary and secondary aliphatic amines including  $\alpha$ -amino acids) [155]. Similarly, primary and secondary amines have been sulfonated via solid-phase-bound 1-sulfonyl-4-dialkylaminopyridines [127] (see also [130]).



**Scheme 3.10.** Amide formation via solid-support-bound activating agent.

Solid-phase-bound reagents have also been used for precipitating excess reagents so that they may be removed by simple filtration. As an example, a combination of a resin-bound calcium sulfonate and a sulfonic acid resin is able to remove excess TBAF from silyl deprotection reactions. This treatment precipitates the fluoride as  $\text{CaF}_2$  whereas the tetrabutylammonium remains bound to the sulfonic acid resin [139].

### 3.4.2 Solid-Phase Extraction

Solid-phase extraction is based on differences in the physico-chemical properties (predominantly basic or acidic properties) or chemical reactivity between starting materials, products, and auxiliaries of a chemical reaction. The reaction solution is brought into contact with a solid phase having chemical groups whose properties are complementary to the reagent that is to be extracted. After the excess starting materials or auxiliaries have been removed, the products are obtained by simply evaporating the solvent. If instead the products are extracted, an additional elution step is required.

Examples in which solid-phase extraction has been applied to solution-phase combinatorial synthesis are listed in Table 3.12 for extraction by ionic interactions, and in Table 3.13 for extraction by covalent bond formation.

Hydrophobic interactions may also serve to purify products of combinatorial syntheses. As an example, penta-O-lauroyl-1-thio- $\beta$ -galactose was subjected to a Michael addition to  $\alpha,\beta$ -unsaturated ketones or to alkylation by  $\alpha$ -chloro ketones followed by reduction of the keto groups or by reductive alkylation with amino acid esters. The reaction mixtures were passed through C18 silica gel, whereby the desired products were adsorbed while unwanted products were washed away. Desorption was carried out with pentane, and the hydrophobic labels were removed with methanolic sodium methoxide. The methyl laurates were finally separated from the desired products by chromatography on silica [148].

For further examples concerning the synthesis and the application of functionalized resins in solid-phase extraction, see also [63].

**Table 3.12.** Solid-phase extraction by ionic interaction.<sup>a)</sup>

Reaction	Extracted material	Solid phase	Reference
acylation of a primary amine having an additional <i>tert</i> -amino group	product (release with MeOH/NH <sub>3</sub> )	cation exchange resin (SCX)	[11]
reductive amination of excess aldehydes with primary and secondary amines	product (elution with MeOH/NH <sub>3</sub> )	cation exchange resin (SCX)	[52, 62]
reductive amination of excess ketones with primary ethanol amines	product (elution with MeOH/NH <sub>3</sub> )	cation exchange resin (SCX)	[89]
opening of epoxides with primary amines	product (release with MeOH/NH <sub>3</sub> )	cation exchange resin (SCX)	[27]
formation of thiazolidinones from aldehydes, amines and $\alpha$ -mercapto acids	unreacted amines and imines	cation exchange resin (SCX)	[109]
formation of ureas from primary amines and isocyanates	excess amine	cation exchange resin (Amberlyst 15)	[93]
nucleophilic aromatic substitution of 2-chloro-4-trifluoromethyl-pyrimidine carboxylic acid	excess amine	cation exchange resin (Amberlyst 15)	[93]
formation of ureas from amino-thiazoles and excess isocyanates	excess isocyanate quenched with 1,2-diaminoethane	cation exchange resin	[16]
acylation of amines using EDC as coupling reagent	excess EDC and the corresponding urea	cation exchange resin	[16]
benzoylation of sulfonamides	excess amine	cation exchange resin (Amberlyst 15)	[136]
formation of hexahydro-isoindole-5,6-dicarboxylic-monoamides from the corresponding dicarboxylic anhydride and of diamides from the monoamides using EDC	unreacted amine, EDC and the corresponding urea and <i>N</i> -acyl urea	cation exchange resin	[118]
formation of 3-aminoimidazo-[1,2-a]-pyridines and parazines from the amines, isonitriles and aldehydes	product (release with MeOH/ NH <sub>3</sub> )	strongly acidic exchange resin (Dowex 50WX 2-200)	[116]
addition of excess organometallics to aldehydes	(protonation of the formed metal alkoxides quenching of excess organo-metallics) binding of the metal ions <sup>b)</sup>	carboxylic acid functionalized resin (Amberlite® IRC-50S)	[20]

Table 3.12. Continued.

Reaction	Extracted material	Solid phase	Reference
thiazole formation from thioureas and excess $\alpha$ -bromoketones	excess $\alpha$ -bromoketone quenched with 4-carboxy-phenyl thiourea	anion exchange resin	[16]
acylation of excess amine with <i>p</i> -nitrophenylesters	1) nitrophenol 2) excess amine	1) anion exchange resin (SAX) 2) cation exchange resin (SCX)	[11]
urea formation from primary amines and isocyanates	excess isocyanate quenched with N,N-dimethylaminoethylamine or 1-(2-hydroxy-phenyl)piperazine	cation exchange resin or anion exchange resin	[62]
urea and thiourea formation between piperidino-thiomorpholines and isocyanates or isothiocyanates	unreacted isocyanate or isothiocyanate unreacted amine	aminomethyl polystyrene Amberlyst 15	[134]
Moffat oxidation of amino alcohols with EDC( $\cdot$ HCl)/DMSO	excess EDC and the formed corresponding urea HCl	cation exchange resin anion exchange resin <sup>e)</sup>	[20]
4-component condensation of succinic monoesters, aldehydes, ammonia and isonitriles	unreacted succinic monoester unreacted isonitrile <sup>e)</sup>	basic ion exchange resin acidic ion exchange resin	[117]
thioureas from secondary amines and excess isothiocyanates	excess isothiocyanates	aminomethylated polystyrene	[56]
acylation of amino-benzophenones with pyrazole carboxylic anhydrides	excess anhydride, formed carboxylic acid unreacted aminobenzophenone	<sup>e)</sup> 1) <sup>f)</sup> , 2) <sup>e)</sup>	[17]
deprotection of trimethylsilyl-ethyl protected carboxylic acids with TBAF	excess TBAF, tetrabutylammonium	resin-bound Ca <sup>++</sup> -sulfonate and sulfonic acid	[102]

a) See also footnotes in Table 3.11.

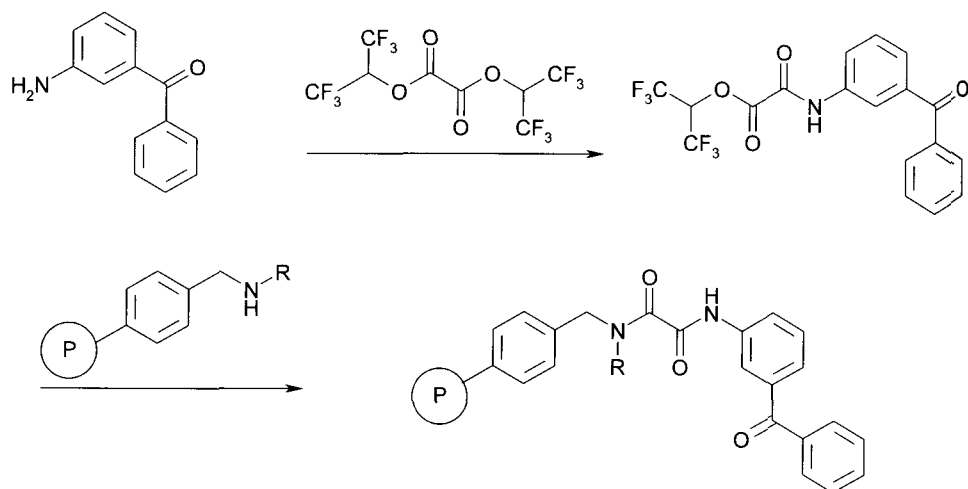
b) With Grignard reagents the addition was incomplete; the remaining aldehyde was removed in a second purification step with 1,2-diaminoethane functionalized resin (Schiff base formation).

c) Both applied simultaneously.

e) Polymer-supported *N*-benzylated bis-(2-aminoethyl)amine.

f) See Scheme 3.11.

g) After conversion to the corresponding amine by HCl/MeOH.

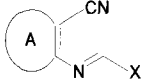


**Scheme 3.11.** Removal of unreacted aminobenzophenone.

**Table 3.13.** Solid-phase extraction by covalent bond formation.

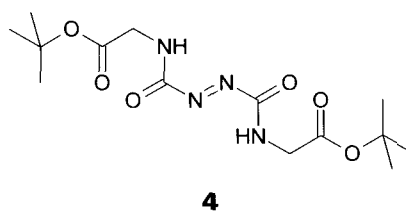
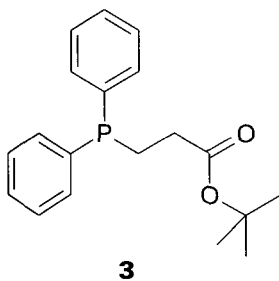
Reaction	Extracted material	Scavenger	Reference
acylation and sulfonation of amines with excess acid chlorides, isocyanates or sulfochlorides	excess acid chlorides, isocyanates or sulfochlorides	aminomethylated polystyrene	[18, 20, 23, 31, 52]
formation of thioureas from amines and isothiocyanates	excess isocyanates	aminomethylated polystyrene	[52]
acylation and sulfonation of anilines with acid chlorides, isocyanates or sulfochlorides <sup>a)</sup>	unreacted aniline and electrophiles	polyamine resin <sup>b)</sup>	[82, 102]
acylation of amines with excess acid chlorides	excess acid chlorides	polyamine resin	[91]
sulfonation of primary aliphatic diamines with protected $\beta$ -aminoalkane sulfonylchlorides	excess sulfonylchloride	polyamine resin	[130]
acylation of aromatic amines	excess acid chlorides or isocyanates	polymer-supported tris(2-aminoethyl)amine	[116]
urea formation from 4-nitrophenyl carbamates and amines	excess nitrophenyl carbamate and amine <sup>c)</sup>	combination of amine or polyamine resin with chlorocarbonyloxy or isocyanate resin <sup>c)</sup>	[150]
acylation of amines	excess amines	polymer-supported isocyanate	[91]
opening of epoxides with excess secondary amines	excess amines	polymer-supported isocyanate	[31]
alkylation of excess secondary amines with alkyl halides or sulfonates	excess amines	polymer-supported isocyanate	[31, 52]

**Table 3.13.** Continued.

Reaction	Extracted material	Scavenger	Reference
thiourea formation from isothiocyanates and amines;	excess amines	isatoic anhydride attached to Merrifield resin via N	[143]
anthranilic acid amides from isatoic anhydrides and amines			
reductive amination of aldehydes and ketones with excess primary amines	excess amines	polymer-supported aldehyde (Schiff base formation)	[31, 52, 91]
formation of fused pyrimidines from	excess amines	polymer-supported aldehyde	[114]
	and primary amines		
reductive amination of excess secondary amines with aldehydes	excess amines	polymer-supported benzoyl chloride	[31, 52]
condensation of thiazolidinones with benzaldehyde	unreacted aldehyde	polyamine resin (Schiff base formation)	[105]
preparation of thiazolidine-4-ones from primary amines, aldehydes (excess) and mercaptoacetic acid (excess)	excess aldehyde and (partly) mercaptoacetic acid	polyamine resin (formation of resin-bound thiazolidine-4-one)	[105]
	remaining mercaptoacetic acid	mercaptoethylamine resin (disulfide formation)	

- a) Excess aniline was used; to mimic incomplete reaction additional electrophile was added before work-up.
- b) At the end of the reaction tetrafluoro-phthalic anhydride was added which reacted with remaining aniline to form phthalic acid monoamide; the polyamine resin bound this latter product (via ionic interactions) and the remaining electrophile (by covalent bond formation).
- c) The amine resins also bound the by-product 4-nitrophenol via ionic interaction.

Conventional reagents that cannot easily be removed by solid-phase extraction may be tagged in such a way that extraction by scavenger resins becomes possible. For example, for Mitsunobu reactions phosphines and azodicarboxylic acid derivatives of types **3** and **4**





have been applied. After the reaction, the tags are unmasked and the excess reagents and their by-products sequestered by an anion exchange resin [141].

A modified version of solid-phase extraction in which the product is covalently bound to the resin has been introduced under the name 'resin capture' ([42]; see also Section 3.3.10). The resin can subsequently be used in solid-phase combinatorial synthesis (see e.g. [33–35]).

Instead of binding excess reagents to a solid phase they can be precipitated via polymer formation. Thus, excess amines used in acylation or sulfonation reactions can be precipitated as insoluble polymeric ureas by first adding excess 1,4-phenylene-diisocyanate and secondly penta-ethylene hexamine [138].

### 3.4.3 Liquid-Phase Extraction

Purification of products or intermediates by liquid-phase extraction has been variously reported. Examples are given in Table 3.14. The scope of liquid-phase extraction may be extended by adding highly polar scavengers before work-up. These react with excess reagents of low polarity to form products that are highly soluble in water and can easily be extracted by aqueous wash (see footnotes in Table 3.14).

**Table 3.14.** Liquid-phase extraction.

Reaction	Extracted material	Extraction conditions	Reference
acylation with acid chlorides or coupling reagents	excess acylating agent or acid	aqueous base	[9, 10, 111]
acylation with acid chlorides <sup>(a)</sup>	unreacted amino and acid components	1) aqueous citric acid 2) aqueous NaHCO <sub>3</sub>	[4a, 94]
reaction of <i>N</i> -methyl-benzyl-amine with excess acid chlorides, isocyanates and sulfochlorides	excess electrophiles	1) addition of potassium sarcosinate <sup>(b)</sup> 2) water	[80]
acylation of amines with cyclic acid anhydrides or with carboxylic acids and coupling reagents	unreacted amine, by-products derived from the coupling reagents <sup>(c)</sup>	aqueous HCl or 1) aqueous HCl, 2) aqueous NaHCO <sub>3</sub>	[7, 13–15, 21, 25, 45, 95]
tetra-ureido xanthenes from the tetra-isocyanate (prepared in situ from tetra-ethoxycarbonyl-aminoxanthene with catechol-chloroborane and NEt <sub>3</sub> ) and protected amino acids <sup>(d)</sup>	NEt <sub>3</sub> , borates, unreacted amino acids	1) aqueous citric acid 2) aqueous NaOH	[5]
acylation and alkylation of secondary amines with the corresponding halides in the presence of tetramethyl guanidine	tetramethyl guanidine	water	[100]

**Table 3.14.** Continued.

Reaction	Extracted material	Extraction conditions	Reference
alkylation of amines with alkyl bromides <sup>a)</sup>	inorganic salts	water	[6]
alkylation of amines with bromoacetamides $\alpha$ -bromo ketones and benzyl bromides <sup>a)</sup>	alkylating agent	aqueous solution of 3-mercaptopropane sulfonic acid	[78, 86]
thiazole formation from thioamides and $\alpha$ -bromo ketone <sup>a)</sup>	HBr	aqueous NaHCO <sub>3</sub>	[12]
nucleophilic aromatic substitution with amines	excess amines	aqueous acid	[10]
nucleophilic aromatic substitution with amines	excess amines	aqueous buffer supported by hydrophilic diatomaceous earth, elution with water-immiscible solvent	[98]
reductive amination	product	quenching of excess cyanoborohydride with HCl, adding NaOH and extracting with DCM	[9]
reductive amination of $\alpha$ -amino acid esters with aldehydes followed by reaction with isocyanates or isothiocyanates and ring closure to form (thio)hydantoin <sup>d)</sup>	borates triethylamine, unreacted aldehyde and iso(thio)cyanate	water water <sup>e)</sup>	[30]
reduction of nitro groups with SnCl <sub>2</sub> <sup>a)</sup>	inorganic materials	aqueous NaOH	[12]
formation of thiazolinones and dihydro thiazinones from amino acid esters, aldehydes and mercapto alkanolic acids	unreacted starting materials	1) aqueous NaHCO <sub>3</sub> 2) aqueous HCl <sup>f)</sup>	[39]
benzodiazepine formation with acetyl chloride (see Table 3.10)	acetic acid, excess acetyl chloride	aqueous NaHCO <sub>3</sub>	[47]
saponification of esters (LiOH in ethanol/dioxane)	inorganic salts	water	[16]

a) Mixtures were synthesized.

b) Sarcosine itself was unsuitable because of insolubility.

c) Cation exchange resin was also used successfully [21].

d) Also 2-thioxo-4-dihydropyrimidinediones from  $\beta$ -aminoacids [56].

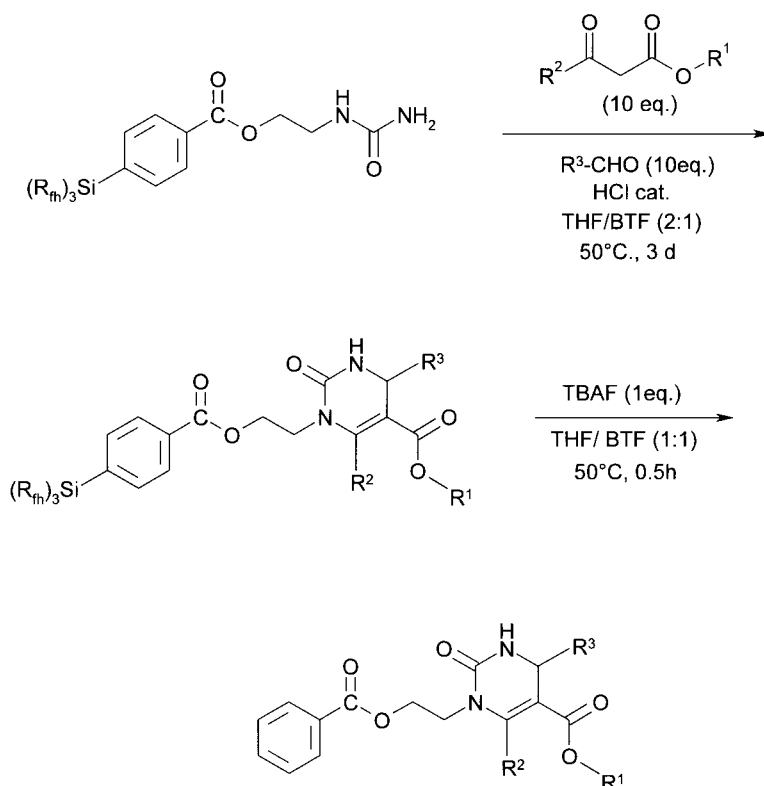
e) Glycine was added which reacted with unreacted aldehyde and iso(thio)cyanate, giving extractable products.

f) Followed by silica gel chromatography.

## 3.4.4 Fluorous Synthesis

Fluorous library synthesis is a technique that uses an interesting combination of liquid-phase extractions for purifying compounds. This technique takes advantage of the specific properties of highly fluorinated compounds. If partitioned between an aqueous or a standard organic phase and a phase consisting of a highly fluorinated liquid (fluorous phase), they appear predominantly in the latter one. Therefore, reaction components or products that carry highly fluorinated labels can be easily separated from 'normal' organic compounds by liquid-phase extractions with binary systems consisting of a fluorous phase and an organic or an aqueous phase. Ternary systems composed of all three types of phases have also been successfully applied [64]. A more detailed description of the synthetic strategies based on fluorous synthesis can be found in [64].

Fluorous synthesis has certain analogies to solid-phase synthesis: the fluorous label and the fluorous phase fulfil functions similar to those of the linker and the solid phase in that both provide a means of removing the attached compound from the reaction mixture (one because of its high affinity to the fluorous phase, the other via a covalent link to the solid support). The similarity, however, also extends to the necessity of additional synthetic steps involved in the attachment to and cleavage from the linker. Otherwise, the fluo-



**Scheme 3.12.** 'Fluorous' Biginelli reaction.

rous technique does combine the advantages of solid phase (notably the possibility of using reagents in large excess) and solution phase (higher reactivity, progress readily checked by conventional analytical methods) synthesis.

Fluorous synthesis protocols for combinatorial chemistry have been developed for various reactions:

- purification of carbinols derived from the addition of Grignard reagents to aldehydes via temporary formation of a tris-(perfluorohexylethyl)silyl-ether [64].
- Stille coupling with aryl tris-(perfluorohexylethyl)stannane as a reaction component (aryls also comprise furyl and pyridyl) [65, 66].
- Ugi reactions (with 4-[tris-(perfluorohexylethyl or perfluorodecylethyl)silyl]benzoic acid as acid component [64, 67].
- Biginelli reaction (Scheme 3.12) [64, 67]. Excess  $\beta$ -ketoester and aldehyde was separated from the fluorous reaction product by partitioning between toluene and FC-84 (mixture of isomers of  $C_7F_{16}$ ). The same system was used after cleavage of the silyl group with TBAF. In this case the toluene extracted the desired product while the fluorous silyl fluoride remained in the fluorous phase.
- 1,3-dipolar cyclo-additions of nitrile oxides to allyl or propargyl alcohols which were labeled as tris-(perfluorohexylethyl)silyl ethers [64, 68].
- 1,3-dipolar cyclo-additions of tris-(perfluorohexylethyl)stannyl azide to nitriles [64].

### 3.4.5 Synthesis on Soluble Polymers

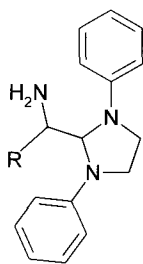
Synthesis on soluble polymers (for recent reviews, see [69, 70]) offers an other possibility to combine the advantages of solid-phase and solution-phase chemistry. Again, however, the convenience has to be paid for by two additional synthesis steps, attachment to and cleavage from the polymer (except if used instead of an otherwise required protective group) as well as (often) the loss of a functional group for the introduction of additional diversity. Various soluble polymer supports have been used in synthesis (a compilation can be found in [69]). Presently, however, PEG seems to be the most versatile and successful polymer support for solution-phase synthesis [69]. PEG is soluble in many organic solvents and in water, but insoluble in hexane, diethyl ether and *tert*-butyl methyl ether. Reaction products that are bound to PEG can therefore be separated from excess starting materials or auxiliary reagents by simply precipitating the polymer through addition of these solvents to the reaction mixture. The method, however, also has its limitations by being restricted to solvents in which the support is soluble. Therefore, the solvents mentioned above as well as cold tetrahydrofuran (THF) are unsuitable. Reagents or by-products must be soluble in the precipitating solvent. For example, an Fmoc protection strategy caused difficulties since the product formed on deprotection (fluorenylmethyl-piperidine) precipitates with the polymer [71].

Reported applications of the technique to combinatorial syntheses include libraries of peptides ([69, 71, 72]), 4-amino-benzenesulfonamides ([69, 71, 72]), and peptidomimetics (poly- $\alpha$ -aza-amino acids or azatides, [69, 71, 73]). The latter synthesis also illustrates one of the advantages of the soluble polymer approach over solid-phase synthesis, namely,

that it is also suitable if solid reagents are present (in this case a Pd catalyst in a debenzoylation [69, 73]). Similarly, MeOPEG-bound benzyloxycarbonyl aniline was deprotected with  $H_2/Pd/C$ . Subsequently, a Schiff base was formed with an aldehyde followed by reaction with a ketene or a titanium enolate of a S-pyridyl thioester [85].

A recently reported synthesis sequence [152] starts from MeOPEG-bound 4-fluoro-3-nitro-benzoic acid. After the nucleophilic aromatic substitution of the fluorine with piperazine, homopiperazine and 4-aminomethyl-piperidine (only the secondary amino group reacted) the remaining free amino function was acylated or sulfonated.

In another application, PEG-bound  $\alpha$ -aminoaldehydes served as a starting point for the combinatorial synthesis of polyamines [154]. Reductive amination with an amine of type **5**, followed by alkylation of the formed secondary amine (via reductive amination) and cleavage of the cyclic aminal yielded the starting aldehyde for the next cycle.

**5**

Various PEGs have been examined for their performance in the synthesis of a biaryl library via Suzuki coupling. Considering both the ease of precipitation and the loading PEG 6000 was found to offer optimum conditions [87].

Soluble polymers may also serve as carriers for catalysts and other auxiliary reagents. In this form they can be easily separated from the reaction products by simple precipitation of the polymer. Thus, a Sharpless asymmetric dihydroxylation with PEG-bound  $(DHQD)_2$ PHAL (bis-ether of 1,4-dihydroxyphthalazine and dihydroquinidine) of various olefins gave the corresponding diols in good yields with high enantiomeric excesses [74].

Soluble polymer-supported phosphines may be a valuable tool for Pd-catalyzed library synthesis in solution [145].

Like conventional solid supports, soluble polymers may have a potential to function as scavengers in solution-phase synthesis. The approach has been applied to the synthesis of  $\beta$ -amino alcohols that are structurally related to propranolol. The crude products (prepared from the corresponding epoxides and amines) and an added PEG-bound borohydride reagent formed a complex which was precipitated and separated from unreacted starting materials and unwanted by-products. Cleavage with HCl in MeOH/DCM and precipitation of the polymer gave a solution of pure (>92%)  $\beta$ -amino alcohols [75].

A related technique is based on soluble dendrimers as support (dendrimer-supported combinatorial chemistry, [76]). The feasibility of the approach was demonstrated by the synthesis of a small library of indoles. The synthesis started from an  $\alpha$ -amino acid that

formed an ester with a benzylic hydroxy group of the dendrimer. After acylation with an  $\omega$ -benzoyl alkanic acid the indole was formed by reaction with aryl hydrazines. The dendrimer-bound products were separated from all unbound materials by size-exclusion chromatography. Cleavage of the products from the dendrimer with methanol/triethylamine yielded the corresponding methyl esters. In an other example, dendrimer-linked 4-chloro-3-nitro-benzoic acid (as an ester) was subjected to a nucleophilic aromatic substitution with phenolates [77].

A related strategy was suggested in which one of the reactants is attached via a linker to tetrabenz[a,c,g,i]fluorene (Tbf). The reaction product can now be purified by taking advantage of the high affinity of Tbf to charcoal in polar solvents (mixture of DCM and methanol). Desorption can be affected with non polar solvents such as toluene [144]. The method was exemplified by the synthesis of a quinolone carboxylic acid derivative (ciprofloxacin).

### 3.5 Conclusions

The examples outlined in this chapter demonstrate that despite the problem of obtaining reasonably pure products, solution-phase synthesis has become a valuable tool in combinatorial chemistry. More advanced techniques that either reduce the need for purification, or allow it to be done in an automated manner, have been developed and will certainly be further improved. Progress in this field is expected to increase the role of solution-phase synthesis in the future. Particularly promising, however, are combinations of solution- and solid-phase techniques such as resin capture or support-bound reagents and auxiliaries. These combinations also imply solid-phase preparation of precursors for solution-phase combinatorial syntheses. Such sequences will be of all the more practical importance the lower the quantities of final products become that are needed for further investigations, e.g., for biological testing in drug research.

### 3.6 References

- [1] M. C. Pirrung, J. H. -L. Chau, J. Chen, *Chem. Biol.* **1995**, *2*, 621–626.
- [2] M. C. Pirrung, J. Chen, *J. Am. Chem. Soc.* **1995**, *117*, 1240–1245.
- [3] B. Déprez, X. Williard, L. Bourel, H. Coste, F. Hyafil, A. Tartar, *J. Am. Chem. Soc.* **1995**, *117*, 5405–5406.
- [4] a) T. Carell, E. A. Wintner, A. Bashir-Hashemi, J. Rebek, *Angew. Chem.* **1994**, *106*, 2159–2162, *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2059–2061. b) T. Carell, E. A. Wintner, A. J. Sutherland, J. Rebek, Y. M. Dunayevskiy, P. Vouros, *Chem. Biol.* **1995**, *2*, 171–183.
- [5] G. W. Shippy, U. P. Spitz, J. Rebek, *Bioorg. Med. Chem.* **1996**, *4*, 655–657.
- [6] a) H. An, P. D. Cook, *Tetrahedron Lett.* **1996**, *37*, 7233–7236. b) H. An, L. L. Cummins, R. H. Grifey, R. Bharadwaj, B. D. Haly, A. S. Fraser, L. Wilson-Lingard, L. M. Risen, J. R. Wyatt, P. D. Cook, *J. Am. Chem. Soc.* **1997**, *119*, 3696–3708. c) H. An, B. D. Haly, A. S. Fraser, C. J. Guinasso, P. D. Cook, *J. Org. Chem.* **1997**, *62*, 5156–5164. d) H. An, B. D. Haly, P. D. Cook, *J. Med. Chem.* **1998**, *41*, 706–716.
- [7] D. L. Boger, W. Chai, R. S. Ozer, C. -M. Andersson, *Bioorg. Med. Chem. Lett.* **1997**, *7*, 463–468.
- [8] P. W. Smith, J. Y. Q. Lai, A. R. Whittington, B. Cox, J. G. Houston, C. H. Stylli, M. N. Banks, P. R. Tiller, *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2821–2824.

- [9] C. D. Garr, J. R. Peterson, L. Schultz, A. R. Oliver, T. L. Underiner, R. D. Cramer, A. M. Ferguson, M. S. Lawless, D. E. Patterson, *J. Biomolec. Screening* **1996**, *1*, 179–186.
- [10] L. Neuville, J. Zhu, *Tetrahedron Lett.* **1997**, *38*, 4091–4094.
- [11] R. M. Lawrence, S. A. Biller, O. M. Fryszman, M. A. Poss, *Synthesis* **1997**, 553–558.
- [12] H. Maehr, R. Yang, *Bioorg. Med. Chem.* **1997**, *5*, 493–496.
- [13] S. Cheng, C. M. Tarby, D. D. Comer, J. P. Williams, L. H. Caporale, P. L. Myers, D. L. Boger, *Bioorg. Med. Chem.* **1996**, *4*, 727–737.
- [14] S. Cheng, D. D. Comer, J. P. Williams, P. L. Myers, D. L. Boger, *J. Am. Chem. Soc.* **1996**, *118*, 2567–2573.
- [15] D. L. Boger, C. M. Tarby, P. L. Myers, L. H. Caporale, *J. Am. Chem. Soc.* **1996**, *118*, 2109–2110.
- [16] A. Chucholowski, T. Masquelin, D. Obrecht, J. Stadlwieser, J. M. Villagordo, *Chimia* **1996**, *50*, 525–530.
- [17] J. J. Parlow, D. A. Mischke, S. S. Woodard, *J. Org. Chem.* **1997**, *62*, 5908–5919.
- [18] J. E. Fritz, S. W. Kaldor (Eli Lilly and Company) WO 97/00244, **1997**.
- [19] C. N. Selway, N. K. Terrett, *Bioorg. Med. Chem.* **1996**, *4*, 645–654.
- [20] D. L. Flynn, J. Z. Crich, R. V. Devraj, S. L. Hockerman, J. J. Parlow, M. S. South, S. Woodard, *J. Am. Chem. Soc.* **1997**, *119*, 4874–4881.
- [21] D. L. Boger, R. S. Ozer, C. -M. Andersson, *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1903–1908.
- [22] S. Hanessian, R. -Y. Yang, *Tetrahedron Lett.* **1996**, *37*, 5835–5838.
- [23] S. W. Kaldor, J. E. Fritz, J. Tang, E. R. McKinney, *Bioorg. Med. Chem. Lett.* **1996**, *6*, 3041–3044.
- [24] C. M. Baldino, D. S. Casebier, J. Caserta, G. Slobodkin, C. Tu, D. L. Coffen, *Synlett* **1997**, 488–490.
- [25] B. MacLachlan, Presentation at the IBC Conference on Synthetic Chemical Libraries in Drug Discovery, London **1995**.
- [26] B. A. Kulkarni, A. Ganesan, *Angew. Chem.* **1997**, *109*, 2565–2567, *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2454–2455.
- [27] A. J. Shuker, M. G. Siegel, D. P. Matthews, L. O. Weigel, *Tetrahedron Lett.* **1997**, *38*, 6149–6152.
- [28] J. J. Parlow, *Tetrahedron Lett.* **1996**, *37*, 5257–5260.
- [29] J. P. Whitten, Y. F. Xie, P. E. Erickson, T. R. Webb, E. B. De Souza, D. E. Grigoriadis, J. R. McCarthy, *J. Med. Chem.* **1996**, *39*, 4354–4357.
- [30] M. M. Sim, A. Ganesan, *J. Org. Chem.* **1997**, *62*, 3230–3235.
- [31] S. W. Kaldor, M. G. Siegel, J. E. Fritz, B. A. Dressman, P. J. Hahn, *Tetrahedron Lett.* **1996**, *37*, 7193–7196.
- [32] B. L. Chng, A. Ganesan, *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1511–1514.
- [33] A. L. Smith, C. G. Thomson, P. D. Leeson, *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1483–1486.
- [34] S. D. Brown, R. W. Armstrong, *J. Am. Chem. Soc.* **1996**, *118*, 6331–6332.
- [35] M. R. Pavia, M. P. Cohen, G. J. Dilley, D. L. Dubuc, T. L. Durgin, F. W. Forman, M. E. Hediger, G. Milot, T. S. Powers, I. Sucholeiki, S. Zhou, D. G. Hangauer, *Bioorg. Med. Chem.* **1996**, *4*, 659–666.
- [36] G. Lindeberg, M. Larhed, A. Hallberg (Labwell AB), WO 97/43230, **1997**.
- [37] a) I. Ugi, A. Dömling, W. Hörl, *Endeavour* **1994**, *18*, 115–122; b) I. Ugi, *J. prakt. Chem.* **1997**, *339*, 499–516.
- [38] I. Ugi, M. Goebel, B. Gruber, M. Heilingbrunner, C. Heiß, W. Hörl, O. Kern, M. Starnecker, *Res. Chem. Intermed.* **1996**, *22*, 625–644.
- [39] C. P. Holmes, J. P. Chinn, G. C. Look, E. M. Gordon, M. A. Gallop, *J. Org. Chem.* **1995**, *60*, 7328–7333.
- [40] R. W. Armstrong (The Regents of the University of California), WO 95/02566, **1995**.
- [41] L. Weber, S. Wallbaum, C. Broger, K. Gubernator, *Angew. Chem.* **1995**, *107*, 2452–2454, *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2280–2282.
- [42] T. A. Keating, R. W. Armstrong, *J. Am. Chem. Soc.* **1996**, *118*, 2574–2583.
- [43] J. C. Hogan (Legomer Partners L. P., now Arqule), WO 94/01102, **1994**.
- [44] J. C. Hogan (Legomer Partners L. P., now Arqule), WO 94/00509, **1994**.
- [45] D. L. Boger (The Scripps Research Institute), WO 96/03424, **1996**.
- [46] A. A. Virgilio, S. C. Schürer, J. A. Ellman, *Tetrahedron Lett.* **1996**, *37*, 6961–6964.
- [47] T. A. Keating, R. W. Armstrong, *J. Org. Chem.* **1996**, *61*, 8935–8939.
- [48] M. A. Marx, A. -L. Grillot, C. T. Louer, K. A. Beaver, P. A. Bartlett, *J. Am. Chem. Soc.* **1997**, *119*, 6153–6167.
- [49] R. A. Houghten, J. M. Ostresh (Torry Pines Institute for Molecular Studies), WO 97/10221, **1997**.

- [50] J. B. Thomas, M. J. Fall, J. B. Cooper, J. P. Burgess, F. I. Carroll, *Tetrahedron Lett.* **1997**, *38*, 5099–5102.
- [51] N. Bailey, A. W. Dean, D. B. Judd, D. Middlemiss, R. Storer, S. P. Watson, *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1409–1414.
- [52] M. Siegel, Presentation at the IBC Conference on Molecular Diversity and Combinatorial Chemistry, San Diego **1996**.
- [53] R. Storer, Presentation at the IBC Conference on Synthetic Chemical Libraries in Drug Discovery, London **1995**.
- [54] N. Bailey, A. W. J. Cooper, M. J. Deal, A. W. Dean, A. L. Gore, M. C. Hawes, D. B. Judd, A. T. Merritt, R. Storer, S. Travers, S. P. Watson, *Chimia* **1997**, *51*, 832–837.
- [55] R. W. Sullivan, C. G. Bigam, P. E. Erdman, M. S. S. Palanki, D. W. Anderson, M. E. Goldman, L. J. Ransone, M. J. Suto, *J. Med. Chem.* **1998**, *41*, 413–419.
- [56] M. M. Sim, C. L. Lee, A. Ganesan, *J. Org. Chem.* **1997**, *62*, 9358–9360.
- [57] C. Hulme, M. M. Morrisette, F. A. Volz, C. J. Burns, *Tetrahedron Lett.* **1998**, *39*, 1113–1116.
- [58] J. Pitlik, C. A. Townsend, *Bioorg. Med. Chem. Lett.* **1997**, *7*, 3129–3134.
- [59] W. H. Weinberg, E. McFarland, I. Goldwasser, T. Boussie, H. Turner, J. A. M. van Beek, V. Murphy, T. Powers (Symyx Technologies), WO 98/03521, **1998**.
- [60] L. M. Gayo, M. J. Suto, *Tetrahedron Lett.* **1997**, *38*, 513–516.
- [61] I. E. Pop, B. P. Déprez, A. L. Tartar, *J. Org. Chem.* **1997**, *62*, 2594–2603.
- [62] M. G. Siegel, P. J. Hahn, B. A. Dressman, J. E. Fritz, J. R. Grunwell, S. W. Kaldor, *Tetrahedron Lett.* **1997**, *38*, 3357–3360.
- [63] R. J. Booth, J. C. Hodges, *J. Am. Chem. Soc.* **1997**, *119*, 4882–4886.
- [64] A. Studer, S. Hadida, R. Ferritto, S. -Y. Kim, P. Jeger, P. Wipf, D. P. Curran, *Science* **1997**, *275*, 823–826.
- [65] D. P. Curran, M. Hoshino, *J. Org. Chem.* **1996**, *61*, 6480–6481.
- [66] M. Hoshino, P. Degenkolb, D. P. Curran, *J. Org. Chem.* **1997**, *62*, 8341–8349.
- [67] A. Studer, P. Jeger, P. Wipf, D. P. Curran, *J. Org. Chem.* **1997**, *62*, 2917–2924.
- [68] A. Studer, D. P. Curran, *Tetrahedron* **1997**, *53*, 6681–6696.
- [69] D. J. Gravert, K. D. Janda, *Chem. Rev.* **1997**, *97*, 489–509.
- [70] D. J. Gravert, K. D. Janda, *Drugs Fut.* **1997**, *22*, 1147–1150.
- [71] A. M. Vandersteen, H. Han, K. D. Janda, *Molecular Diversity* **1996**, *2*, 89–96.
- [72] H. Han, M. M. Wolfe, S. Brenner, K. D. Janda, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 6419–6423.
- [73] H. Han, K. D. Janda, *J. Am. Chem. Soc.* **1996**, *118*, 2539–2544.
- [74] H. Han, K. D. Janda, *Tetrahedron Lett.* **1997**, *38*, 1527–1530.
- [75] M. Hori, K. D. Janda, *J. Org. Chem.* **1998**, *63*, 889–894.
- [76] R. M. Kim, M. Manna, S. M. Hutchins, P. R. Griffin, N. A. Yates, A. M. Bernick, K. T. Chapman, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10012–10017.
- [77] Y. Pan (the Procter & Gamble Company), WO 98/03452, **1998**.
- [78] P-P Kung, R. Bharadwaj, A. S. Fraser, D. R. Cook, A. M. Kawasaki, P. D. Cook, *J. Org. Chem.* **1998**, *63*, 1846–1852.
- [79] T. A. Keating, R. W. Armstrong, *J. Org. Chem.* **1998**, *63*, 867–871.
- [80] S. S. Nikam, B. E. Kornberg, S. E. Ault-Justus, M. F. Rafferty, *Tetrahedron Lett.* **1998**, *39*, 1121–1124.
- [81] I. C. Cotterill, A. Ya. Usyatinski, J. M. Arnold, D. S. Clark, J. S. Dordick, P. C. Michels, Y. L. Khmel'nitsky, *Tetrahedron Lett.* **1998**, *39*, 1117–1120.
- [82] J. J. Parlow, W. Naing, M. S. South, D. L. Flynn, *Tetrahedron Lett.* **1997**, *38*, 7959–7962.
- [83] B. O. Buckman, M. M. Morrissey, R. Mohan, *Tetrahedron Lett.* **1998**, *39*, 1487–1488.
- [84] P. Hodge, *Chem. Soc. Rev.* **1997**, *26*, 417–424.
- [85] V. Molteni, R. Annunziata, M. Cinquini, F. Cozzi, M. Benaglia, *Tetrahedron Lett.* **1998**, *39*, 1257–1260.
- [86] P. D. Cook, A. M. Kawasaki, P. P. Kung (Isis Pharmaceuticals, Inc.), WO 98/05332, **1998**.
- [87] C. G. Blettner, W. A. König, W. Stenzel, T. Schotten, *Synlett*, **1998**, 295–297.
- [88] M. T. Fiorini, C. Abell, *Tetrahedron Lett.* **1998**, *39*, 1827–1830.
- [89] M. G. Siegel, A. J. Shuker, C. D. Droste, P. J. Hahn, C. D. Jesudason, J. H. McDonald, D. P. Matthews, C. J. Rito, A. J. Thorpe, *Molecular Diversity* **1998**, *3*, 113–116.
- [90] V. V. Mozhaev, C. L. Buddle, J. O. Rich, A. Ya. Usyatinsky, P. C. Michels, Y. L. Khmel'nitsky, D. S. Clark, J. S. Dordick, *Tetrahedron* **1998**, *54*, 3971–3982.



- [91] M. W. Creswell, G. L. Bolton, J. C. Hodges, M. Meppen, *Tetrahedron* **1998**, *54*, 3983–3998.
- [92] H. An, T. Wang, V. Mohan, R. H. Griffey, P. D. Cook, *Tetrahedron* **1998**, *54*, 3999–4012.
- [93] M. J. Suto, L. M. Gayo-Fung, M. S. S. Palanki, R. Sullivan, *Tetrahedron* **1998**, *54*, 4141–4150.
- [94] J. Dickhaut (Boehringer Mannheim GmbH), WO 98/06386, **1998**.
- [95] D. L. Boger, W. Chai, *Tetrahedron* **1998**, *54*, 3955–3970.
- [96] P. C. Tang, G. McMahon, K. P. Hirth, L. K. Shower (Sugen Inc.), WO 98/07695, **1998**.
- [97] K. E. Pryor, G. W. Shipps, D. A. Skyler, J. Rebeck, *Tetrahedron* **1998**, *54*, 4107–4124.
- [98] C. R. Johnson, B. Zhang, P. Fantauzzi, M. Hocker, K. M. Yager, *Tetrahedron* **1998**, *54*, 4097–4106.
- [99] D. G. Powers, D. S. Casebier, D. Fokas, W. J. Ryan, J. R. Troth, D. L. Coffen, *Tetrahedron* **1998**, *54*, 4085–4096.
- [100] Y. F. Xie, J. P. Whitten, T. Y. Chen, Z. Liu, J. R. McCarthy, *Tetrahedron* **1998**, *54*, 4077–4084.
- [101] G. R. Gustafson, C. M. Baldino, M.-M. E. O'Donnell, A. Sheldon, R. J. Tarsa, C. J. Verni, D. L. Coffen, *Tetrahedron* **1998**, *54*, 4051–4065.
- [102] J. J. Parlow, D. L. Flynn, *Tetrahedron* **1998**, *54*, 4013–4031.
- [103] R. A. Houghten, J. M. Ostresh (Trega Biosciences, Inc.), WO 98/11438, **1998**.
- [104] S. Kobayashi, S. Komiyama, H. Ishitani, *Biotechnol. Bioeng.* **1998**, *61*, 23–31.
- [105] S. E. Ault-Justus, J. C. Hodges, M. W. Wilson, *Biotechnol. Bioeng.* **1998**, *61*, 18–22.
- [106] B. A. Kulkarni, A. Ganesan, *Chem. Commun.*, **1998**, 785–786.
- [107] D. Fokas, W. J. Ryan, D. S. Casebier, D. L. Coffen, *Tetrahedron Lett.* **1998**, *39*, 2235–2238.
- [108] T. Wang, H. An, T. A. Vickers, R. Bharadwaj, P. D. Cook, *Tetrahedron* **1998**, *54*, 7955–7976.
- [109] J. Mendoza (Eli Lilly and Company) WO 98/21584, **1998**.
- [110] O. Hinds Gaul (Synsorb Biotech, Inc.) WO 98/22487 **1998**.
- [111] R. Maltais, D. Poirier, *Tetrahedron Lett.* **1998**, *39*, 4151–4154.
- [112] R. Baudelle, P. Melnyk, B. Déprez, A. Tartar, *Tetrahedron* **1998**, *54*, 4125–4140.
- [113] D. A. Klumpp, K. Y. Yeung, G. A. S. Prakash, G. A. Olah, *J. Org. Chem.* **1998**, *63*, 4481–4484.
- [114] G. E. Jagdmann (Eli Lilly and Company) WO 98/27087, **1998**.
- [115] J. S. Mendoza (Eli Lilly and Company) WO 98/27093, **1998**.
- [116] C. Blackburn, B. Guan, P. Fleming, K. Shiosaki, S. Tsai, *Tetrahedron Lett.* **1998**, *39*, 3635–3638.
- [117] C. D. Floyd, L. A. Harnett, A. Miller, S. Patel, L. Saroglou, M. Whittaker, *Synlett* **1998**, 637–639.
- [118] L. H. Caporale (CombiChem Inc.) US 5,767,238, **1998**.
- [119] D. L. Boger, W. Chai, Q. Jin, *J. Am. Chem. Soc.* **1998**, *120*, 7220–7225.
- [120] S. W. Kim, Y. S. Shin, S. Ro, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1665–1668.
- [121] P.-P. Kung, P. D. Cook, *Biotechnol. Bioeng. (Comb. Chem.)* **1998**, *61*, 119–125.
- [122] K. Berlin, R. K. Jain, C. Richert, *Biotechnol. Bioeng. (Comb. Chem.)* **1998**, *61*, 107–118.
- [123] D. L. Boger (The Scripps Research Institute), WO 98/35231, **1998**.
- [124] M. Panunzio, M. Villa, A. Missio, T. Rossi, P. Seneci, *Tetrahedron Lett.* **1998**, *39*, 6585–6588.
- [125] S. W. Kim, S. M. Bauer, R. W. Armstrong, *Tetrahedron Lett.* **1998**, *39*, 7031–7034.
- [126] M. Marder, H. Viola, J. A. Bacigaluppo, M. I. Colombo, C. Wasowski, C. Wolfman, J. H. Medina, E. A. Rúveda, A. C. Paladini, *Biochem. Biophys. Res. Commun.* **1998**, *249*, 481–485.
- [127] S. V. Ley, M. H. Bolli, B. Hinzen, A. -G. Gervois, B. L. Hall, *J. Chem. Soc., Perkin Trans. 1* **1998**, 2239–2241.
- [128] B. Hinzen, R. Lenz, S. V. Ley, *Synthesis* **1998**, 977–979.
- [129] M. H. Bolli, S. V. Ley, *J. Chem. Soc., Perkin Trans. 1* **1998**, 2243–2246.
- [130] C. Gennari, S. Ceccarelli, U. Piarulli, C. A. G. N. Montalbetti, R. F. W. Jackson, *J. Org. Chem.* **1998**, *63*, 5312–5313.
- [131] F. Hauernt, M. H. Bolli, B. Hinzen, S. V. Ley, *J. Chem. Soc., Perkin Trans. 1* **1998**, 2235–2237.
- [132] a) D. L. Boger, P. Ducray, W. Chai, W. Jiang, J. Goldberg, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2339–2344, b) D. L. Boger, J. Goldberg, W. Jiang, W. Chai, P. Ducray, J. K. Lee, R. S. Ozer, C. -M. Andersson, *Bioorg. Med. Chem.* **1998**, *6*, 1347–1378.
- [133] M. Falorni, G. Giacomelli, L. Mameli, A. Porcheddu, *Tetrahedron Lett.* **1998**, *39*, 7607–7610.
- [134] J. Habermann, S. V. Ley, J. S. Scott, *J. Chem. Soc., Perkin Trans. 1* **1998**, 3127–3130.
- [135] C. Hulme, J. Peng, B. Louridas, P. Menard, P. Krolkowski, V. Kumar, *Tetrahedron Lett.* **1998**, *39*, 8047–8050.
- [136] C. F. Sturino, M. Labelle, *Tetrahedron Lett.* **1998**, *39*, 5891–5894.
- [137] J. W. Nieuwenhuijzen, P. G. M. Conti, H. C. J. Ottenheim, J. T. M. Linders, *Tetrahedron Lett.* **1998**, *39*, 7811–7814.
- [138] A. G. M. Barrett, M. L. Smith, F. J. Zecri, *Chem. Commun.*, **1998**, 2317–2318.

- [139] J. J. Parlow, M. L. Vazquez, D. L. Flynn, *Bioorg. Med. Chem. Lett.* **1998**, 8, 2391–2394.
- [140] H. An, B. D. Haly, P. D. Cook, *Bioorg. Med. Chem. Lett.* **1998**, 8, 2345–2350.
- [141] G. W. Starkey, J. J. Parlow, D. L. Flynn, *Bioorg. Med. Chem. Lett.* **1998**, 8, 2385–2390.
- [142] D. L. Coffen, J. C. Hogan (Arqule, Inc.), WO 98/46551, **1998**.
- [143] G. M. Coppola, *Tetrahedron Lett.* **1998**, 39, 8233–8236.
- [144] A. M. Hay, S. Hobbs-DeWitt, A. A. MacDonald, R. Ramage, *Tetrahedron Lett.* **1998**, 39, 8721–8724.
- [145] D. E. Bergbreiter, Y. -S. Liu, S. Furyk, B. L. Case, *Tetrahedron Lett.* **1998**, 39, 8799–8802.
- [146] A. R. Tunoori, D. Dutta, G. I. Georg, *Tetrahedron Lett.* **1998**, 39, 8751–8754.
- [147] K. D. Combrink, H. B. Gülgeze, N. A. Meanwell, B. C. Pearce, P. Zulan, G. S. Bisacchi, D. G. M. Roberts, P. Stanley, S. V. Seiler, *J. Med. Chem.* **1998**, 41, 4854–4860.
- [148] U. J. Nilsson, E. J. -L. Fournier, O. Hindsgaul, *Bioorg. Med. Chem.* **1998**, 6, 1563–1575.
- [149] A. L. Castelhana, B. McKibben (Cadus Pharmaceutical Corp.), WO 98/54116, **1998**.
- [150] B. Raju, J. M. Kassir, T. P. Kogan, *Bioorg. Med. Chem. Lett.* **1998**, 8, 3043–3048.
- [151] C. Hulme, J. Peng, S. -Y. Tang, C. J. Burns, I. Morize, R. Labaudiniere, *J. Org. Chem.* **1998**, 63, 8021–8023.
- [152] P.-C. Pan, C.-M. Sun, *Tetrahedron Lett.* **1998**, 39, 9505–9508.
- [153] S. W. Kaldor, M. G. Siegel, *Current Opin. Chem. Biol.*, **1997**, 1, 101–106.
- [154] N. M. J. Vermeulin, C. L. O'Day, H. K. Webb, M. R. Burns (Oridigm Corp.), WO 99/03823, **1999**.
- [155] M. Adamczyk, J. F. Fishpaugh, P. G. Mattingly, *Tetrahedron Lett.* **1999**, 40, 463–466.
- [156] P. Brooking, A. Doran, P. Grimsey, N. W. Hird, W. S. MacLachlan, M. Vimal, *Tetrahedron Lett.* **1999**, 40, 1405–1408.

This Page Intentionally Left Blank

# 4 Combinatorial Chemistry of Multicomponent Reactions

*Ivar Ugi, Alexander Dömling and Birgit Ebert*

*In memoriam of Sir Derek H. R. Barton*

## 4.1 Classical and Modern Chemistry of Isocyanides and Multicomponent Reactions

The classical syntheses correspond to multistep reactions of one or two educts. These were introduced in the first half of the last century [1]. At this time, the chemistry of the multicomponent reactions (MCRs) was also initiated, whose products are essentially formed by condensing reactions from more than two educts with reactive functional groups. In this way, the resulting products contain at least some part of all educts [2–6]. Essentially, the early chemistry of the one-pot MCRs related to  $\alpha$ -aminoalkylation of nucleophiles and the secondary formation of heterocycles [7]. The chemistry of the isocyanides [8] also began at this time, and in 1921 the discovery of P-3CR (P = Passerine) [8–10] was a unique exception, as for a long time this was not only the first but also the only MCR of the isocyanides. Fifty years later, in the 1970s, the chemistry of the four-component isocyanide MCRs and their libraries was introduced. But by the 1990s, this has led to essential changes in synthetic chemistry such that the reaction has become one of the most fundamental facets of industrial chemical research [6].

The two early types of MCRs remained separate entities for a whole century [7, 8], but in January 1959 [11] they were combined [12] and became the U-4CR (U = Ugi) [2–4, 8]. This soon also included the libraries of combinatorial chemistry [8, 13a]. The educts and products of the U-4CR and its unions with other chemical reactions are by many orders of magnitude more variable than those of any other chemical reactions, and even of the entire spectrum of other MCRs. Under suitable reaction conditions, the U-4CR may produce almost quantitative yields of pure product, and with minimal preparative work [4, 8, 10]. Nevertheless, this profound preparative progress was of no general chemical interest until early 1995 [2–5, 14], when almost overnight the one-pot MCR chemistry of the isocyanides and their libraries became one of the most intensely used areas of industrial chemical research. The technique is used not only to find new chemical products, but also to produce them far more conveniently and in higher yields than might other multistep syntheses.

Ten years later there emerged a different, albeit now very popular type of chemical process, whose products are formed by the polycyclization of unsaturated bonds [15, 16]. These so-called ‘tandem’ [16, 17] or ‘domino’ [18] procedures differ profoundly from those of the MCRs. Although both the MCRs and the tandem–domino chemistry are one-pot reactions that form many new chemical bonds [19], they differ fundamentally in all other important aspects. These two methods of chemistry are, therefore, complementary.

This somewhat new area of preparative chemistry corresponds to syntheses of polycyclic natural and related compounds, the essential step of which always corresponds to a polycyclization of unsaturated educts. These are often associated with further derivatives, which undergo secondary reactions.

The industrial quest for new, desirable chemical products is now accomplished by the combinatorial chemistry of the MCRs and their libraries, including those of the isocyanides [20–22]. The reactions of the isocyanides proceed especially well, as all correspond to irreversible transitions of divalent carbon atoms, C<sup>II</sup>, into their C<sup>IV</sup> atoms. If no by-products are formed by some competing reactions, then quantitative yields of MCR products are formed.

Nowadays, the search for new products using the chemistry of MCR libraries, and particularly using automated equipment [20–22], can be accomplished up to 1000 times faster than by any previous methodology. Thus, a much greater number of new chemical compounds has already been formed than by any previous approach.

This chemistry will not be replaced in the near future by some other methodology, as no other more efficient procedure is yet known, or is predictable. Any widely used 'new' method of progress is based on a concept which has been proposed many years previously, and the chemical industry was quick to recognize the great advantages of producing and investigating products using one-pot MCRs and their libraries, especially if isocyanides are involved. Hence, this area is one of the most active in current chemical research [23].

As we enter the next millennium, the chemistry of MCRs and their libraries will furnish new ideas that will lead to the creation of new types of desirable products, and in an increasingly efficient manner [20].

## 4.2 Early Studies and Concepts of MCR Chemistry

Although organic chemistry as a science, including that of the MCRs, began many billions of years ago [24], the first MCR was accomplished in 1838 [7, 25]. Progress in MCRs was rather slow for the next century, but in the past few years the combinatorial chemistry of the MCRs has become unusually active, and their libraries are currently a vital area of industrial chemical research [22]. Following the identification of MCRs, it was soon realized that this was a particularly efficient method of searching for new products [3–5, 20–22], as chemical compounds formed by a suitable MCR were always synthesized more efficiently than by any other multistep procedure.

When life on earth first began, many sequences of chemical reactions formed a vast variety of biochemical compounds as their libraries. Within the natural atmosphere many chemical compounds were formed, and even today the chemistry of nature proceeds not only by reactions of two components, but also by MCRs of three and more educts. In this way, many different chemical compounds are formed simultaneously as libraries. In his famous experiments, Miller showed how some  $\alpha$ -amino acids and their derivatives were formed by the S-3CR (S = Strecker) [7, 31].

It is assumed that adenine, an essential component of DNA and RNA, was formed from five units of hydrogen cyanide, and this process undoubtedly played a major role in the development of the first living cells [24].

It is highly probable that in nature, no adequate liquid-phase reactions could proceed, as sufficient concentrations of educts are not usually in solution. It is likely that such reactions occurred on solid surfaces where sufficient concentrations of participating chemical compounds could exist [24a].

As living cells first began to exist, they contained libraries of the products of many different biochemical MCRs. Today, these are still formed by sequences of enzymatically accelerated selective subreactions, and continuously generate vast collections of chemical compounds as libraries. However, if these enzymatic processes are insufficiently specific, then undesirable impurities may also be formed, but these will be removed selectively by other enzymatic reactions [27].

Although the recent activities of MCRs and their libraries have been mainly initiated by the chemical industry [22, 27], the basic research of preparative MCR chemistry [8, 10], together with the techniques' mathematically oriented computer methodology [28, 29], was developed long before MCR chemistry was considered a viable scientific prospect.

During this time, the one-pot MCRs and their libraries became widely used, and were not merely fashionable on a temporary basis, as were many previous methods of chemical research. MCR chemistry has many basic advantages over conventional methods of chemistry, and is now widely used as it incorporates areas of science and technology such as biology, physics, mathematics, computer methodology and automation methods [20, 21]. It is inconceivable that MCR chemistry will be superseded by alternative technologies, and the technique is now applied more intensely than any other in the search for new products [2, 6, 22].

### 4.3 Conceptual Differences between Conventional Chemical Reactions and MCRs

In principle, all chemical reactions correspond to equilibria between one or two educts and their products. Their preferred preparative processes are virtually irreversible reactions, and if no competing formation of by-products occurs, then quantitative yields of pure products can result.

Conventional syntheses are conversions of  $N \geq 3$  different starting materials into target compounds of  $M \geq N-1$  preparative steps. Such sequences of reactions include the isolation and purification of intermediates and final products. After each such step, its resulting compound must be combined with the next educt, which ultimately leads to the final product. The amount of preparative work increases with each preparative step, and thus its resulting final yield decreases.

In contrast, any one-pot MCR is a conversion of  $N \geq 3$  different starting materials into its products, and must contain at least some part of each educt. No MCR can directly convert its educts into its products, as this corresponds to many steps, so that one or two components of their sub-reaction also participate. Therefore, this is always a sequence of elementary reactions. Only one direct chemical reaction of three participating components seems possible, namely the formation of the subsequently rearranging  $\alpha$ -adducts of the isocyanides, the  $\alpha$ -aminoalkyl cations **5** and the anions **6**. Such a three-component reac-

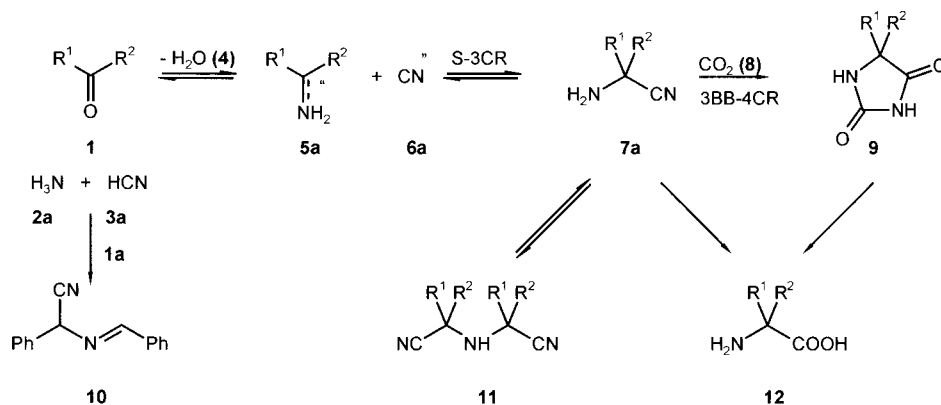
tion is conceivable, as these pairs of ions can be closely tied together, particularly if high concentrations of such ions participated in suitable solvents [8, 28].

The MCRs can lead to quantitative yields of products, if their early steps equilibrate, though their final product-forming steps are in practice irreversible. In contrast to the multistep syntheses, a one-pot MCR requires almost no work, and usually also requires much less chemical materials, including solvents.

Normally, a 'library' is a room that contains many books, but in the past decade the term has been used to define a collection of different chemical compounds, while its 'combinatorial chemistry' enables it to form many different chemical compounds using the same type of preparative procedure [2, 6].

#### 4.4 The Different Types of MCRs

MCR chemistry first began when Laurent and Gerhardt [25] recognized that ammonia **2a** reacts with bitter almond oil, forming first glucose, benzaldehyde **1a** ( $R^1 = H$ ;  $R^2 = Ph$ ) and hydrogen cyanide **3a**, whose components **1a–3a** are converted into the intermediate product **7a** ( $R^1 = H$ ;  $R^2 = Ph$ ) by a 3CR, whose primary product **7a** undergoes a secondary reaction with additional **1a** and forms the crystalline product **10** [7]. This was the first example of the well-known name reaction, S-3CR.



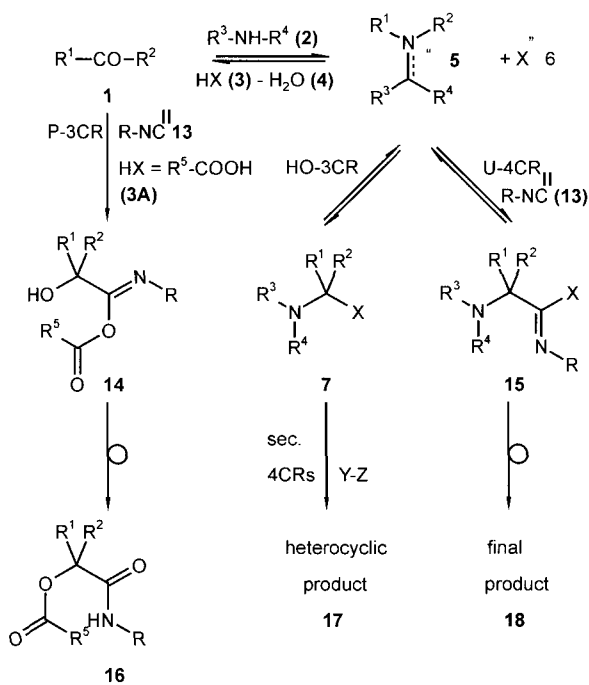
**Scheme 4.1.**

Twelve years later, the general reaction **1** + **2a** + **3a** → **7a** (usually  $R^1 = H$ ;  $R^2 = alkyl$ ) was introduced under the name of S-3CR [26]. Often, toxic by-products like **11** are formed, which depends on the structures of their participating carbonyl compounds **1**. Nevertheless, the  $\alpha$ -amino acids **12** were prepared by the S-3CR and the subsequent hydrolysis of their products **7a**.

In 1929, Bergs introduced the syntheses of the hydantoin derivatives **9**, and shortly after this method was improved by Bucherer [30]. The chemical industry preferentially produces the  $\alpha$ -amino acids **12** by BB-4CR, whose components are hydrolyzed, and hence excellent yields of pure  $\alpha$ -amino acids can be obtained [31].

The BB-4CR corresponds to the equilibrating S-3CR whose product **7a** reacts further with  $\text{CO}_2$ . In the final step, an irreversible ring closure takes place. This fact might have led to the insight, that two different basic types of MCR exist, and a general insight that irreversible MCRs have great preparative advantages over equilibrating MCRs.

In the 1960s it became clear that, in general, several different types of MCRs can take place. The U-4CR was discovered and its reaction mechanism investigated. It was realized that the early part of the U-4CR corresponds to an equilibrium of a 3CR, where **1-3** and **4-6** participate. Subsequently, **5** and **6** form the  $\alpha$ -adduct **15** of an isocyanide **13**, and this intermediate rearranges into its final product [8, 10, 28].



Scheme 4.2.

Thus, the U-4CR and some related reactions of the isocyanides can form a wide variety of structurally different products, as the final steps of their rearranging reaction mechanisms of **15** into their products **18** can differ profoundly, depending of the nature of the acid component **6**.

We now know that the S-3CR, the BB-4CR and the U-4CR correspond to different types of MCRs. This led to the insight that all MCRs correspond to collections of sub-reactions, some of which equilibrate, while others proceed irreversibly. Their collections determine the purity and the yield of its final product. Consequently, several totally different types of MCRs can take place:

- Type IA: All starting materials, intermediate and final products participate with mobile equilibria, so that their products are usually not isolatable.



- Type IB: All starting materials, intermediates and products participate with mobile equilibria, but the products are so stable that they can be isolated.
- Type IIA: The intermediate products of type IIA of IB react with a further educt, forming the final products irreversibly. The number of formed chemical bonds increases.
- Type IIB: The pre-final products of type IA or IB react irreversibly with further multifunctional components and form heterocyclic products irreversibly. The formed number of chemical bonds remains constant.
- Type III: The educts of MCR form one-pot products by sequences of irreversible sub-reactions.

Most MCRs of type IA are 3CRs whose equilibrating products react with further educts by higher MCRs of type II. In 1960, Hellmann and Opitz published their  *$\alpha$ -Aminoalkylierung* book [7], which was a comprehensive list and discussion of all so-far-known MCRs. They realized, as did others previously, that the ‘name reactions’ [2, 5] S-3CR, the M-3CR [32] and many other 3CRs belong to the mechanistic family of  $\alpha$ -aminoalkylations of nucleophiles, which are usually the anions of deprotonated weak acids components. This collection of 3CRs is now also referred to as the HO-3CR [4].

This formation of heterocycles by 4CRs was first performed in the 1880s. The resultant  $\alpha$ -aminoalkyl compounds **7** of type I can also be intermediate products that react with further bifunctional, five- or six-membered heterocycles **17** forming educts **Y–Z**. Usually, the final ring-forming step of such 4CRs, e.g., Hantzsch, Biginelli, Bergmann–Bucherer and Asinger [7], proceed irreversibly and are thus MCRs of type II. Under certain instances the A-3CR and A-4CR [33] and the related N and O reactions [34] seem to be reversible.

Most MCRs of types I and III do not form their products well. Any chemical compound that can be formed by an MCR of type II can be accomplished better than by any other chemical reaction or multistep sequence of preparative procedures. In modern, often automated preparative chemistry the MCRs of type IIA and IIB and their libraries are now used to great extent, and are more efficiently accomplished than ever before [20, 21].

Some type I MCRs correspond to collections of equilibrating sub-reactions, including the formation of their products, such reactions often having basic preparative disadvantages. Type II MCRs have many ‘early’ equilibrating sub-reactions, but their prefinal and/or final products can subsequently follow irreversible final steps. A wide variety of educts can be converted into their products efficiently by MCRs of type II. Only a few preparative procedures by type III MCRs [35] could be developed, and reactions forming unpleasant by-products were often encountered.

In living cells, however, where the majority of chemical compounds are formed by type III MCRs, the sub-reactions are largely selectively accelerated by suitable enzymes. Usually, they proceed only in one desirable direction, though if in an exceptional case an undesirable by-product is formed, it can be removed using selective procedures [27].

## 4.5 The First Century of Isocyanide Chemistry

Until the late 1950s, the chemistry of the isocyanides had remained a rather ‘empty’ area of organic chemistry, despite several important reactions of the isocyanides having

been introduced and investigated. In most textbooks, the isocyanides were mentioned only very briefly, if at all.

The chemistry of the isocyanides began in earnest in 1859, when Lieke [8, 36] prepared allyl isocyanide from allyl iodide and silver cyanide. Eight years later, Meyer [37] produced the methyl and ethyl isocyanides. Such syntheses of the alkylisocyanides were generalized in 1867 by Gautier [38], who with Hofmann [39] introduced the preparation of the isocyanides by mixing primary amines, chloroform and alkali (thus aryl isocyanides could also be produced). Gautier then also attempted to dehydrate the formats of the primary amines using  $P_2O_5$ , in order to produce the isocyanides. These were not obtained, as the isocyanides cannot exist in an acidic media [28], and equivalents of base had not been added to the reaction.

For many years, the availability of the isocyanides was very poor, as only these two classical methods of preparing them were known, and each had moderate yields. The smell of the then-available isocyanides was described as 'horrifying' and having 'continuing complaints in the neighborhood about the vile odor' [8]. It was also assumed that the isocyanides were generally toxic.

It is no wonder therefore that very few investigations were made on the reactivity of isocyanides. However, in 1910 Oliveri-Mandala and Alagna [40] discovered the formation of the N-substituted tetrazoles by the  $\alpha$ -additions of hydrazoic acid onto the isocyanides and the subsequent rearrangements into their final products.

Between 1921 and 1931, Passerini introduced and investigated the reaction which bears his name [8–10], forming the products **16** from carbonyl compounds **1**, carboxylic acids **3A** and isocyanides **13**. The discovery of the P-3CR was a milestone in the development of isocyanide-based MCRs, as it was the introduction of the first practicably irreversible MCR of these compounds.

In the 1940s, Baker and Linn [41], and later Dewar [42], each proposed mechanisms of the P-3CR. The important role of the intermediate hydrogen bond between the carboxylic acid and the carbonyl compound in suitable solvents was mentioned. The most plausible mechanism of the Passerini reaction is depicted in Scheme 4.2 (**1** + **3A** + **13**  $\rightarrow$  **16**) [43].

The greatest variety of the MCRs of type II are those of the isocyanides **13**. They are the only compounds in organic chemistry, that contain a divalent carbon atom  $C^{II}$ . All of its chemical reactions correspond to strongly exothermic transitions of  $C^{II}$  into  $C^{IV}$  [8]. Usually, an  $\alpha$ -adduct of the isocyanide is formed, such as **15**. This can be an isolatable compound, or it may undergo rearrangement into its final product **18**.

It was recognized early on that the functional isocyano group differs profoundly from any other functional class of chemical compounds. In the 1890s, Nef [44] had mentioned the fact that the functional group  $-NC$  of the isocyanides contains a divalent carbon atom  $C^{II}$ , and therefore their chemistry differs very much from other stable organic chemical compounds that contain only tetravalent carbon atoms  $C^{IV}$ . Any synthesis of the isocyanides corresponds to a conversion of  $C^{IV}$  into  $C^{II}$ , and all characteristic chemical reactions of the isocyanides correspond to exothermic irreversible transitions of the carbon atoms  $C^{II}$  into  $C^{IV}$ . Under the correct conditions, when no competing formation of by-products takes place, such reactions yield the products quantitatively.

Any products that can be formed by a one-pot U-4CR, or related MCRs of type IIB [3, 4], can also be formed as a library of that certain backbone type. This has profound prep-

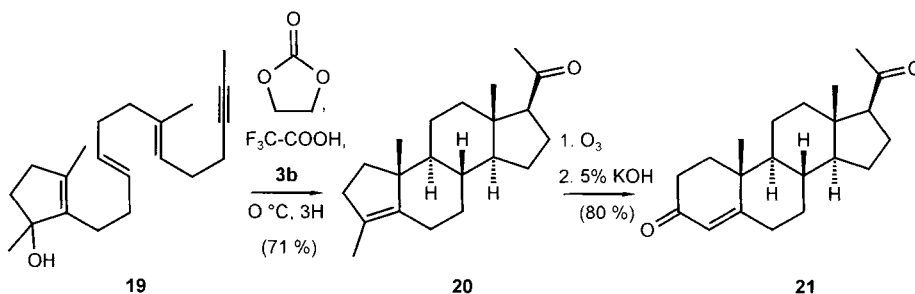
arative advantages over any multistep sequence of conventional chemical reactions of one or two components per step.

#### 4.6 Complementary Aspects of Natural Product Syntheses by Tandem–Domino Reactions and MCR Chemistry

In 1986, Posner [16, 19] introduced the general scheme of preparing a variety of natural and structurally related chemical compounds by ‘multicomponent one-pot annulation forming three to six bonds’ by converting educts with many multiple bonds into their polycyclic products with five- and six-membered rings. These syntheses were accomplished by polycyclizations of highly unsaturated educts. Then, many new bonds are formed in one step and, surprisingly, the product appears very different from the starting material.

Although the ‘tandem–domino’ chemistry by such polycycloadditions reduces the total number of steps for preparative procedures, a profound disadvantage of this approach is that the starting material for this type of conversion can only be made in a time- and material-consuming multistep synthesis. The MCR starting materials are either produced by a short synthesis, or are available commercially in a wide variety. Also, MCRs are far more variable than the ‘tandem–domino’ syntheses, as more different products can be prepared by the one-pot MCRs than by any other one-step procedure. About 30 different backbones are known to be accessible by isocyanide-based MCRs, and the number of novel backbones is growing continuously. Thus, isocyanide-based MCRs provide access to the largest chemical space of different products known in total organic chemistry.

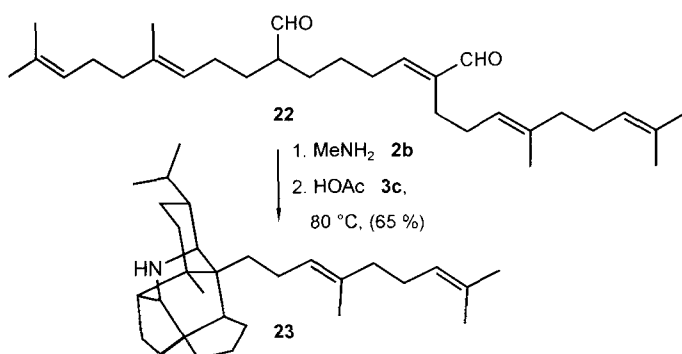
The idea of preparing polycyclic compounds from educts with many multiple bonds was stimulated by the formation of steroids in living cells via the biochemical conversion of squalene epoxide into lanosterol, as discovered in the 1950s [45]. In the early 1970s, Johnson [15] began the new technique of annulating syntheses of polycyclic products by synthesizing tetracyclic progesterone **20** from the highly unsaturated educt **19**, using its many new C–C bonds to induce polycycloaddition. As can be seen by the number of relevant reviews and original articles, this chemistry is very ‘en vogue’, and is entirely comprehensive with regard to the often unexpected and beautiful transformations.



Scheme 4.3.

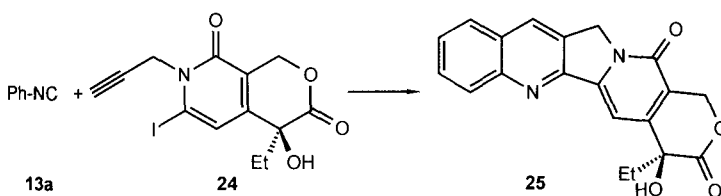
Often, tandem–domino chemistry has essential aspects in common with the polymerization of olefines. There are also some common features with MCRs, as both types of reaction form many new bonds. However, MCRs correspond to the conversions of many educts into products by exchange of functional groups, and not by cyclo-additions.

The modern annulating preparative chemistry is illustrated by the unique one-pot, one-step synthesis of the dihydroprotodaphniphyllite **23** [46]. It is noteworthy that the preparation of the essential starting material **22** corresponds to six preparative steps that are not at all trivial.



**Scheme 4.4.**

Nowadays, many elegant syntheses of steroids, alkaloids and affiliated products are prepared from several five- or six-membered rings. Usually, the preferred steps of such syntheses correspond to the polycyclizations of educts, that contain many multiple bonds, and one or two other functional groups of further educts can often also participate [17]. However, in recent years a new type of the radical chemistry of the isocyanides has been introduced [47]. This is illustrated by the one-pot synthesis of (20*S*)-camptothecin **25** from phenyl isocyanide **13a** and **24**. Such preparations of cyclopenta-fused quinolines are seldom found in tandem–domino chemistry [16], and this procedure is also closely related to the MCRs of the isocyanides. The previous multistep synthesis of **25** was accomplished by Danishefsky and Volkmann [47b], and uses a large number of linear synthetic steps with only small overall transformations towards the target.

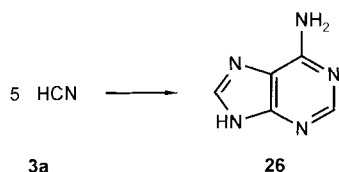


**Scheme 4.5.**

Often, such elegant syntheses of products need not only awkward steps of forming polycyclic compound from their simpler educts, but also much preparative work, as most

of their polycyclic educts must be prepared by difficult procedures. The excellent, creative proposal of Posner [19] has played an important role in this new era of preparing polycyclic products. In this article, some syntheses were also included that were polycondensations, and these differed somewhat from polycycloadditions by pairs of unions of M-3CR [19, 48, 49] which are not polycycloadditions of tandem-domino chemistry.

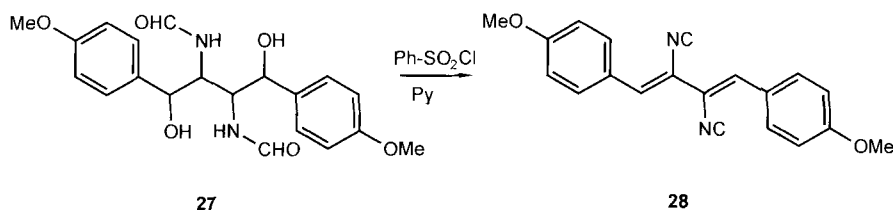
An unusual annulating reaction is the formation of adenine **26** from five units of hydrogen cyanide **3a**. This is almost a MCR, particularly as each of its five educts reacts in a different way [50], whereas the educts of MCRs have also different functional groups of educts. The formation of adenine is also closely related to some type of polymerization, as the unsaturated educts form their chains of molecules by repetition of analogous steps.



Scheme 4.6.

## 4.7 Modern Synthesis of the Isocyanides

In 1948, Rothe [51] discovered the antibiotic xanthocillin, a di-isocyanide which is formed biosynthetically from tyrosine [52] by cultures of *Penicillium notatum* and *Penicillium chrysatum*. Presumably, these isocyanide groups are formed by the dehydration of their formylamine groups through the polyphosphate derivatives such as ATP.

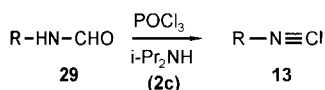


Scheme 4.7.

The O,O'-dimethyl-Xanthocillin **28** was prepared from its N, N'-diformyl-diamine **27**. This was dehydrated by phenyl-sulfonyl-chloride in the presence of pyridine. The procedure was published in *Pharmazie* [53] in East Germany in 1956, but was not available to Western countries. The method can be used particularly if isocyano-polymers are formed [54]. Two years later, Hertler and Corey [55] reported a similar conversion of a formylamino-steroid into its isocyano-derivative.

In 1957, Ugi et al. compared the spectroscopic data of comparable tetrazole and pentazole derivatives [56], and noted that the 1-substituted tetrazoles were easily produced from isocyanides and hydrogen acid [29], but that the desired isocyanides were less easily prepared. Ugi then successfully investigated the formation of isocyanides by dehydrating

their formylamines [8, 57], testing many different reagents and methods of preparing the isocyanides **13** [58, 59]. In 1985 it was found that the combination of POCl<sub>3</sub> and di-isopropylamine **2c** (Scheme 4.8) can convert the majority of formylamines into the isocyanides, usually with a highly satisfactory yield [60].



**Scheme 4.8.**

An increasing number of different isocyanides was produced, such that 329 known isocyanides were described in *Isonitrile Chemistry* in 1971 [8]. By comparison, during the era of classical isocyanide chemistry before 1960, only about 12 isocyanides were known.

It was then also recognized [61] that many isocyanides – especially the di- and tri-isocyanides – have plant-protecting and antibiotic properties. Moreover, it was also realized that these properties remained active, even after 200 generations.

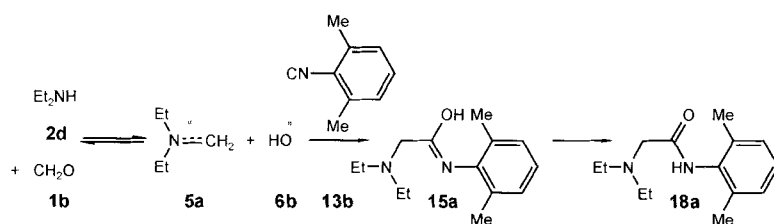
In the past few years, many new chemical compounds have been made prepared by the MCRs of the isocyanides than by all other types of chemical reactions combined, as isocyanide chemistry – and especially one-pot MCR procedure – has far greater variability than does any other facet of synthetic chemistry.

Barton et al. [62] and Baldwin et al. [63] each produced many interesting isocyanides that are natural products or related compounds. In recent years, Chang and Scheuer [64] discovered products with one to three isocyanic groups as natural products of the marine life in the Pacific Ocean. Since 1961, many isocyanides have been monitored toxicologically by Bayer AG, and shown to be non toxic in mammals [8]. Such non toxicity is confirmed by the occurrence of isocyanides in living organisms, and especially in marine animals [6]. In recent years, very many different mono-, di- and tri-isocyanides have been prepared and tested. Among these, more than 30 have unusual antibiotic and/or fungicidal and/or cytostatic properties [65], and may well represent successful future products for either pharmaceutical use or plant protection.

## 4.8 The Introduction of the New Isocyanide MCRs and their Libraries

After the Second World War, the local anesthetic Xylocain™ was introduced by A.B. Astra, Sweden, and this rapidly gained a major international market. Xylocain remains a particularly suitable pain-protecting product, notably for tooth extraction. Thus, Astra had achieved an exclusive business situation, as all 27 conceivable chemical procedures of forming Xylocain **18a** were protected by suitable patents [66].

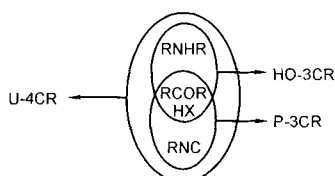
As soon as the isocyanides were easily available [8], Ugi and his co-workers discussed their potential applications and recognized the possibility to produce Xylocain from **1b**, **2b** and the isocyanide **13b**. Indeed, in January 1959 Xylocain was produced in feasible amounts by a one-pot reaction [11].



Scheme 4.9.

Formaldehyde **1a** and the secondary amine diethylamine **2b** equilibrate with the  $\alpha$ -aminoalkylation **5a** and the OH-anion **6b**, which can form the  $\alpha$ -adduct **13b** of the isocyanide **13b** whose OH rearranges into its NH-product **18a**. Since its OH-anion **6a** comes from water **4**, it became clear that this reaction was not a 3CR à la HO-3CR, but that a 4CR took place.

Since 1962, this 4CR has been officially referred to as the Ugi reaction [10a, 67–70], abbreviated as 4CC, U-4CC [2a, 8] and U-4CR [2b, 4, 6]. The U-4CR corresponds essentially to a union [12] of the HO-3CR and the P-3CR [8, 10].



Scheme 4.10.

A. B. Astra soon realized that Ugi's group had introduced this new 4CR, but had rather different interests and concepts. The company was mainly interested in avoiding competition from others, whereas Ugi proposed that, by this new 4CR, similar collections of compounds with the same reaction and thus backbone could be produced simultaneously. These compounds should be tested in order to find many new chemical entities with similar or even better biological properties. Hopefully, many local anesthetics can be discovered in future using this approach.

While working as a consultant at A. B. Astra, Ugi proposed ideas of parallel synthesis of Xylocain analogs by the 4CR. Obviously, the chemists could not detect an advantage over the classical stepwise and one-at-a-time synthesis of compounds. Moreover, they were not totally wrong at the time, as the entire collection of complementary methods required was not yet known. However, many decades later in the chemical industry the formation of such scaffolds of U-4CR products became one of the most widely and efficiently applied methods of searching for new desirable chemical compounds. Ugi assumed that this idea was, in principle, realistic, but it had been proposed too early, especially as since the necessary methods and equipment did not yet exist. The chemistry of MCR libraries had not yet been applied, though its essential formulation was mentioned in 1961 [13a], and again in 1971 [8]. Despite this, the idea was not realized for many decades. Perhaps reassuringly, 16 related new local anesthetics have been produced in recent years, besides oth-

er new products, and these were most likely discovered using the chemistry of the U-4CR and its product libraries [6].

## 4.9 The Great Variability of the U-4CR

Following the early studies of the U-4CR, it was realized that the yields and purities of their products depend very much on their reacting conditions, and in particular on the concentrations and ratios of their educts [28].

In 1963, McFarland [69] investigated the preparation of tetracycline derivatives by U-4CRs and found that besides its expected products, which depend on the reaction conditions, different concentrations of a great variety of by-products were formed [71–73]; in addition to the usual by-products, autoxidized compounds may be formed, particularly if aldehydes of the U-4CR proceed slowly, and oxygen is present [74].

Despite claims of 'Improved Ugi-reactions', no generally preferable method for the U-4CR exists whose conditions optimally convert the educts into their particular products. Therefore, satisfactory preparative conditions must be confirmed by suitable experimental prestudies [3, 8, 10, 73]. Usually, each educt and product of a chemical reaction has its characteristic structural features, whereas those of the U-4CR and related MCRs can have a range of different types of molecular arrangement. Hence, as many different combinations may participate, a much wider variety of U-MCR products and their libraries can be formed than by the entire range of chemical reactions [2]. The U-4CR can proceed stereoselectively, and sometimes even stereospecifically [5, 75]. All chemical reactions that correspond to the highly exothermic transitions  $C^{II} \rightarrow C^{IV}$ , and also their sterically 'overcrowded' products **18b–e**, can be formed from the sterically hindered combination of educts, that can be 'driven' by the energetic force of the isocyanide of their U-4CRs. The products shown in Figure 4.1 can be formed successfully by the U-4CR, but not by the usual preparative methods [10, 75–79].

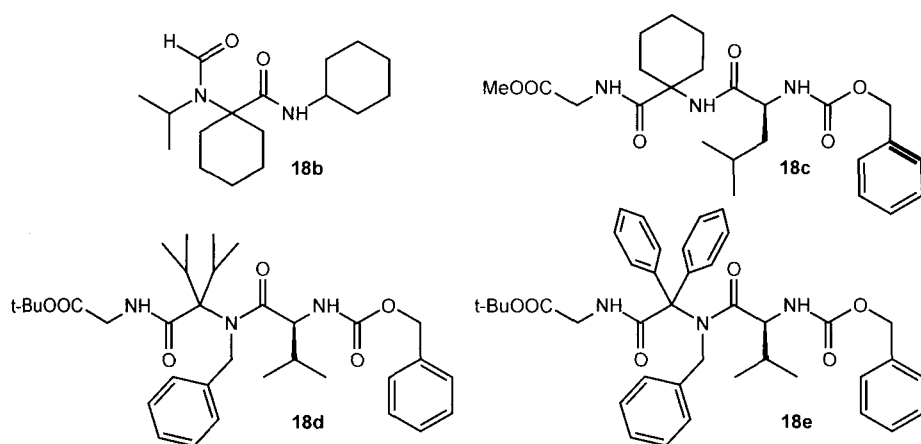
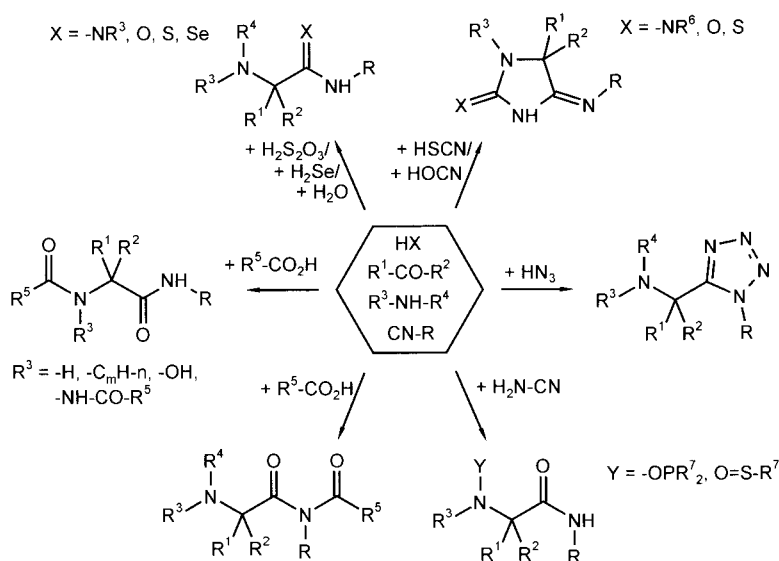


Figure 4.1.



Since 1995 the MCR chemistry of the isocyanides and their libraries has become one of the most popular and important areas of organic chemistry, notably as the products can be synthesized and investigated much more rapidly and efficiently than can those created by previously used chemical methods.

The U-4CR converts carbonyl compounds **1**, amines **2**, a great variety of acid components **3** and isocyanides **13** into their  $\alpha$ -adducts **15**, which then rearrange into different types of product **18** (Schemes 4.2 and 4.11) [6]. The U-4CRs can form products which are structurally more different than those of other chemical reactions, mainly due to the different types of acid components involved **3** [8, 10, 72, 73].



**Scheme 4.11.**

In addition, the wide variety of amine components **2** and related educts can participate in U-4CR to form many constitutionally different types of product (Fig. 4.2) [4, 6].

Furthermore, stereoselective U-4CRs can form an even greater variety of products. Occasionally, such reactions can act stereospecifically, this being accomplished efficiently if chiral alkylamine components **2** can participate under suitable reaction conditions [10]. Only in a few exceptional cases may some P-3CRs of chiral isocyanides proceed with efficient stereoselectivity [80].

The U-4CR and related MCRs can form almost quantitative yields of pure products if optimal reaction conditions are used, though the search for these often requires difficult experimental procedures.

Scheme 4.12 shows that at 0 °C in methylene chloride the educts **1c**, **2e**, **3d** and **13c** of the U-4CR form an almost quantitative yield of the P-3CR product **16a**, whereas in methanol equally well. Thus, the product **18f** is formed [73].

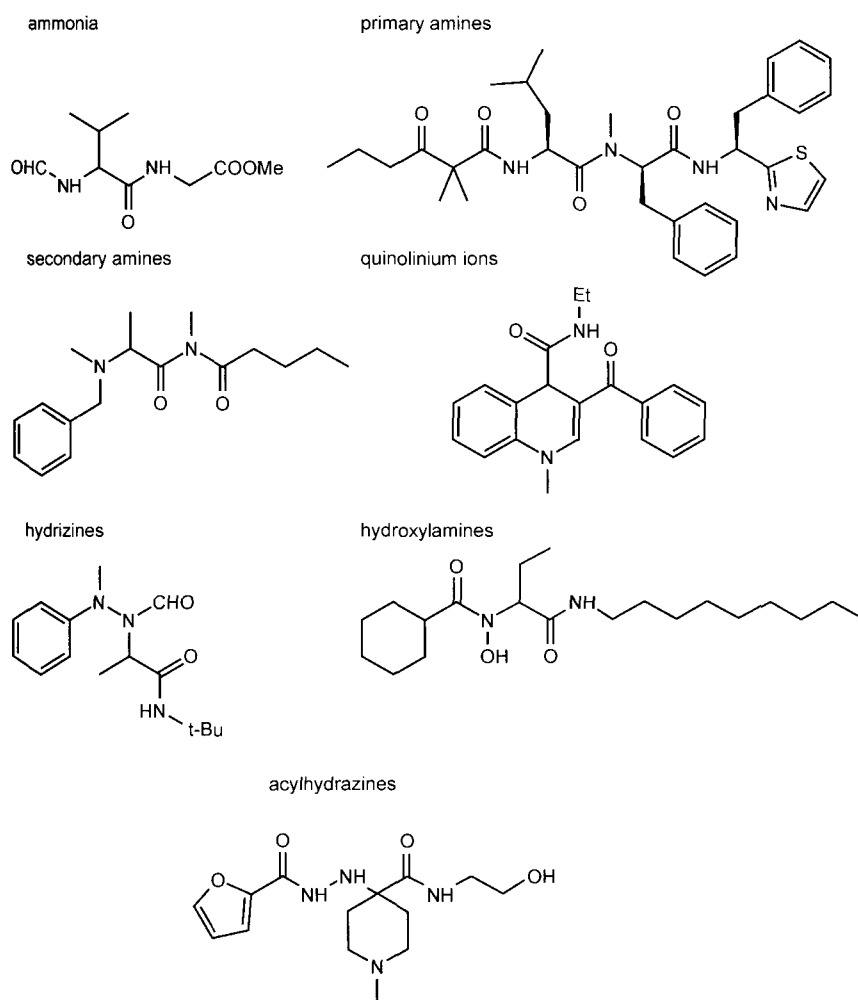
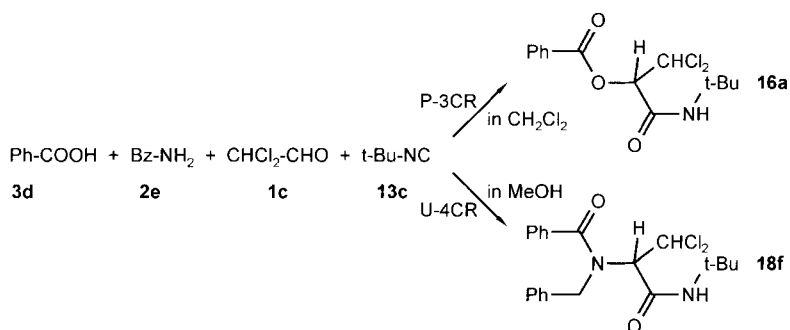


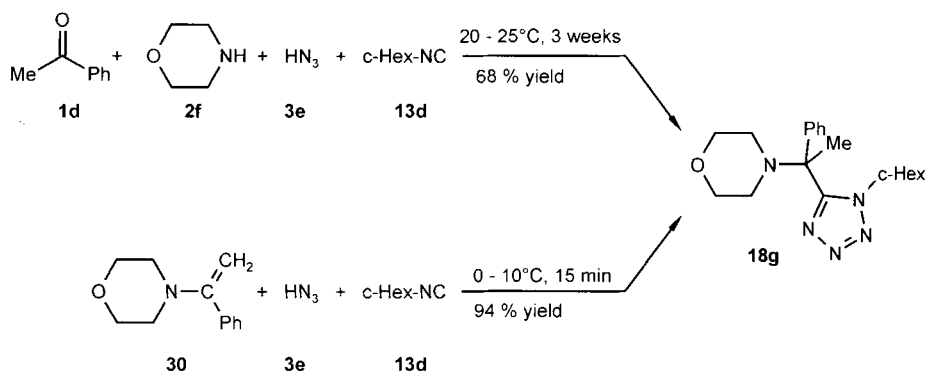
Figure 4.2.

According to Scheme 4.13, the educts **1d**, **2f**, **3e** and **13d** of the U-4CR proceed rather slowly and form a low yield of **18c** with some autoxidation by-products. In contrast, the precondensed compound **30** of **1d** and **2f** reacts very rapidly with **3e** and **13d**, to obtain an almost quantitative yield of **18g** [73].

The chemical compounds of the four functional groups of the connected pairs of atoms C, H, O, and N react especially well if those of the pairs CO, CN, NH and a variety of acidic compounds HX are present. This concept has important roles in organic and bioorganic chemistry. If these four atom pairs are represented by the four corners of a tetrahedron (Fig. 4.3), then each of the six connected lines of such a tetrahedron corresponds to a variety of chemical reactions of two components, or the four triangles of such a tetrahedron



Scheme 4.12.



Scheme 4.13.

correspond to their 3CRs. The whole tetrahedron of these four pairs of functional groups represents the chemistry of the U-4CR. This corresponds to a union of these pairs of atom reactions. This feature of chemistry was proposed by Ugi, and investigated and confirmed by Heck [81].

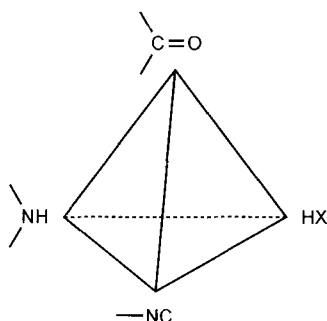


Figure 4.3.

## 4.10 Usual and Unusual Peptide Chemistry

Twenty-one  $\alpha$ -amino acids are contained in peptides or proteins, and can be prepared from these sources. Certain  $\alpha$ -amino acids can be produced by fermentations using suitable microbes, while the chemical synthesis of unnatural  $\alpha$ -amino acids may require many preparative steps [31, 82, 83].

The classical syntheses of peptide derivatives from easily obtainable  $\alpha$ -amino acids correspond to converting these into suitably protected  $\alpha$ -amino acid derivatives. They are connected and deprotected, and ultimately, after many coupling/deprotection cycles, the planned peptide derivatives are formed. Preparative difficulties increase with the number of cycles, with one of the main problems being the significant degrees of racemization that occur during some of these reactions. Consequently, chemical procedures must be carefully selected and carried out under suitable conditions [31, 72, 82–84].

If  $N$   $\alpha$ -amino acids are converted into the corresponding peptides, at least  $3N - 1$  steps of preparative procedures must occur in order to protect, connect and deprotect the functional groups of the  $\alpha$ -amino acids.

In 1966, Bodanszky and Ondetti [82] published their textbook, *Peptide Synthesis*, which contains the essential concepts and methods of peptide chemistry. Since that time, no further profound preparative progress has been accomplished in peptide chemistry, though many previously developed chemical methods have been improved and modified.

In 1963, Merrifield introduced the last profound new methodology of preparing peptides from  $\alpha$ -amino acids [85], and today his technique of connecting  $\alpha$ -amino acid derivatives by solid-phase procedures is the preferred means for the chemical preparation of peptides.

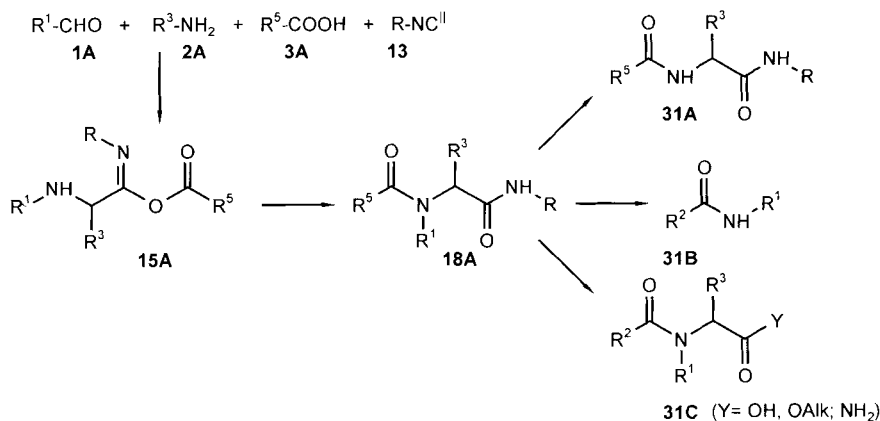
Two decades later, Furka [24, 86] introduced the production of solid-phase peptide libraries, forming these by extending Merrifield's procedure. Such peptide libraries were mainly used as an improved method of searching for new pharmaceutically applicable peptide derivatives. It was gradually realized, that such compounds have limited variability and that, in general, peptides are often at disadvantages when administered orally. Consequently, it was also realized that widely differing libraries of other types of chemical compounds can be formed by solid-phase multistep syntheses [87].

Rather than the usual multistep procedures, peptide derivatives may be prepared by one-pot U-4CRs. Bodanszky and Ondetti [82] first recognized the potential preparative advantages of this approach more than 30 years ago, suggesting that peptide derivatives might be prepared much more easily in this manner than by any other method [31].

With regard to the chemistry of the  $\alpha$ -amino acids and peptides, there are various ways of using the U-4CR which might be advantageous. However, in all cases special methods of cleaving the U-4CR products are required, and consequently special types of U-4CR components must always participate. This led to the insight [72, 73] that in its early phase, the main problem is to find a suitable primary amine component **2A** whose stereoselective U-4CR forms products **18A** that are cleavable under such mild conditions that the desired product **28A** is not racemized.

A major problem here is that a highly stereoselective (or even better a stereospecific U-4CR) must be accomplished under specifically controlled reaction conditions. According to preliminary studies, only suitable chiral amine components **2A** can influence the

steric aspects of the U-4CRs sufficiently well. It was also found that, depending on the reaction conditions, the one or the other diastereomeric product can be formed preferentially. Furthermore, the amine component **2A** of these components must contain an alkyl group  $R^3$ , whose product **18A** can be cleaved to form the product **31A**. Such a process would be particularly advantageous if, simultaneously with this cleavage, **18A**, **31A** and the amine **2A** can be re-synthesized. Therefore, preparing unnatural  $\alpha$ -amino acids and their peptides by stereoselective U-4CRs requires suitable primary amine components **2A**. Special amine components or isocyanides can also form U-4CR products **18A**, which can be converted into **31B** and **31C**.



**Scheme 4.14.**

Figure 4.4 shows some examples of one-pot U-4CRs elucidating the variety of products **18h–18m** possible by U-MCR. Naturally occurring 14-membered cyclic peptide alkaloids **18h** were prepared by Joullié et al. [88] using the U-4CR as a central step. Other natural products synthesized by U-MCR are bicyclomycin, dimethyldysidenin, willardiin, nikkomycin, sinefugin, polyoxins, plumbemycin, glutathione and eloidisine. Glycosamine can be synthesized by the U-4CR and subsequent reductions. Libraries of C-glycoside peptide ligands for cell surface carbohydrate receptors **18i** using rink resin have been described. Under the quasi-high-dilution conditions of the resin, dicarboxylic acids can be mono-converted in the U-4CR. An amine *ras-raf* element directed library **18j** has been created using amino acids bond onto polyethylene glycol. The yields when using the PEG-method could not be improved over those of the liquid-phase reaction in methanol, but the resulting products could be purified more easily. Anti-virus-directed nucleobase libraries **18k** have been described, while a powerful combination of the genetic algorithm and the U-4CR has been introduced by Weber et al. [21] for the rapid discovery of new leads. In a milestone publication, Weber et. al. described a combinatorial approach for the discovery of novel products like **18l** by such MCRs. This is a highly interesting continuation of the application of combinatorial chemistry to related areas such as drug discovery, and multiple drug-resistant reverting compounds **18m** could be found by applying combinatorial U-4CR [22].

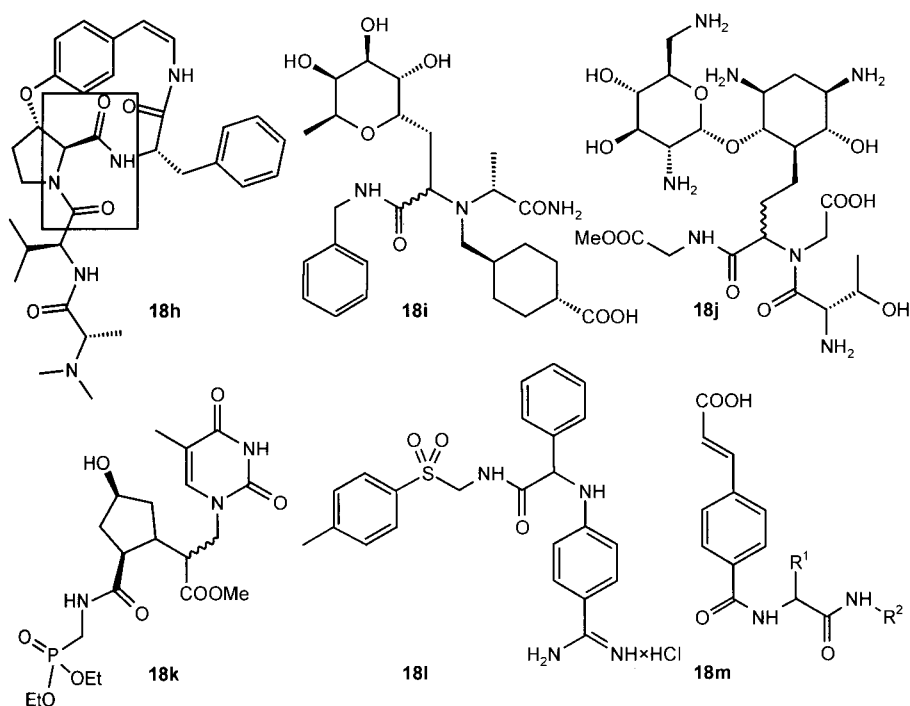
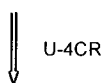
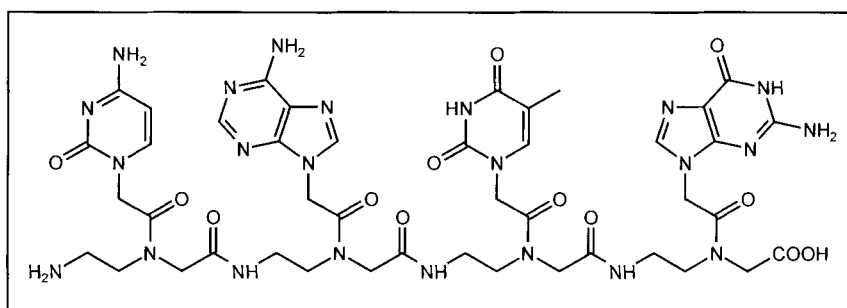


Figure 4.4.

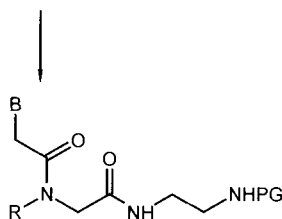
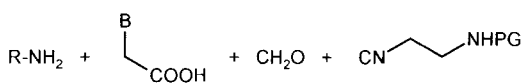
In 1982, the first syntheses of DNA/RNA-related peptide nucleic acids (PNAs) were described by Shvachkin et al. [89], and ten years later Nielsen et al. [90] represented a novel PNA backbone with an enhanced binding to the complementary DNA and RNA strands. This has many potential uses [91]. The very good protease and nuclease stability makes this PNA promising in antisense therapy. Unfortunately, some properties, including the tendency of self-aggregation, poor water solubility and an inability to cross the cell membrane, hamper the in-vivo application of PNAs. Therefore, many scientists have tried to enhance these properties by introducing modifications in the backbone. The usual method of preparing such compounds corresponds to the coupling of monomers by conventional multistep peptide chemistry.

The PNA syntheses can also be accomplished by the U-4CRs according to Scheme 4.15, including deprotected procedures **18A**  $\rightarrow$  **29C**. Thus, not only such PNAs but also their libraries can be accomplished much more quickly and easily [92].

In contrast to standard methods of PNA synthesis, this approach allows the introduction of much more diversity, and eventually this should improve the properties of PNAs for antisense applications.

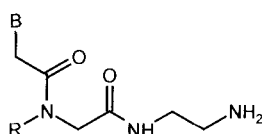


1 step



remove  
protecting  
group

2 step



3 step

use resulting primary amine as amine component  
in step 1, repeat step 1 - 3 n times

**Scheme 4.15.**

## 4.11 Stereoselective U-4CRs and their Secondary Reactions

As a result of early pre-studies of the U-4CR, it was realized that the yield and stereoselectivity of this reaction depend greatly on the ratio of educts and on the reaction conditions [8, 93]. The individual goals of these syntheses can differ widely, especially if certain stereoisomers of products are formed preferentially [3, 11b, 73]. Planning of syntheses by the U-4CR and related MCRs, and the suitable choice and availability of its educts, is even more important than in most other chemical reactions, particularly as these MCRs do not always function under optimal conditions, but may differ from case to case.

The reaction mechanisms of the MCRs cannot be investigated with the usual methods. Some essential mechanistic information of a MCR can be obtained from a suitable combination of several methods, and not all of the usual detailed information can be expected. It is even more difficult to identify optimal reaction conditions to create certain isomers stereoselectively by U-4CRs. Nevertheless, studies of the reaction mechanism of the U-4CR can be accomplished more easily if a stereoselective U-4CR takes place, than without the stereochemical differences of its different products. Therefore, a particularly suitable model reaction of a stereoselective U-4CR was investigated.

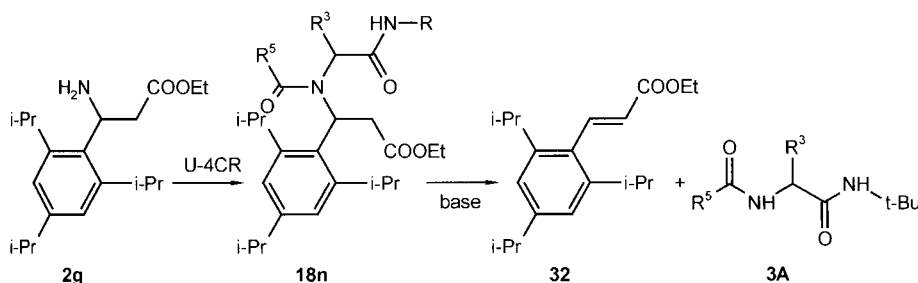
The number of essential steps was reduced by precondensing its aldehyde **1A** and its amine component **2A** into its Schiff base. Thus, the U-4CR can also proceed sufficiently well if somewhat low concentrations of educts are used.

Between 1964 and 1967, many different ratios and concentrations of isobutyraldehyde-(S)- $\alpha$ -phenylethylamine, benzoic acid and tert-butylisocyanide in methanol at 0 °C were investigated [4, 8, 28]. In one series of experiments the dependence of the electrical conductivity of the Schiff base and its carboxylic acid was determined, while in a second series the relationship between educt concentration and the ratio of diastereoisomeric products was investigated by measurement of their optical rotation.

The large quantity of data obtained was evaluated using a system of mathematical equations that were solved using a computer program of the reaction mechanism [29, 30]. Thus, not only were the rate constants of the essential steps of this specific U-4CR mechanism determined, but also the general concept of such reaction mechanisms. These findings showed that this U-4CR started with an equilibrium of a 3CR of type IA, followed by an irreversible  $\alpha$ -addition onto its isocyanide. The U-4CR belonged to type IIB, though only in recent years has the general importance of the different types of MCRs been fully realized [4].

In the early 1960s, in addition to the early finding of stereoselectively reacting amine components, several new types of achiral amines were also found whose U-4CR products could be cleaved selectively according to **18**  $\rightarrow$  **28A** [73]. However, even today, despite an ongoing search for optimal amine components, the simultaneous fulfillment of all conditions has not been accomplished. Indeed, it transpires that this problem remains one of the most difficult in preparative chemistry.

The first amine component **2g** that had all necessary properties. Such U-4CRs can form (proportionally) up to 95:5 cations of their diastereomeric products **18n**, and the unde-



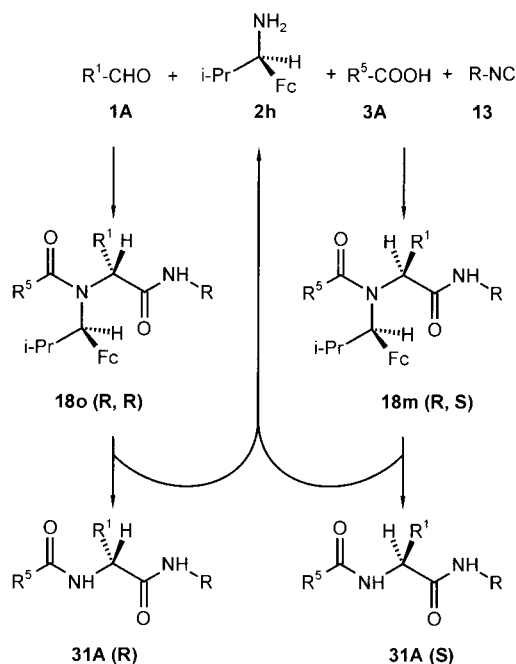
Scheme 4.16.



sired isomer can selectively be destroyed [73]. In the past, the synthesis of  $\alpha$ -amino acid derivatives using methods of the U-4CR has involved many preparative disadvantages, though nowadays the selective removal of minor proportions of unrequired isomers and related compounds is an accepted general method in preparative chemistry [32].

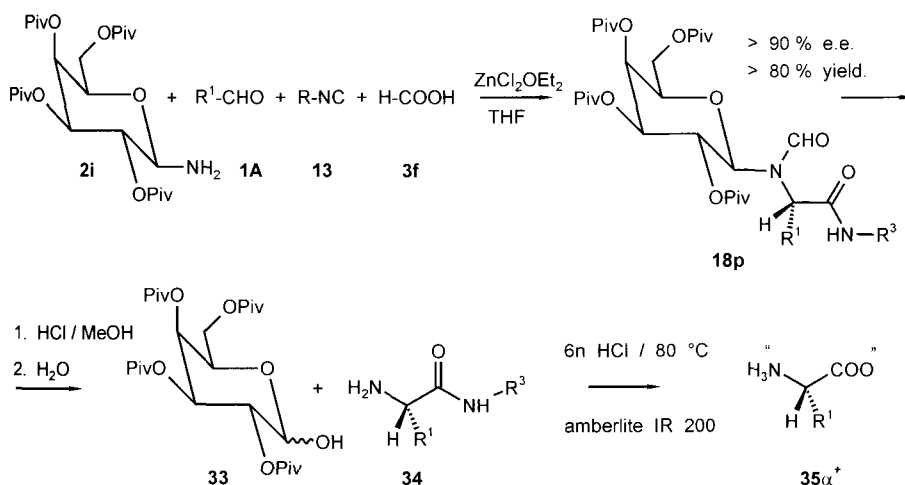
In principle, the synthesis of peptide derivatives by the U-4CR in the presence of the  $\alpha$ -ferrocenyl alkyl amines is a perfect way of accomplishing such syntheses. However, in practice the optimal ferrocene derivatives and their reaction conditions have not yet been found, and ongoing attempts seek to find the ideal way of applying such chemical technique.

The initial problem was in the attempt to prepare a variety of correct intermediate and final derivatives of ferrocene. These compounds first became available in the 1970s [73, 94], when the  $\alpha$ -ferrocenyl alkyl amines were shown also to play an essential role, and when the He-3CR of ferrocene was first introduced [94]. These derivatives are now widely used as catalysts in organic chemistry [95], as illustrated by the U-4CR of **2h** being the first chiral amine component that fulfilled all essential aspects (Scheme 4.17; see also Scheme 4.14); however, the individual steps of such U-4CRs did not yet proceed in sufficient yield and stereoselectivity. The final yields of cleaving **18o** into **31A** and the simultaneous re-synthesis of the amine **2h** were poor, though hope remains (with some justification) that closely related U-4CRs with  $\alpha$ -ferrocenyl- $\beta$ -hydroxyethylamine derivatives will be developed.



**Scheme 4.17.**

Kunz and Pfrengle [96] introduced the formation of  $\alpha$ -amino-acid derivatives by U-4CR with chiral O-acylated amino-carbohydrates like **2i** and formic acid **3f**. Usually, their products are formed relatively stereoselectively, and generally in good yield. It seems that the use of formic acid **3f** proceeds in the U-4CRs better than do other acid components of **3A**, as its amine components like **2i** are sterically hindered. The essential disadvantage of the products **18p** is that subsequent cleavage of its C-N bond into **33** and **34** can be accomplished only by strong acids such as HCl. Thus, it is doubtful if delicate chiral  $\alpha$ -amino acid derivatives or peptides can be prepared successfully in this manner.



Scheme 4.18.

Later, several other types of 1-amino-carbohydrate derivatives like **2j** and **2k** were used, in order to obtain products of peptide derivatives **18A** by the U-4CR. These could be cleaved more efficiently, but not yet well enough [97].

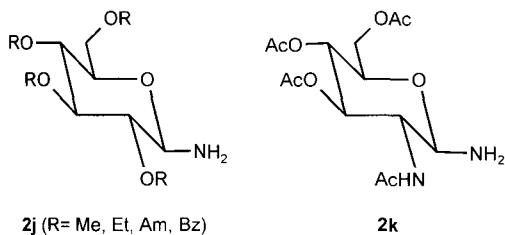
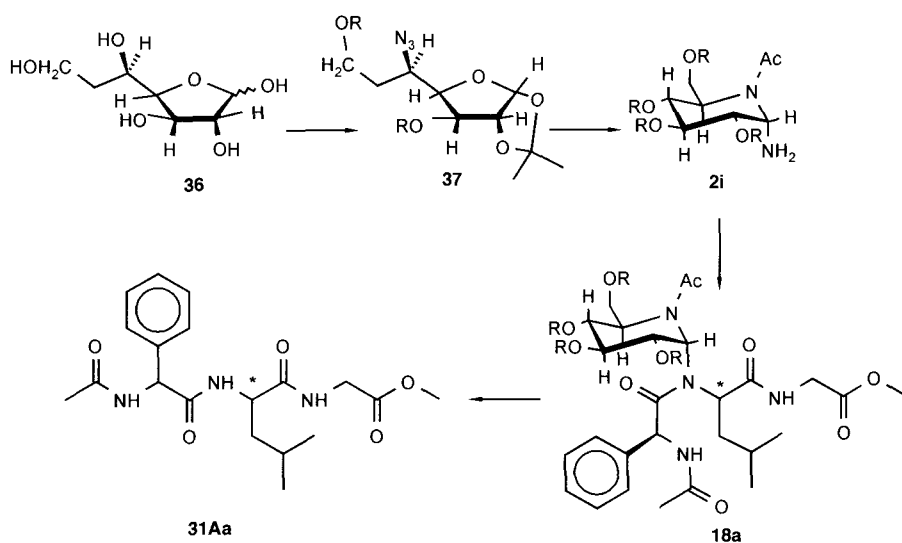


Figure 4.5.

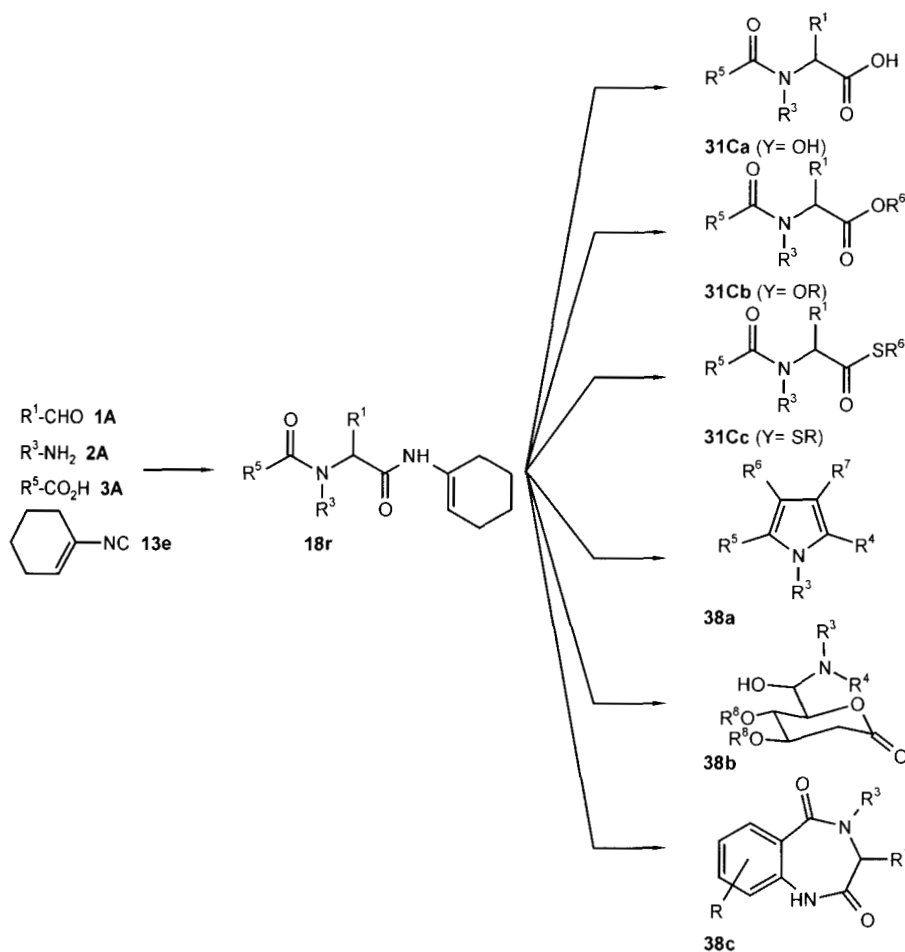
Recently, some other amino carbohydrates (e.g. **2l**) were introduced whose endocyclic oxygen was replaced by nitrogen derivatives like **36**  $\rightarrow$  **2l**. It appears that all such  $\alpha$ -amine components of peptide syntheses by U-4CRs have desirable properties [75].



Scheme 4.19.

Occasionally, it is advantageous to form a peptide derivative by a U-4CR so that N- and C-protected  $\alpha$ -amino acids are used as **2A** and **3A** with a new connected C-N bond of the other functional groups of their product **18A**. The component **1A** of such a U-4CR must be selected so that **18A** can be converted into **31B** without any cleavage or stereoisomerization, and **31B** can be obtained with ease as a pure product.

Occasionally it is also suitable to have an isocyanide **13e** whose product **18n** contains a -CO-NH-R group that can be converted into a derivative **31C** with the group -CO-Y, which can correspond to a carboxylic acid (Y = OH), a carboxylic ester (Y = OR with an alkyl group), or an amide (Y = NH<sub>2</sub>) (Scheme 4.20).



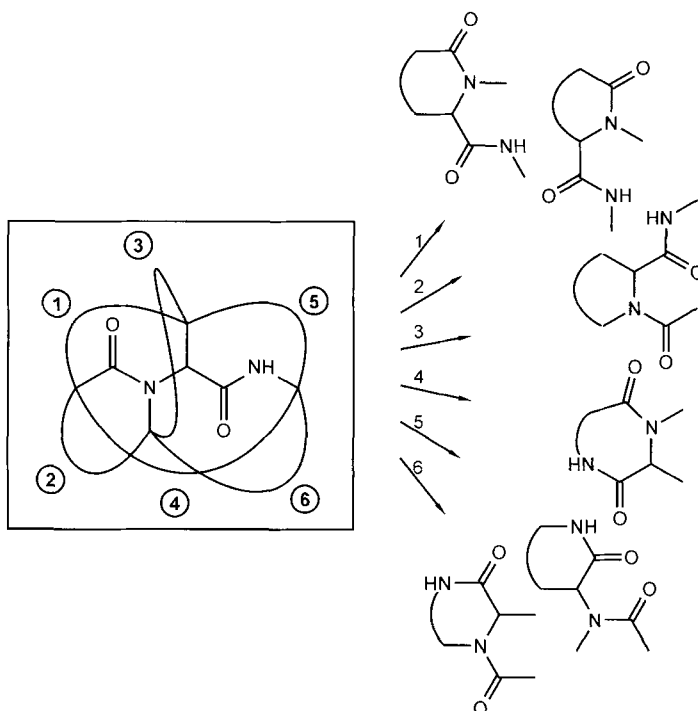
Scheme 4.20.

Armstrong et al. [98] formed various products of the type **31C** by hydrolyzing under mild conditions the products of the U-4CRs with  $\Delta^1$ -cyclohexenyl isocyanide **13d** [99] as a particularly suitable educt. Many different products of the types **31Ca–c** and the heterocycles **38a–c** can be obtained from **18A**. The first such reactions were carried out in 1963. The pharmaceutically interesting products **38c** can also be formed directly by a U-4CR of anthranilic acid [100].

## 4.12 MCRs of Educts with Two or Three Functional Groups

A U-4CR converts its four characteristic educts into its product, so that each functional group of the participating educt plays a particular role. However, a great variety

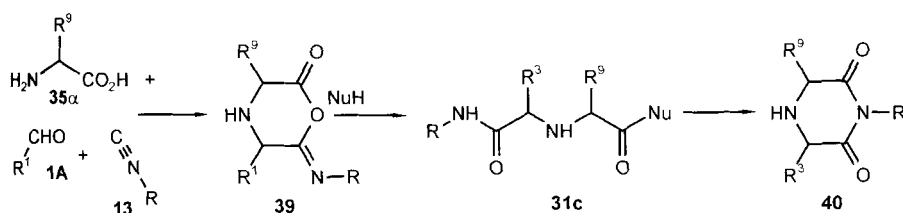
of somewhat different but also closely related reactions can take place, if two or three of their participating reactive groups belong to the same educt. As a consequence, six ( $6 = 4! : 2! \times 2!$ ) pairs of them can form a minimum of six structurally different types of products; thus a much larger variety of intramolecular products can be formed [6].



Scheme 4.21.

The bifunctional educts can form five- or six-membered cyclic  $\alpha$ -adducts of the isocyanides, which either rearrange into their final heterocyclic products, or these cyclic intermediate  $\alpha$ -adducts react with further nucleophilic reagents or anions of weak acids **NuH**.

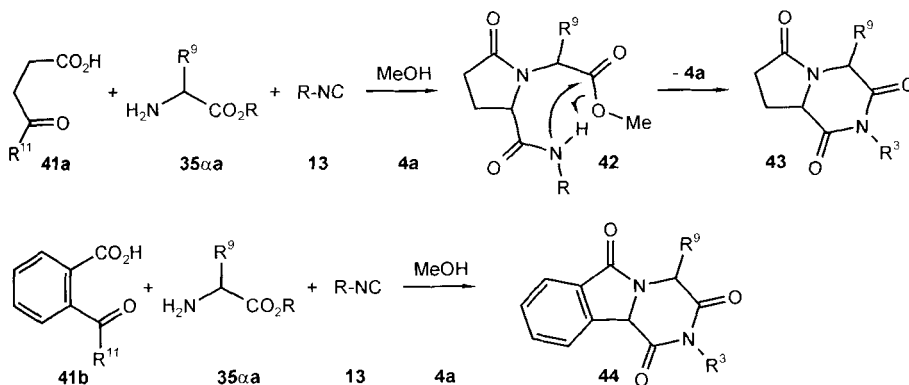
However the  $\alpha$ -amino acids **35a** reacts with aldehydes **1A** and isocyanides **13** in a rather different way than the other amino acids, since the six-membered cyclic intermediates **38** do not directly rearrange into their final products, but instead **38** reacts with some additional nucleophilic compounds. If methanol is used as solvent, this can be its fifth educt of this 5CR [3, 4, 100a]. At a higher temperature, their methoxy-products **31C** ( $3g = 4a$ ) form 2,6-piperazinediones **40** [101a] (Scheme 4.22).



Scheme 4.22.

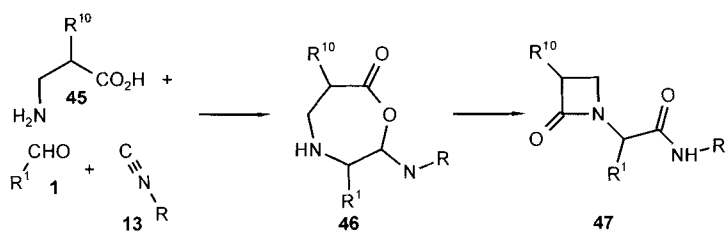
In the 1960s, a somewhat stable five-membered cyclic  $\alpha$ -adduct could also be formed by an isocyano-alkylcarboxylic and a Schiff base. At  $>80^\circ\text{C}$  this  $\alpha$ -adduct can be rearranged into a 2,5-piperazinedione derivative. At lower temperatures, a peptide derivative can be formed from this  $\alpha$ -adduct, in addition to a secondary amine [8, 102].

In recent years a wide variety of U-4CRs [103, 104] educts of the carbonyl and carboxylic groups have been studied. Besides a few U-4CRs of  $\alpha$ - and  $\beta$ -oxoacids [101b], many  $\gamma$ -oxoacids were thus converted into their  $\gamma$ -lactam derivatives, like **43** and **44** [102] (Scheme 4.23).



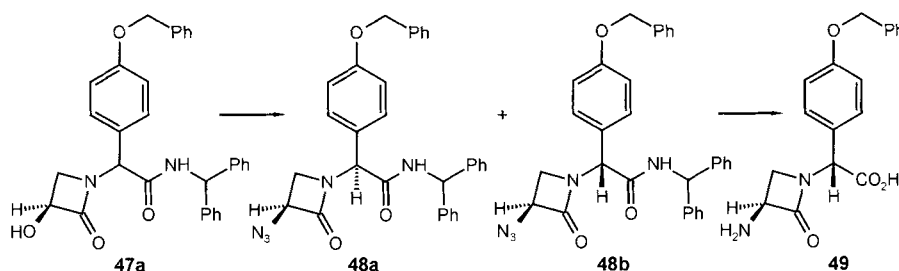
Scheme 4.23.

Thus,  $\beta$ -lactam derivatives **47** are formed by the U-4CR from  $\beta$ -amino acids **45** [8, 105–108] (Scheme 4.24). However, the anthranilic acid reacts with aldehydes and isocyanides so that two molecules of each educt are formed, forming eight-membered cyclic products [100b].



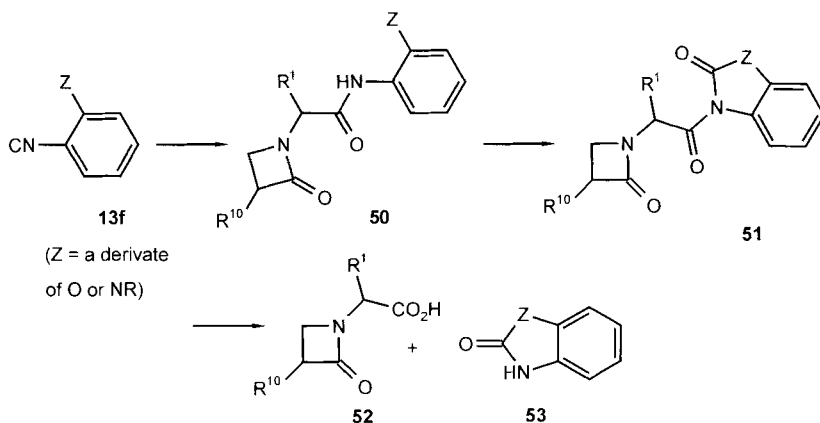
Scheme 4.24.

Shortly after this it was also realized that it is sometimes of interest to convert the U-4CR products into other chemical compounds. In 1983, Isenring and Hofheinz [109] of Hoffmann-La Roche produced a  $\beta$ -lactam derivative **47** from the diphenylmethyl isocyanide, which could be converted into its derivative **49** [108, 109] (Scheme 4.25).



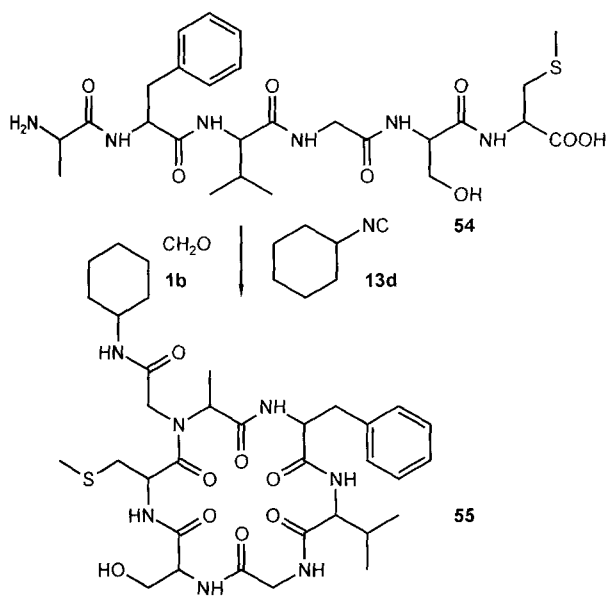
Scheme 4.25.

The U-4CR of the isocyanide **13f** form the products **49** ( $Z = NH, O,$  or  $S$ ), which under very mild conditions are convertible into their acylating derivatives **51**, which can also be hydrolyzed under mild conditions, such that even  $\beta$ -lactam derivatives **52** can be prepared [110].



Scheme 4.26.

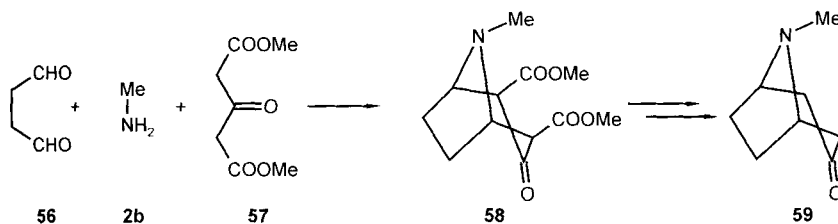
Faili et al. [73, 111] accomplished a U-4CR of a hexapeptide **54** which forms an 18-membered cyclic hexapeptide product **55** (Scheme 4.27). This macrocyclization was not further optimized and proceeded fairly well by the U-4CR.



Scheme 4.27.

### 4.13 Unions of MCRs and Related Reactions

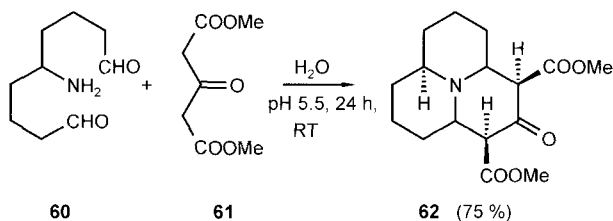
Several one-pot syntheses are closely related to MCRs that have irreversible final steps, although they are unions of two usually reversible M-3CRs, since their products are rather stable polyheterocyclic compounds. The first such procedure was introduced by Robinson in 1917 [48a], who synthesized the tropinone **59** from succindialdehyde **56**, methylamine **2b** and the methyl ester of acetonedicarboxylic acid **57** (Scheme 4.28). Two decades later, Schöpf et al. accomplished some progress with a closely related synthesis [48b].



Scheme 4.28.



The even more sophisticated stereospecific formation of the chiral alkaloid **63** by Stevens and Lee [49], and the previous synthesis of **58** and related procedures have in common their essential ring-closing steps corresponding to unions of M-3CRs. The  $\alpha$ -aminoalkylating MCRs of type I [4, 7] correspond to collections of equilibrating subreactions. Such positive formation of polyheterocyclic products correspond to almost irreversible cyclocondensations. These elegant syntheses correspond to an area of preparative chemistry whose products can be formed successfully, and where unions of reversible MCRs can have virtually irreversible final steps.



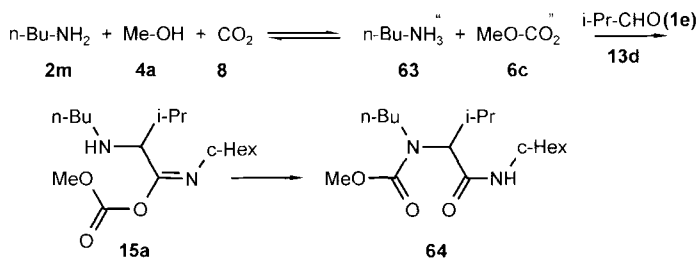
Scheme 4.29.

#### 4.14 MCRs of More than Four Functional Groups

In 1993 it was realized that one-pot MCRs of five and more educts can take place as unions of the U-4CR and other reactions or MCRs. A great variety of their educts and products can participate, if the irreversible U-4CR can be combined with a great variety of further types of participating chemical compounds.

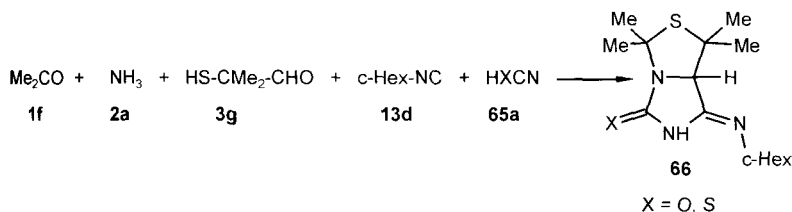
Usually, the higher MCRs are unions of the U-4CR and further chemical reactions or MCRs. However, these procedures are the rare exceptions. The majority of MCR unions proceeds only well if at least one of its participating MCRs belongs to type II. In these exceptional cases, an irreversible final step of MCR unions takes place, although all of the participating subreactions of MCRs are reversible.

When the U-4CR was introduced, some 5CRs were also introduced. Instead of the usual carboxylic acids **3A** of the U-4CR, the methoxycarboxylic acids were also used. Since the acid components of this reaction equilibrate with methanol **4a** and  $\text{CO}_2$  **8**, as well as with the ions **63** and **6c** [8, 13b, 112], five different educts can participate as the first unior of an U-4CR and the reaction of **4a** and  $\text{CO}_2$  **8**.



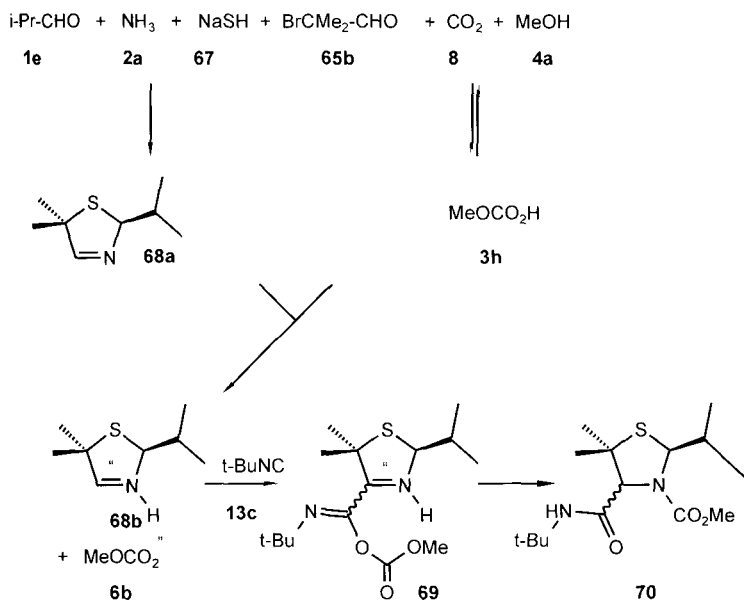
Scheme 4.30.

In the early studies relating to anions of acid components in the U-4CR, it was also found that cyanic acid and thiocyanic acid can participate [113, 114] in the U-4CR. It was then realized that products of the type **67** can also be formed directly via A-3(4)CRs [36, 37]. Such 5CRs and 6CRs [115, 116] correspond to the union of the A-MCRs and the U-4CR [2, 3].



Scheme 4.31.

The chemistry of even higher MCRs began when the concept of unions was fully realized [2], and when it was shown that higher unions of one-pot reactions can succeed if their last procedures are MCRs of type II. The first 7CRs were accomplished in 1993 [116, 117]. In Britain [118] and the USA [119], it was soon recognized that such MCRs were the start of a new era of MCR chemistry and their libraries. The first 7CR took place via the  $\alpha$ -addition of **69b** and **6b** onto **13c**, forming **70**. This corresponds to the isocyanide **13c  $\alpha$ -adduct **3h** of **4a** + **8** and the intermediate A-4CR product **68a** of **1d**, **2a**, **66b** and **68** rearranges into **70**, as its last event, just like an U-4CR (Scheme 4.32).**



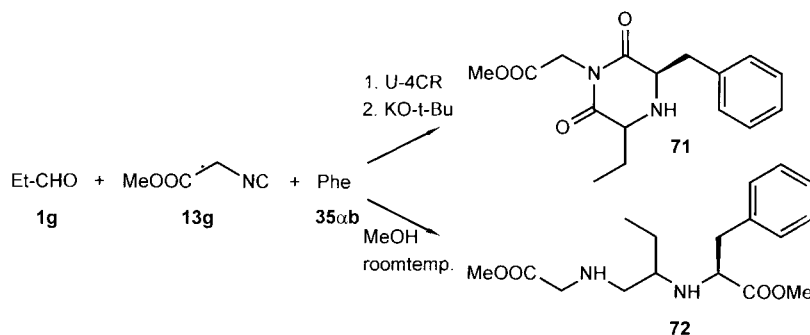
Scheme 4.32.

Until now, all MCRs of more than five different educts have been unions of MCRs, in whose last steps isocyanides have participated.

#### 4.15 New Unions of the U-4CR and Further Chemical Reaction

Recently, many new types of chemical products have been described that were produced by one-pot MCRs of by-functional educts [6]. Since then, further progress has been made. In particular, those MCRs will be mentioned here which are unions of the U-4CR and related chemical reactions, and that have been developed recently by several chemical companies.

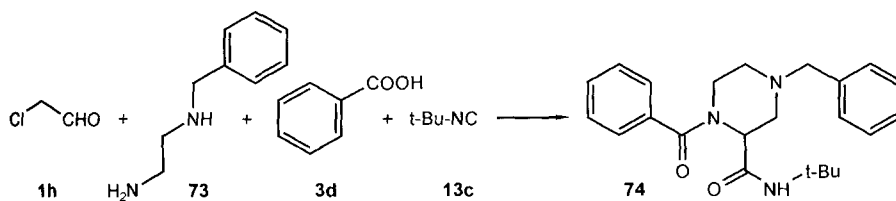
Depending on the reaction conditions,  $\alpha$ -iminodicarboxylic acid derivatives or 2,6-dipiperazinediones [120] are formed from  $\alpha$ -amino acids, oxocompounds and isocyanides. The isomeric 2,4-piperazinediones can thus also be formed by some of the U-4CR of vicinal diamines, isocyanides, oxocompounds and glyoxylic monoesters [121, 122]. The libraries of highly substituted 2-piperazones were produced by the chemists of Merck Inc. from N-monosubstituted ethylenediamines, carboxylic acids, isocyanides and chloroacetaldehyde.



**Scheme 4.33.**

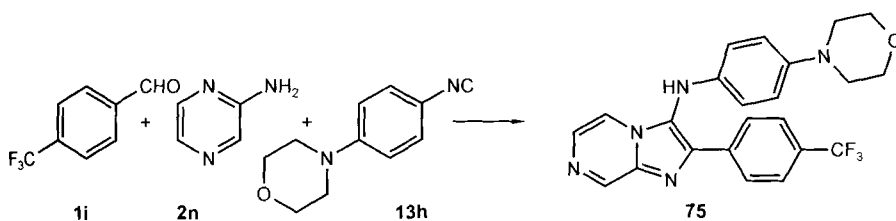
Chloroacetaldehyde combined with  $\alpha$ -amino acids and isocyanides, then arizidine derivatives can thus be formed [123].

Instead of the U-4CR, a 3CR takes the place of its secondary reaction when 2-amino-pyridine or pyrazine, oxocompounds and isocyanides form their products, so that its carbonic acid components do not enter this product [124, 125]. In all probability, the intermediate protonated Schiff base reacts with its isocyanides so that, ultimately, the bicyclic 3-aminoimidazole [1,2a] pyridine is cyclized. Some further diversifications can proceed, if secondary amines can be acylized. No other simple way of producing a comparable collection of such a great variety of complex heterocycles can be accomplished [126].

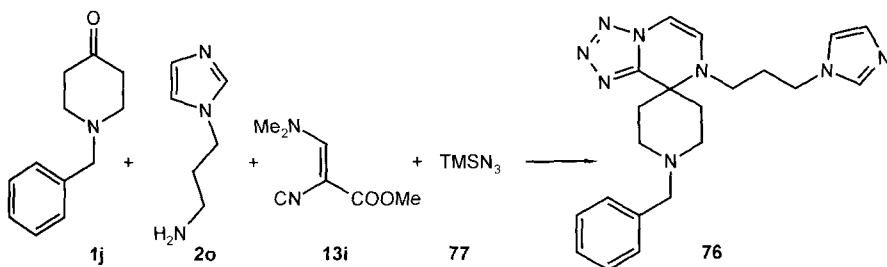


Scheme 4.34.

The early study of the U-4CR introduced the preparation of monosubstituted tetrazole derivatives from the isocyanides and hydrogen azide, and since the introduction of the U-4CR a great variety of 1,5-disubstituted tetrazoles can thus be formed. A research group of Rhône-Poulenc introduced an interesting version of the U-4CR of producing intermediate products from Schöllkopf's alkyl- $\beta$ -(N,N-dimethyl amino)- $\alpha$ -isocynoacrylate which subsequently forms a bicyclic piperazine derivative [127] by replacing it dimethylamino group. This method resulted in high-purity and high-yielding products.

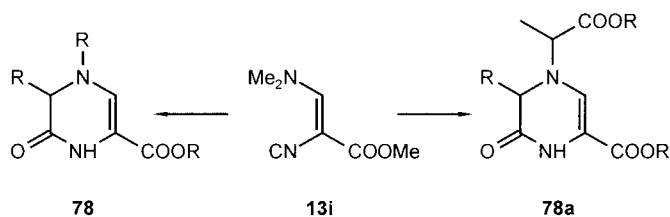


Scheme 4.35.



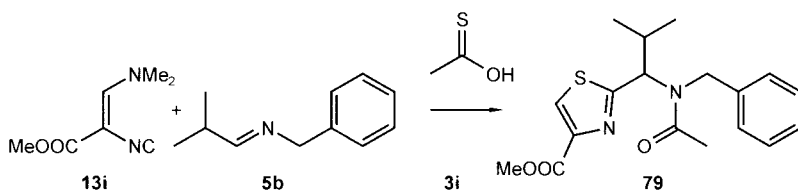
Scheme 4.36.

Similar reactions of  $\alpha$ -amino acids, as well those of  $\alpha$ -amino amines are conceivable [128].



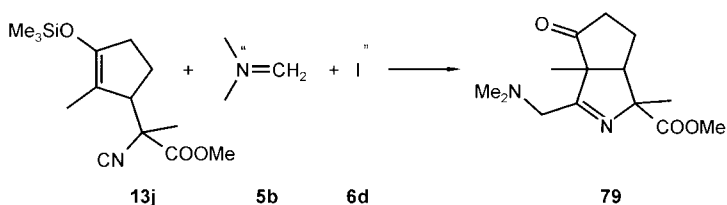
Scheme 4.37.

Recently, we found that thioacids work in Ugi's MCR forming monothioabisamides. Together with Schöllkopf's isocyanide, this resulted in a new thiazole synthesis. For that particular substitution pattern of thiazoles it seems to be the most versatile synthesis so far known [129]. First experiments show that oxazoles and imidazoles are also accessible [128].



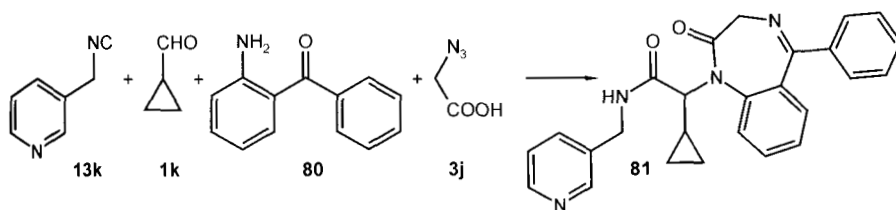
Scheme 4.38.

A Japanese group described the first U-4CR-related C-nucleophilic reaction [130]. The 4-isocyano-/silyloxy-alkenes react with Eschenmoser's salt into their bicyclic pyrroline derivatives.



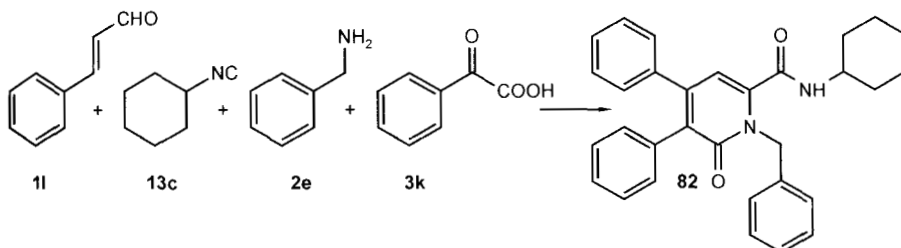
Scheme 4.39.

In 1995, Keating and Armstrong reported their combinatorial synthesis of benzodiazepines by their 'convertible' isocyanides [131]. An anthranilic acid reacts with cyclohexeny isocyanide, and an oxo-component by a U-4CR. Under acidic conditions, a Münchnon intermediate results and subsequently a benzodiazepine is formed.



Scheme 4.40.

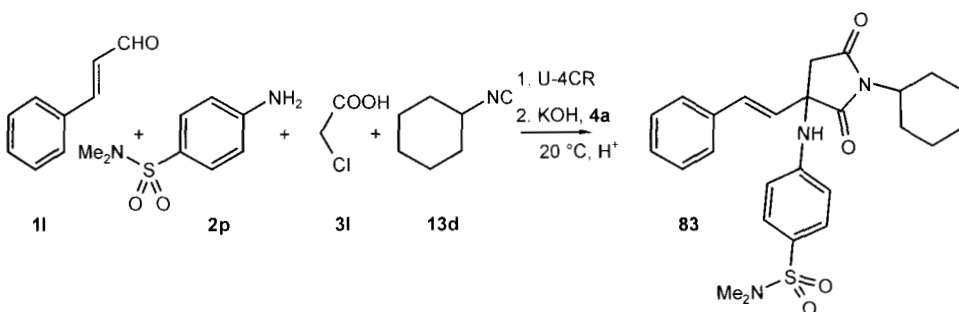
Recently, a research group of Hoffmann-La Roche described syntheses of benzodiazepine by a different kind of MCR [132], where 2-aminobenzophenone,  $\alpha$ -azido-carboxylic acid, oxocomponents and isocyanides react. The cyclization into benzodiazepines occurs via an Aza-Wittig reaction, this synthesis of benzodiazepines seeming to be far more variable and proceeding more easily than by the method of Armstrong.



Scheme 4.41.

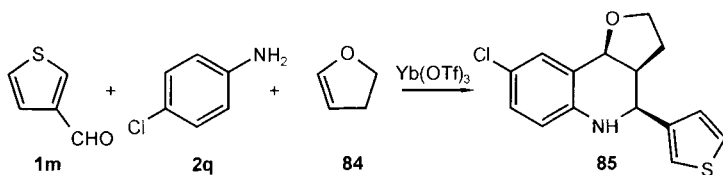
Natural product-related 1,6-dihydro-6-oxopyridine-2-carboxylic acid derivatives originate from the reaction of aldehyde, amines, phenylglyoxylic acid and isocyanides [133].

Highly substituted imides can also be produced starting from cinnamon aldehydes, primary amines, isocyanides and chloroacetic acid [134]. Their ring close procedures proceed in excellent yields and under mild conditions in an alkaline milieu.



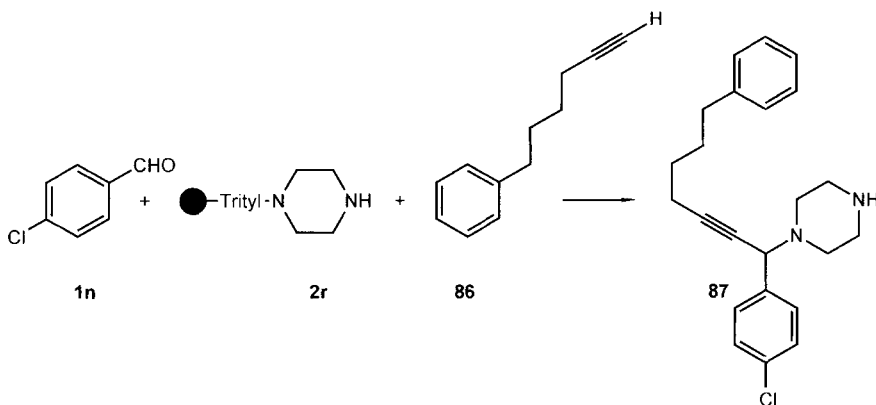
Scheme 4.42.

Many newly described syntheses of complicated products are based on isocyanide-based MCR, while others use totally different types of MCRs. A Japanese group reported a 3CRs of enoethers, aldehydes and amines under Lewis acid catalysis of lanthanoid derivatives. An  $\alpha$ -aminoketone results, and if an aniline derivative was used, then quinoline derivatives were formed [135].



**Scheme 4.43.**

A solid-phase variation of the M-3CR of acetylene, aldehydes and amines was recently described [136].



**Scheme 4.44.**

## 4.16 Progress in Combinatorial MCR Chemistry

In the combinatorial chemistry of the MCRs, recent progress has been achieved in two ways. At Hoffmann-La Roche, Weber et al. were able to prepare and analyze all products formed by automated combinatorial chemistry of seven different reactive educts. All possible mixtures of seven, six, five, etc. components were reacted and subsequently evaluated automatically by a HPLC-MS protocol. By abstraction of the starting materials signals [137], several new MCRs could be found in the complex reaction mixtures. Previously in 1961, and again in 1971, it was shown that  $M^4$  new constitutionally different products can be formed by the U-4CR, if  $M$  different educts of each class of their educts are formed [8].

Thus, the usual U-4CR can generate  $2 \cdot 10^{10}$  constitutionally different products, if 1000 commercially available aldehydes or ketones, amines, acid components and 20 isocyanides are brought to reaction. No combinatorial chemist would ever be able to synthesize so many chemical compounds. Nevertheless, searching for desirable products, such spaces of chemistry can be investigated. Weber's research group has demonstrated how large chemical spaces can be investigated by the means of genetic algorithms, although only a minor collection of chemical compounds must be produced and tested [23a]. Thus, it was possible to produce thrombin inhibitors in nanomolar activities. In a short time, a space of 160000 compounds could be investigated, and results confirmed by an independent experimental investigation [21]. Again, a genetic algorithm of a range of 15000 compounds and its target was studied. It could be confirmed that the same inhibitors resulted from this collection of chemicals, although two independent methods of searching were used.

## 4.17 Perspectives

Although the chemistry of the liquid-phase MCR libraries was introduced in the early 1960s, their profound preparative advantages were realized only recently. The solid-phase peptide libraries have been used since 1982, with various other solid-phase multistep libraries being produced subsequently.

In early 1995, the chemistry of the MCRs and its libraries suddenly became one of the most active areas of industrial research, despite only very few chemists being interested in this field before this time. This new, highly active area of chemistry was not introduced suddenly, but has been developed in a number of steps since the mid-1800s, the essential stages of progress being recognized as:

- the introduction of  $\alpha$ -aminoalkylating HO-3CRs of the type I B since 1838;
- the subsequent syntheses of heterocycles by 4CRs of the type IA;
- the beginning of isocyanide chemistry in 1859;
- the discovery of the P-3CR as the first 3CR of the type IIB of the isocyanides in 1921;
- the good availability of the isocyanides since 1958;
- the development of the U-4CR of type IIB as a union of the HO-3CR and the P-3CR;
- the determination of the reaction mechanism of the U-4CR, and the general insight that three basic types of MCRs existed;
- the proposal of the MCR libraries in 1961 and 1971;
- the development of automated equipment and methods of producing and investigating the chemistry of libraries since 1982;
- the publication of many new algebraic and group theoretical theorems as the basis of new types of chemistry-oriented computer programs as a part of progress in the MCR chemistry since 1973;
- the demonstration that, in principle, an unlimited number of different educts can form products by their MCRs since 1993; and
- the mathematically based new methodology of the MCRs and their libraries since 1994.



## 4.18 Acknowledgements

The authors are grateful to R. Lazarus for kindly checking the manuscript.

## 4.19 References

- [1] E. Hjelt, *Geschichte der Organischen Chemie*, Vieweg & Söhne, Braunschweig, 1916.
- [2] a) I. Ugi, A. Dömling, W. Hörl, *Endeavour*, **18**, 115 (1994); b) I. Ugi, A. Dömling, W. Hörl, *GIT Fachzeitschrift für das Laboratorium*, **38**, 430 (1994).
- [3] I. Ugi, *Proc. Estonian Acad. Sci. Chem.*, **44**, 237 (1995).
- [4] I. Ugi, *Journal für Praktische Chemie*, **339**, 499 (1997).
- [5] Z. Hippe, I. Ugi (eds.), *MultiComponent Reactions & Combinatorial Chemistry*; German-Polish Workshop, Rzeszów/Technical University, Munich, 1998.
- [6] A. Dömling, *Combinatorial Chemistry & High Throughput Screening*, **1**, 1 (1998).
- [7] H. Hellmann, G. Opitz,  *$\alpha$ -Aminoalkylierung*, Verlag Chemie, Weinheim, 1960;
- [8] I. Ugi, *Isonitrile Chemistry*, Academic Press, New York, 1971, p. 1–278.
- [9] M. Passerini, *Gazz. Chim. Ital.*, **51 II**, 126 (1921); **51 II**, 181; **56**, 826 (1926); M. Passerini, G. Ragni, **61**, 964 (1931).
- [10] I. Ugi, S. Lohberger, R. Karl, *Comprehensive Organic Synthesis: Selectivity for Synthetic Efficiency*, vol. 2, chap. 4.6, B. M. Trost, C. H. Heathcock (eds), Pergamon, Oxford 1991, p. 1083; a) p. 1090.
- [11] I. Ugi, R. Meyr, U. Fetzer, C. Steinbrückner, *Angew. Chem.*, **71**, 386 (1959); b) I. Ugi, *Angew. Chem.*, **74**, 9 (1962); *Angew. Chem. Int. Ed. Engl.*, **1**, 8 (1962).
- [12] S. Mac Lane, G. Birkhoff, *Algebra*, MacMillan Company, New York, 1967 [p. 3: Given sets of  $R$  and  $S$  have the intersection  $R \cap S$  with the common elements  $R$  and  $S$ . This means  $R \cap S = \{x \otimes R$  and  $x \otimes S\}$ , whereas a union  $R \cup S$  is  $R \cup S = \{x \otimes R$  or  $x \otimes S\}$ ].
- [13] I. Ugi, C. Steinbrückner, *Chem. Ber.*, **94**, a) 734, b) 2802 (1961).
- [14] I. Ugi, A. Dömling, B. Gruber, M. Heilingbrunner, C. Heiß, W. Hörl, *Software-Entwicklung in der Chemie*. R. Moll (ed.), Gesellschaft Deutscher Chemiker, Frankfurt, 1995, p. 113–128. (Tagungsband der GdCh-Workshop 16.–18.11.1994).
- [15] W. S. Johnson, *Angew. Chem.*, **88**, 33 (1976); *Angew. Chem. Int. Ed. Engl.*, **15**, 9 (1976).
- [16] *Chem. Rev.*, **96**, 1–600 (1996).
- [17] T. L. Ho, *Tandem Organic Chemistry*, J. Wiley & Sons, New York, 1992.
- [18] L. F. Tietze, U. Beifuss, *Angew. Chem.*, 105, 137 (1993); *Angew. Chem. Int. Ed. Engl.*, **32**, 131 (1993); L. F. Tietze, *Chem. Rev.*, **96**, 115 (1996).
- [19] G. H. Posner, *Chem. Rev.*, **86**, 831 (1986).
- [20] a) S. Waltbaum, C. Broger, K. Gubernator, *Angew. Chem.*, **107**, 2452 (1995); *Angew. Chem. Int. Ed. Engl.*, **34**, 2280 (1995); b) O. Lacke, L. Weber, *Chimia*, **50**, 445 (1996); c) L. Weber, *Current Opinion in Chemical Biology*, **2**, 381 (1998); K. Groebke, L. Weber, F. Mehlin, *Synlett*, **6**, 661 (1998).
- [21] L. Weber, Lecture at the ANALYTICA 98, conference at Munich, June 1998.
- [22] F. Balkenhohl, C. v. Buschen-Hünnefeld, A. Lansky, C. Zechel, *Angew. Chem.*, **108**, 3436 (1996), p. 3445; *Angew. Chem. Int. Ed. Engl.*, **35**, 2288 (1996), p. 2299.
- [23] I. Ugi, A. Dömling, B. Gruber, *Nachr. aus Chemie, Techn. und Lab.*, submitted.
- [24] J. W. Schopf, *Earth's Earliest Biosphere: Its Origin and Evolution*, Princeton Univ. Press, Princeton, N.J., 1983; S. L. Miller, *Science*, **117** (1953) p. 528; A. D. Keefe, S. L. Miller, G. McDonald, K. L. Bada, *Proc. Natl. Acad. Sci. USA*, **92**, 11904 (1994); M. P. Robertson, S. L. Miller, **268**, 702 (1995); J. Sutherland, J. N. Whitfield, *Tetrahedron*, **53**, 1149 (1997); a) G. Wächtershäuser, *Microbiol. Rev.*, **52**, 452 (1988); C. Huber, G. Wächtershäuser, *Science*, **276**, 245 (1997).
- [25] A. Laurent, C. F. Gerhardt, *Ann. Chem. et Physique*, **66**, 181 (1838); *Liebigs Ann. Chem.*, **28**, 265 (1838).
- [26] A. Strecker, *Ann. Chem.*, **75**, 27 (1850).
- [27] J. Brandt, C. Jochum, I. Ugi, P. Jochum, *Tetrahedron*, **33**, 1353 (1977).
- [28] I. Ugi, G. Kaufhold, *Liebigs Ann. Chem.*, 709, 11 (1967); I. Ugi, *Rec Chem. Progr.*, **30**, 289 (1969).

- [29] I. Ugi, M. Wochner, E. Fontain, J. Bauer, B. Gruber, R. Karl, *Chemical Similarity, Chemical Distance, and Computer-assisted Formalized Reasoning by Analogy*, Wiley & Sons, New York, 1990, p. 239; I. Ugi, J. Bauer, K. Bley, A. Dengler, A. Dietz, E. Fontain, B. Gruber, R. Herges, M. Knauer, K. Reitsam, N. Stein, *Angew. Chem.*, **105**, 210 (1993); *Angew. Chem. Int. Ed. Engl.*, **32**, 201 (1993).
- [30] H. Bergs, *Ger. Pat.* 566,094 (1929); *Chem. Abstr.*, **27**, 1001 (1933); Bucherer et al., *J. Prakt. Chem.*, **140**, 69, 151, 129, 291 (1934); **141**, 5 (1934).
- [31] E. Wunsch, *Methoden der Organischen Chemie. Synthese von Peptiden*, Vol. I and II, Georg Thieme Verlag, Stuttgart, 1974.
- [32] C. Mannich, I. Krötsche, *Arch. Pharm.*, **250**, 647 (1912); F. F. Blick, *Organic Reaction*, R. Adams, (ed.), Vol. 1, John Wiley & Sons, New York, 1942, p. 303.
- [33] F. Asinger, *Angew. Chem.*, **68**, 413 (1956); F. Asinger, M. Thiel, E. Pallas, *Liebigs Ann. Chem.*, **602**, 37 (1957); F. Asinger, M. Thiel, *Angew. Chem.*, **79**, 953 (1967); F. Asinger, W. Leuchtenberg, H. Offermanns, *Chem. Zeitung*, **94**, 6105 (1974); F. Asinger, K. H. Gluzek, *Monats. Chem.*, **114**, 47 (1983).
- [34] A. Dömling, I. Ugi, *Tetrahedron*, **49**, 9495 (1995).
- [35] S. Hünisch, W. Richter, I. Ugi, J. Chattopadhyaya, *Liebigs Ann. Chem.*, **269**, 269 (1994); N. Puri, S. Hünisch, C. Sund, I. Ugi, J. Chattopadhyaya, *Tetrahedron*, **51**, 2991 (1995); J. Chattopadhyaya, A. Dömling, K. Lorenz, I. Ugi, B. Werner, *Nucleosides & Nucleotides*, **16**, 843 (1997); publication submitted soon.
- [36] W. Lieke, *Justus Liebigs Ann. Chem.*, **112**, 316 (1859).
- [37] E. Meyer, *Justus Liebigs Ann. Chem.*, **67**, 147 (1866).
- [38] A. Gautier, *Justus Liebigs Ann. Chem.*, **142**, 289 (1867); *Ann. Chim. (Paris) [4]*, **17**, 103, 203 (1869).
- [39] A. W. Hofmann, *Ann. Chem.*, **144**, 114 (1967); *Chem. Ber.*, **3**, 653 (1870).
- [40] E. Oliveri-Mandala, B. Alagna, *Gazz. Chim. Ital.*, **40 II**, 441 (1910); Pergamon, Oxford, 1991, p. 1083; a) p. 1090.
- [41] R. H. Baker, L. E. Linn, *J. Amer. Chem. Soc.*, **70**, 3721 (1948); R. H. Baker, A. H. Schlesinger, *J. Amer. Soc.*, **73**, 699 (1951).
- [42] M. J. S. Dewar, *Chem. Soc.*, **67**, 1499 (1945); *Theory of Organic Chemistry*, Oxford Univ. Press (Clarendon), London and New York, 1949.
- [43] I. Ugi, R. Meyr, *Chem. Ber.*, **94**, 2229 (1961 und 1958).
- [44] I. U. Nef, *Justus Liebigs Ann. Chem.*, **270**, 335 (1972); *ibid.* **309**, 126 (1899).
- [45] A. Eschenmoser, H. Heusser, *Experientia* **9**, 357 (1953); A. Eschenmoser, L. Ruzicka, O. Jeger, D. Arigoni, *Helv. Chim. Acta*, **38**, 1890 (1955); see also: D. H. R. Barton, G. R. Moss, *Chem. Commun.*, **1966**, 261; E. J. Corey, W. E. Russey, R. R. Ortiz de Mortellano, *J. Amer. Chem. Soc.*, **88**, 4750 (1966).
- [46] C. H. Heathcock, *Angew. Chem.*, **104**, 6751 (1992); *Angew. Chem. Int. Ed. Engl.*, **31**, 665 (1992).
- [47] a) I. Ryu, N. Sonoda, D. P. Curran, *Chem. Rev.*, **96**, 177 (1996); see also b) S. Danishefsky, R. Volkman, *Tetrahedron Letter.*, **1973**, 2521.
- [48] a) R. Robinson, *J. Chem. Soc.*, **762** (1917) (II); b) C. Schöpf, G. Lehmann, W. Arnold, *Angew. Chem.*, **50**, 779 (1937).
- [49] R. V. Stevens, A. V. M. Lee, **101**, 7032 (1979).
- [50] J. Ferris, P. C. Joshi, E. H. Edelson, J. G. Lawless, *J. Mol. Evol.*, **11**, 293 (1978); J. Oro, A. P. Kimball, *Arch. Biochem. Biophys.*, **94**, 217 (1961); S. Drendard, J. Ferris, A. Eschenmoser, *Helv. Chem. Acta*, **93**, 1373 (1990).
- [51] W. Rothe, *Pharmazie*, **5**, 190 (1950).
- [52] H. Griesebach, H. Achenbach, *Z. Naturforsch.*, **B20**, 137 (1965).
- [53] I. Hagedorn, H. Tönjes, *Pharmazie*, **11**, 409 (1956); **12**, 567 (1957); I. Hagedorn, U. Eholzer, A. Lüttringhaus, *Chem. Ber.*, **93**, 1584 (1960);.
- [54] R. Arshadi, I. Ugi, *Polymer*, **31**, 1164 (1990).
- [55] W. R. Hertler, E. J. Corey, *J. Org. Chem.*, **23**, 1221 (1958).
- [56] I. Ugi, H. Perlinger, L. Behringer, *Chem. Ber.*, **91**, 2324 (1958); **92**, 1864 (1959); I. Ugi, Houben-Weyl, *Methoden der Organischen Chemie*, vol. E8d, Heterene III/part 4; E. Schaumann (ed.), Georg Thieme Verlag, Stuttgart, (1994) p. 796–803.
- [57] I. Ugi, R. Meyr, *Angew. Chem.*, **70**, 702 (1958); I. Ugi, R. Meyr, *Chem. Ber.*, **93**, 239 (1960); I. Ugi, U. Fetzter, U. Eholzer, H. Knapfer, K. Offermann, *Angew. Chem.*, **77**, 492 (1965); *Angew. Chem. Int. Ed. Engl.*, **4**, 452 (1965).
- [58] G. Skorna, I. Ugi, *Angew. Chem.*, **89**, 267 (1977); *Angew. Chem. Int. Ed. Engl.*, **16**, 259 (1977).
- [59] H. Eckert, B. Forster, *Angew. Chem.*, **99**, 922 (1987); *Angew. Chem. Int. Ed. Engl.*, **26**, 1221 (1987).

- [60] R. Obrecht, R. Herrmann, I. Ugi, *Synthesis*, 400 (1985).
- [61] e. g. I. Ugi, *Ger. Publ. Pat. Apl.* 1,158, 500, Farbenfabriken Bayer A. G. (1962–1963).
- [62] D. H. R. Barton, T. Bowles, S. Husinec, J. E. Forbes, A. Llobera, A. E. A. Porter, Z. Zard, *Tetrahedron Lett.*, **29**, 3343 (1988).
- [63] I. E. Baldwin, R. M. Aldinger, J. Chandragiamu, M. S. Edenbarough, J. W. Keeping, C. B. Ziegler, *J. Chem. Soc. Chem. Comm.*, 816–817 (1985).
- [64] Chang, C. W. J., Scheuer, P. J., Isocyanides in nature, *Top. Can. Chem.* **167**, 33–75 (1993); Scheuer, P. J., Isocyanides and cyanides as natural products, *Acc. Chem. Res.* **25**, 433–439 (1992)
- [65] Derwent World Drug Index, 1990–95.
- [66] K. Lindquist, S. Sundling, *Xylocain*, A. B. Astra, 1993; R. Dahlbon, *Det häanden på lullen*, (Astra 1940–1960), A. B. Astra, 1993.
- [67] G. Opitz, W. Merz, *Justus Liebigs Ann. Chem.*, **659**, 163 (1962).
- [68] K. Sjöberg, *Sv. Kem. Tidskr.*, **75**, 493 (1963).
- [69] J. W. McFarland, *J. Org. Chem.*, **28**, 2179 (1963).
- [70] G. Zinner, W. Kleigel, *Arch. Pharm. (Weinheim)*, **299**, 746 (1966).
- [71] S. Borman, *C&EN*, February, 24 (1997), p. 43; April, 6 (1998), p. 47
- [72] I. Ugi, *The Peptides*, Ed.: M. Meinhofer, E. Gross, Academic Press, New York, 1979, p. 365.
- [73] I. Ugi, D. Marquarding, R. Urban, *Chemistry and Biochemistry of Amino Acids, Peptides and Proteins*, Vol. 6, B. Weinstein (ed.), Marcel Dekker, New York 1982, p. 245.
- [74] G. George, *Doctoral Thesis*, T.U. München, 1974; A. Gieren, B. Dederer, G. George, D. Marquarding, I. Ugi, *Tetrahedron Lett.*, **18**, 1503 (1977); S. Lohberger, E. Fontain, I. Ugi, G. Müller, J. Lachmann, *New Journal of Chemistry*, **15**, 913 (1991).
- [75] A. v. Zychlinsky, *Doctoral Thesis*, Technical University of München, 1998.
- [76] T. Yamada, Y. Nakamura, T. Miyazawa, S. Kuwata, K. Mazumoto, *Japan Chem. Express*, **8**, 161 (1993).
- [77] T. Yamada, N. Motoyama, T. Taniguchi, T. Kazuta, T. Miuzawa, S. Kuwata, K. Matsumoto, M. Sugiura, *Chem. Lett.*, 723 (1987).
- [78] P. M. Hardy, I. N. Lingham, *Int. J. Protein Res.*, **21**, 406 (1983).
- [79] K. Matsumoto, T. Uchida, *Organic Synthesis at High Pressure*, K: Matsumoto, R. M. Acheson (eds.), Wiley, New York, 164 (1991).
- [80] H. Bock, I. Ugi, *Journal für Praktische Chemie*, **339**, 3859 (1997).
- [81] S. Heck, *Doctoral Thesis*, Technical University of München, 1998.
- [82] M. Bodanszky, M. A. Ondetti, *Peptide Synthesis*, G. A. Olah (ed.), J. Wiley & Sons, New York, 1966.
- [83] J. Jones, *Amino Acid and Peptide Syntheses*, Oxford University Press, Oxford, 1992.
- [84] G. Jung, *Angew. Chem.*, **103**, 1067 (1991); *Angew. Chem. Int. Ed. Engl.*, **30**, 1051 (1991).
- [85] R. B. Merrifield, *J. Am. Chem. Soc.*, **85**, 2149 (1963).
- [86] A. Furka, *Drug Dev. Res.*, **36**, 1 (1995).
- [87] K. Frobels, T. Krämer, *Chemie in unserer Zeit*, **30**, 270 (1996); N. K. Terrett, *Combinatorial Chemistry*, Oxford Univ. Press 1998; J. S. Früchtel, G. Jung, *Angew. Chem.*, **108**, 19 (1996); *Angew. Chem. Int. Ed. Engl.*, **35**, 17 (1996).
- [88] M. M. Joullié, *Synth. Comm.*, **20**, 459 (1990).
- [89] Y. P. Shvachkin, G. P. Mishin, G. A. Korshunova, *Russian Chem. Rev.*, **51**, 178 (1982); A. P. Smirnova, N. M. Ermak, *Z. Obsh. Khim.*, **65**, 1230 (1995).
- [90] P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt, *Science*, **254**, 1497 (1991).
- [91] L. Good, P. E. Nielsen, *Antisense Nucleic Acid Drug Res.*, **7**, 431, (1997).
- [92] A. Dömling, *Nucleosides & Nucleotides*, 1998, in press.
- [93] I. Ugi, K. Offermann, H. Herlinger, D. Marquarding, *Liebigs Ann. Chem.*, **709**, 1 (1967).
- [94] R. Herrmann, I. Ugi, *Tetrahedron*, **37**, 1001 (1981).
- [95] G. Wagner, R. Herrmann, *Ferrocenes*, A. Togni, T. Hayashi (eds.), VCH, Weinheim, 1995.
- [96] H. Kunz, W. Pfirngle, *J. Am. Chem. Soc.*, **110**, 651 (1988); *Tetrahedron*, **44**, 5487 (1988).
- [97] M. Goebel, I. Ugi, *Tetrahedron Letters*, **36**, 6043 (1995); S. Lehnhoff, M. Goebel, R. M. Karl, K. Klösel, I. Ugi, *Angew. Chemie*, **107**, 1208 (1995); *Angew. Chem. Int. Ed. Engl.*, **34**, 1104, (1995); M. Goebel, H.-G. Nothofer, G. Roß, I. Ugi, *Tetrahedron*, **53**, 3123 (1997).
- [98] R. W. Armstrong, *Combinatorial libraries related to natural products at the ACS National Meeting in Anaheim*, Calif. on April 2, 1995; T. A. Keating, R. W. Armstrong, *J. Am. Chem. Soc.*, **117**, 7842 (1995); **118**, 2574 (1996).
- [99] I. Ugi, F. K. Rosendahl, *Liebigs Ann. Chem.*, **666**, 65 (1963).

- [100] a) B. M. Ebert, I. Ugi, *Tetrahedron*, (1998) in press; b) B. M. Ebert, *Doctoral Thesis*, T. U. München, 1998.
- [101] a) A. Demharter, *Doktoral Thesis*, T. U. München, 1992; A. Demharter, W. Hörl, E. Herdweck, I. Ugi, *Angew. Chemie*, **108**, 185 (1996); *Angew. Chem. Int. Ed. Engl.*, **35**, 173 (1996); W. Hörl, *Doktoral Thesis*, T. U. München, 1996; b) I. Ugi, W. Hörl, T. Schmid, C. Hanusch-Kompa, E. Herdweck, *Heterocycles*, **47**, 965 (1998).
- [102] C. Hanusch-Kompa, *Doktoral Thesis*, T. U. München, 1998.
- [103] A. v. Zychlinski, I. Ugi, *Tetrahedron Letters*, submitted.
- [104] C. Hanusch-Kompa, I. Ugi, *Tetrahedron Letters*, **39**, 2725 (1998).
- [105] I. Ugi, *Angew. Chemie*, **94**, 826 (1982); *Angew. Chem. Int. Ed. Engl.*, **21**, 810 (1982).
- [106] E. Wischhöfer, *Chem. Ber.*, **95**, 136 (1962); A. Schutz, I. Ugi, H. F. Kabbe, *J. Chem. Res.*, **157**, 1979; A. Schutz, I. Ugi, *Z. Naturforsch.*, **34b**, 1303, 1979; G. Neyer, I. Ugi, *Synthesis*, **743**, 1991.
- [107] K. Kehagia, A. Dömling, I. Ugi, *Tetrahedron*, **51**, 139 (1995); K. Kehagia, I. Ugi, *Tetrahedron*, **51**, 9523 (1995); A. Dömling, M. Starnecker, I. Ugi, *Angew. Chemie*, **107**, 2352 (1995); *Angew. Chem. Int. Ed. Engl.*, **34**, 2238, (1995).
- [108] I. Ugi, H. Eckert, *Natural Product Chemistry*, A. ur Rahman (ed.), Vol. 12, Elsevier, Science Publ. 1000AE, Amsterdam, 1992, p. 113.
- [109] H. P. Isenring, W. Hofheinz, *Synthesis*, **37**, 385 (1981); **39**, 2591 (1983).
- [110] I. Ugi, J. Geller, *Chem. Scripta*, **22**, 85 (1983).
- [111] A. Failli, H. Immer, M. Götz, *Can. J. Chem.*, **57**, 3257 (1979).
- [112] H. Gross, J. Gloede, I. Keitel, D. Dunath, *J. Prakt. Chem.*, **37(4)**, 192 (1968).
- [113] I. Ugi, F. K. Rosendahl, *Liebigs Ann. Chem.*, **670**, 80 (1963).
- [114] I. Ugi, K. Offerman, *Chem. Ber.*, **97**, 2276 (1964).
- [115] K. Offerman, *Doktoral Thesis*, L. M. U. München, 1963; in 1994 S. Lehnhoff, PhD Thesis, Technical University Munich, 1994, C. Hanusch-Kompa, PhD Thesis, Technical University Munich, 1998.
- [116] A. Dömling, E. Herdweck, I. Ugi, *Acta Chemica Scandinavica*, **52**, 107 (1998).
- [117] A. Dömling, I. Ugi, *Angew. Chem.*, 105, 634 (1993); *Angew. Chem. Int. Ed. Engl.*, **32**, 563 (1993).
- [118] D. Bradley, *New Scientist*, 16, July 3 (1993).
- [119] *C&EN*, **32**, April 19 (1993).
- [120] I. Ugi, W. Hörl, C. Hanusch-Kompa, T. Schmid, E. Herdweck, *Heterocycles*, **47**, 965 (1998).
- [121] A. v. Zychlinsky, I. Ugi, I., *Tetrahedron Lett.*, (1998) in press.
- [122] K. Rosen, J. Sager, L. DiMichele, *Tetrahedron Lett.*, **38**, 3183 (1997).
- [123] T. Schmid, I. Ugi, *J. Prakt. Chem.*, 339 (1997).
- [124] K. Groebke, L. Weber, F. Mehlin, *Synlett*, 661 (1998).
- [125] C. Blackburn, B. Guan, P. Fleming, K. Shiosaki, S. Tsai, *Tetrahedron Lett.*, **39**, 3635 (1998).
- [126] U. Schöllkopf, P.-H. Porsch, H.-H. Lau, *Liebigs Ann. Chem.*, 1444 (1979).
- [127] H. Bienaymé, K. Bouzid, *Tetrahedron Lett.*, **39**, 2735 (1998).
- [128] A. Dömling, unpublished.
- [129] A. Dömling, S. Heck, E. Herdweck, *Angew. Chem.*, 1999, submitted for publication
- [130] M. Murakami, N. Hasegawa, T. Nagatomi, Y. Ito, *Heterocycles*, **47**, 85 (1998).
- [131] T. A. Keating, R. W. Armstrong, *J. Org. Chem.*, **61**, 8953 (1996).
- [132] A. Chucholowski, D. Heinrich, B. Mathys, C. Müller, *Abstr. Pap. Am. Chem. Soc.*, **214**, 287 (1997).
- [133] R. Bossio, C. F. Marcos, S. Marcaccini, R. Pepino, *Heterocycles*, **45**, 1589 (1997).
- [134] R. Bossio, C. F. Marcos, S. Marcaccini, R. Pepino, *Synthesis*, 1389 (1997).
- [135] S. Kobayashi, S. Komiyama, H. Ishitani, *Biotech. Bioeng. (Combinatorial Chemistry)*, **61**, 23 (1998).
- [136] J. J. McNally, M. A. Youngman, S. L. Dax, *Tetrahedron Lett.*, **39**, 967 (1998).
- [137] O. Lack, L. Weber, *Chimia*, **50**, 445.

This Page Intentionally Left Blank

# 5 Solid-Phase Anchors in Organic Chemistry

*Ralf Warrass*

## 5.1 Introduction

At quite an early point, it was realized that Merrifield's idea [1] for peptide chemistry on solid-phase supports had advantages for organic chemists other than just those interested in peptides [2–5], oligonucleotides [6–8] and oligosaccharides [9, 10]. Since the beginning of the 1970s, additional methods for organic chemistry on solid-phase supports (resins) have been actively sought. In the past few years, an explosion in research in this area has been aimed at establishing the entire spectrum of organic chemical reactions on solid-phase supports [11–13]. The basic incentive for this line of research has been the growing interest in pharmaceutical chemistry to develop new medical compounds more rapidly, and in a more cost-effective manner.

When one compares 'classical syntheses' with those done on solid-phase supports, a number of advantages become obvious. First, the ease with which many syntheses can be carried out and the possibility of their automation both result in a substantial saving in the time required; concomitantly, costs per compound are reduced. Second, the carrier molecule itself may act to protect a particular functional group of the desired final molecule (*e.g.*, aldehyde, alcohol, amine, carboxylic acid, ketone, etc.) from further chemical reaction. Third, the growing molecule can be protected from interaction with itself, since low loading of the carrier will prevent contact (and hence reaction) of individual molecules. Fourth, through the use of large excesses of reagents, one can force the reaction of a particular synthesis step further towards complete conversion. However, the attachment of building blocks to the solid support is an additional reaction step and one must ensure that all following reaction conditions will harm neither the resin nor the anchor. As a corollary to this last point, one must also ensure that the cleavage conditions necessary to remove the product from the resin do not damage the final product. It becomes obvious, that the choice of the solid support and the cleavage strategy is important for a successful synthesis.

Should it be possible to assay the final products in a carrier-bound form, then cleavage from the resin is not necessary; unfortunately, in most cases the compounds synthesized must be obtained dissolved in solution in order to carry out, for example, biological testing. The basis for this requirement include (but are not limited to) the following factors: compounds in solution have better chances to interact with (for example) membrane-bound proteins, the conformation of soluble compounds will not be affected by the carrier material when it has been removed and, in particular, bioassays can then be carried out under physiological conditions.

In order to ensure that there can be no misunderstandings of the meanings of the terms linker, anchor and spacer in this overview, we would like to take this opportunity to define the following (Fig. 5.1):

- The **spacer** is located between the resin carrier and the linker and, when present, should reduce factors such as steric hindrance of reactions brought about by the bulk of the resin or to modify features such as the hydrophilicity/hydrophobicity of the local environment and promote one or more of the reaction steps. Following cleavage of the final product, the spacer remains attached to the resin carrier. The use of spacers is optional; they are not always necessary, but may often be advantageous.
- The **linker** is a bifunctional molecule which, on the one hand can be bound irreversibly to the carrier phase (resin) and, on the other hand, offers a reversible binding site for the coupling of desired molecules so that further chemical reactions may be carried out. In the simplest cases the linker remains unchanged during the synthesis and, after the cleavage of the product, remains attached to the carrier so that the resin can be reused and further syntheses on the same linker are possible.
- The **anchor** is defined as a resin-immobilized functional group which forms a cleavable coupling to the first building block used in the synthesis. This means that a linker becomes an anchor after it is immobilized on a resin. The different names are important,

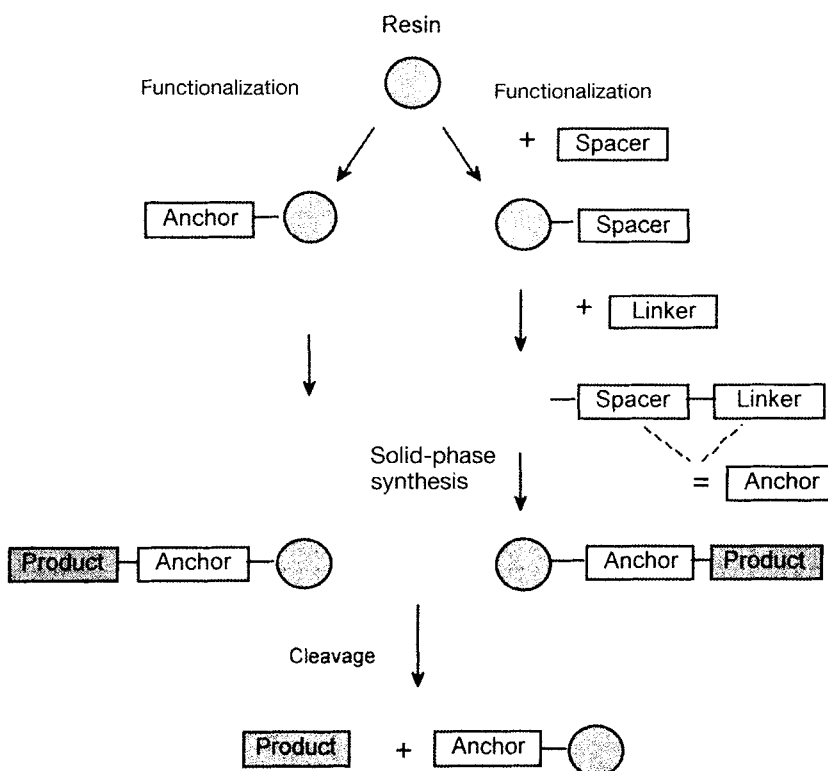


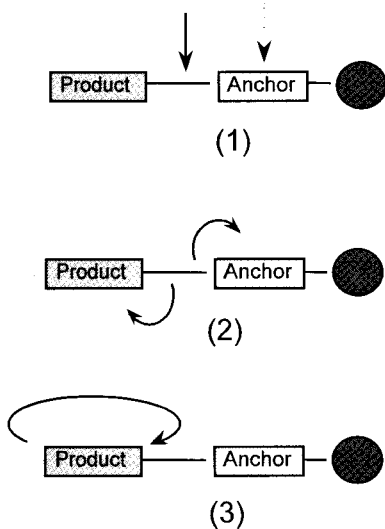
Figure 5.1. Explanation of the terms anchor, linker and spacer.

since many anchors are synthesized directly on the resin (*e.g.*, trityl- and benzhydryl-resins) and not as a soluble bifunctional molecule (linker) prior to its immobilization.

During the period of development of solid-phase peptide synthesis (SPPS), many new synthesis strategies, as well as novel linkers, have been developed; many of these can also be used for solid-phase organic chemistry (SPOC). The most commonly used strategies in SPPS are the Boc/Bnz and Fmoc/tBu strategies. In these two cases, the temporary N-terminal protecting group of the growing peptide chain is specifically cleaved prior to each new coupling cycle; Boc groups are removed with trifluoroacetic acid (TFA) and Fmoc with secondary amines, such as piperidine. In order to use these strategies, it is essential that the anchor remains stable under the conditions used for removal of this N-terminal-protecting group (*i.e.*, under acidic or basic conditions, respectively). In order to characterize linkers and anchors that have already been used in SPPS, the compatibility to either or both of these strategies is mentioned under 'observation'.

In order to promote the final cleavage of the product from the resin, a large number of possible chemical strategies have been found (Fig. 5.2). The most common methods among these involve a cleavage reagent, which attacks a labile bond between the solid-phase carrier and the product (**1**). The cleavable bond may lie within the anchor and allows a predictable modification of the final product (*e.g.*, see multifunctional anchor, silicon anchor, etc.). A second possibility is to use a reagent cocktail for the cleavage (*e.g.*, see aminolysis and hydrazinolysis) to replace the resin in a substitution reaction (**2**). A variation of this procedure involves the formation of a free functional group on the resin which reacts with the anchor and results in a cyclization, concomitantly allowing the cleavage of the product from the resin (**3**) [14, 15].

An interesting development has been the design of anchors which allow the formation of multiple products after cleavage. This possibility has been achieved by resins which carry a number of different anchor structures requiring orthogonal cleavage conditions [16,



**Figure 5.2.** Different possibilities for the cleavage of the final product from the polymeric carrier.



17]. Another variation depends on the slow cleavage kinetics of products when treated with gaseous TFA. The cleaved products are then brought into solution during a wash step. These two steps can then be repeated a number of times and the uncleaved products can be stored still attached to the resin [18].

Depending on the chemical structure of the anchor and chemistry of its attachment to the resin, the product can be cleaved at the end of the synthesis either with acid, base or nucleophilic cleavage reagents, hydrogenolysis, enzymatic, palladium-catalyzed or photochemical, oxidative and reductive cleavage methods. In addition, the 'safety-catch' anchors can be chemically modified at the end of the synthesis to provide a structure which is subsequently cleavable.

All these differences are considered in the following summary, sorted according to the cleavage conditions and structural characteristics of the different linkers and anchors which have been reported in the literature. The list of the structures is alphabetic in each section. However, such an alphabetical listing can prove difficult in finding a particular structure, since names of many anchors/linkers differ in the literature. Many trivial names are currently in use, and the naming of many of the structures using IUPAC nomenclature is often complicated and inconvenient. For this reason, a detailed index has been provided, in which the structures are listed by both systematic and trivial names. To further facilitate searching the list, the structures of the respective linkers/anchors are shown below the names in the list. To provide a better overview, the functional groups used for immobilization of the linker on the resin are always shown at the right-hand side of each structure; when an anchor is featured in the structure, the resin is located on the right-hand side. The cleavage reagents and conditions and the product(s) obtained can be found under the headings 'Cleavage' and 'Product after cleavage', respectively. In those cases where multiple methods for the cleavage from the same anchor have been reported, this is marked under the heading 'Observations' and the reader is directed to the appropriate sections. Under 'Chemical Reactions', typical syntheses which have been carried out on the particular anchor are mentioned. These examples should describe the principal applications of the anchor, but represent in no case an exhaustive listing. Even if the searched synthesis conditions are not mentioned, these observations should assist in finding an appropriate anchor for the planned synthesis.

Since the carrier material (*e.g.*, polystyrene, controlled pore glass, polyacryl, cellulose, etc.) also affects the success of the synthesis, they will be considered in the following sections, together with the linkers and anchors under 'Carrier' [19].

An increasing interest can also be found for functionalized resins with immobilized catalysts or activating reagents. In addition, the so-called inverse solid-phase chemistry, in which the product is synthesized in solution and the resin is then used to pull out the product from the reagents and by-products, is currently finding a number of applications [20, 21]. However, these resins will not be considered in the following list since, in these cases no reactions actually take place on the resin itself. However, they have particular applications for combinatorial chemistry for the work-up procedures if by-products are a serious problem.

The list '*Solid-phase Anchors in Organic Chemistry*' can never be absolutely complete. Negative results obtained with novel anchors will be very rarely published and many possible combinations (resin/linker/reaction) have not yet been tried. However, this list can

be of considerable assistance to the advancement of combinatorial chemistry on solid supports.

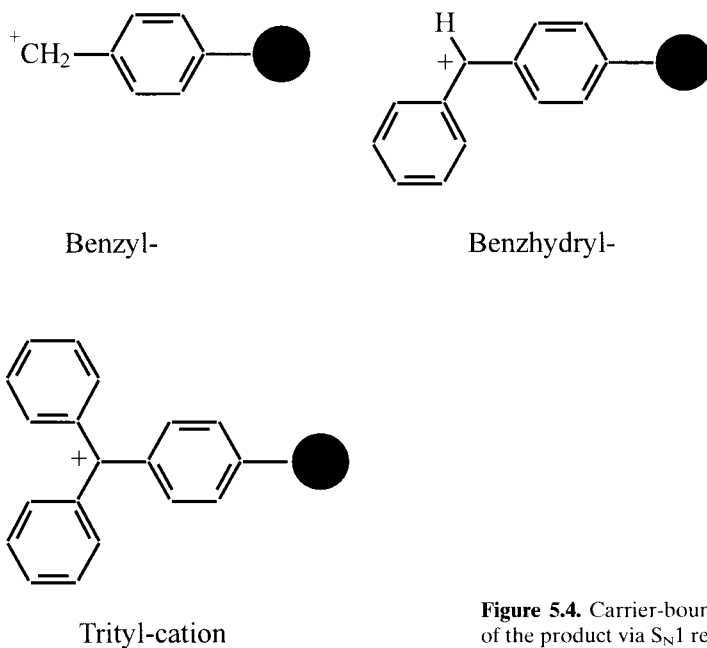
## 5.2 Acid-Labile Anchors

Most synthesis strategies in both peptide and combinatorial chemistry involve acid-labile anchors. Compounds such as TFMSA, HF, HBr, TFA, PPTS, acetic acid and HFIP can be employed as cleavage reagents and generate a number of products, including carboxylic acids, carboxamides, thiocarboxylic acids, alcohols, thiols, ketones, amines and hydrazides. The greatest part of acid-labile anchors can be subdivided into two categories. The first subgroup is characterized by an acid-labile acetal group which is obtained following addition of an alcohol to a 2,3-dihydro-4H-pyran (Fig. 5.3).



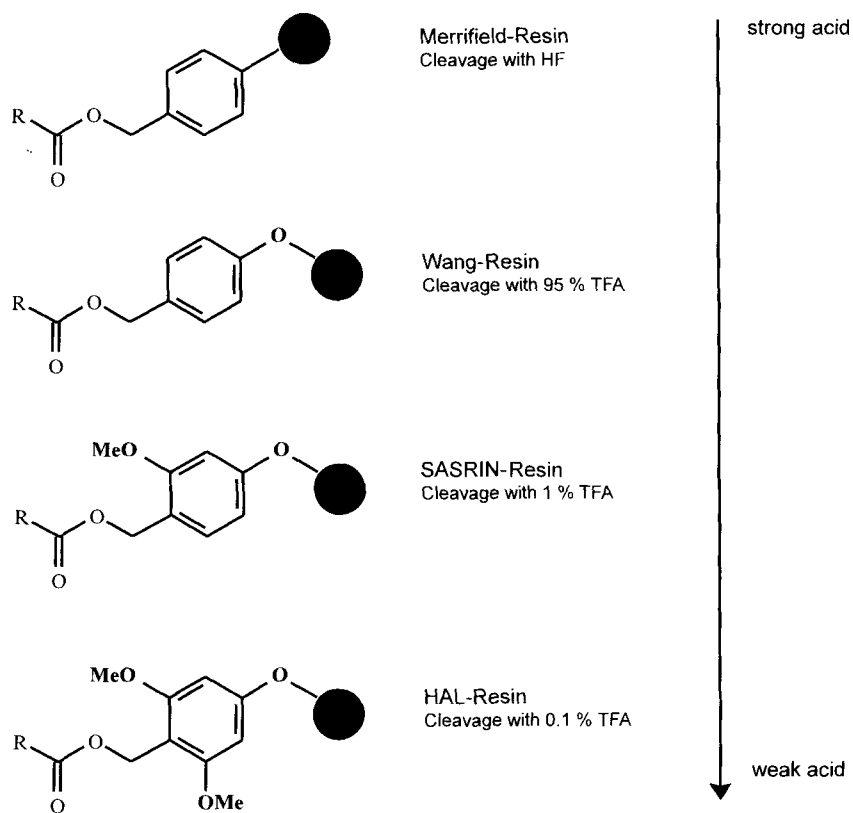
**Figure 5.3.** Formation of a carrier-bound acid-labile acetal.

The second subgroup, which contains the largest number of acid-labile anchors, can be characterized by their ability to form stable cations. Typical members of this group include functional groups linked with benzyl-, benzhydryl- and trityl-anchor (Fig. 5.4).



**Figure 5.4.** Carrier-bound cations after the cleavage of the product via  $S_N1$  reaction.

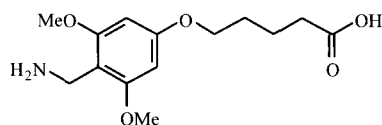
The desired acid lability of the anchor can be altered through the substitution of different aromatic substituents. In general, it can be said that electro-positive substituents (*e.g.*, methoxy, amino or hydroxy groups) raise the stability of the intermediate cation and therefore the use of weaker acids is possible for cleavage. This fine-tuning of the reaction conditions of the cleavage is depicted in Figure 5.5, in which the benzyl linker is used as an example. One disadvantage of increased electron density, is the concomitantly higher reactivity of the anchor system and the subsequent danger of losing the product before the last synthesis step as a result of premature cleavage. In planning a synthesis, the stability of the anchor during the synthesis and the sensitivity of the anchor during cleavage must be carefully considered.



**Figure 5.5.** Dependence of the cleavage conditions on the aromatic substituents.

Because of the very mild cleavage conditions required when using SASRIN- or HAL-resins, even Boc-protected amines remain unaffected during cleavage. Thus, further reactions such as head-to-tail cyclization or fragment condensations are possible, where protected reactive functions are required.

**5-(4-Aminomethyl-3,5-dimethoxy-phenoxy)-valeric acid**  
ADPV linker



<b>Cleavage:</b> TFA 0.1% in DCM 60 min, 25 °C	<b>Product after cleavage:</b> Carboxamide [22–24]
--	---

**Observations:**

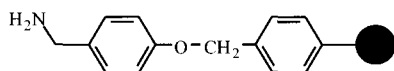
- Compatible with Fmoc/tBu SPPS strategy (Boc-N-protective groups stable during cleavage)
- To avoid side products, a low conc. TFA/high conc. TFA cleavage strategy is recommended [25].

**Chemical reaction:**

- Pd-catalyzed deallylation [26].
- Phase transfer catalysis, Heck reaction [27], Mitsunobu reaction [28], nucleophilic substitution [28], Wittig reaction [29].

- Carrier:
- Aminomethyl-PS/DVB (PAL resin)
  - MBHA-PS/DVB
  - HMP-PS/DVB
  - PS-PEG
  - CLEAR (Cross-linked ethoxylate acrylate resin) [30]

**4-(Aminomethyl)phenoxyethyl-PS/DVB**

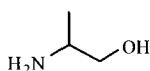


<b>Cleavage:</b> 50% TFA/DCM	<b>Product after cleavage:</b> Carboxamide [31, 32]
---------------------------------	--

**Observations:**

- Stable against 1% TFA.
- Peptide synthesis with Ppoc-N-protection.

**2-Aminopropanol**



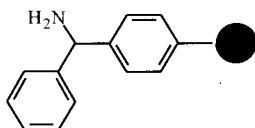
Carrier: Chloromethyl-PS/DVB

<b>Cleavage:</b> Weak acids	<b>Product after cleavage:</b> Ketone [33]
--------------------------------	---

**Observations:**

The linker is immobilized through the hydroxy group to the resin. A ketone is immobilized on the resin by condensation with the amine.

**Benzhydrylamine-PS/DVB**  
BHA-Anchor



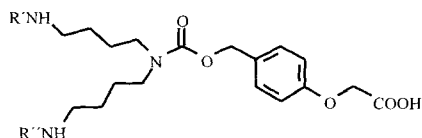
<b>Cleavage:</b> TFMSA	<b>Product after cleavage:</b> Carboxamide [32, 34, 35]
---------------------------	--

**Observations:**

Compatible with Boc/Bnz SPPS strategy.

**Chemical reaction:**

Reductive amination [36]

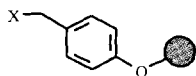
**N,N-Bis-(4-aminobutyl)-aminocarboxy-oxymethyl-phenyloxyacetic acid**

Carrier: Aminofunctionalized -PEG/PS

<b>Cleavage:</b> TFA	<b>Product after cleavage:</b> Amine [37]
-------------------------	--

**Observations:**

R' and R can be two orthogonal protecting groups (Fmoc, Boc, Allyl, etc.).

**4-Bromo- (Iodo-) methylphenoxy-PS/DVB**

<b>Cleavage:</b> TFA	<b>Product after cleavage:</b> Amine [38]
-------------------------	--

**Observations:**

- X = Br, I
- Anchor is generated by treating Hydroxymethylphenoxy-PS/DVB with  $\text{Ph}_3\text{PBr}_2$  or  $\text{Ph}_3\text{PI}_2$ .

**Chemical reactions:**

N-Alkylation

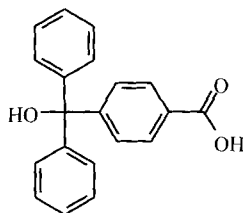
**4-Carboxybenzaldehyde [39]**

Carrier: Aminofunctionalized PS/DVB

<b>Cleavage:</b> TFA	<b>Product after cleavage:</b> Diol
-------------------------	--

**Observations:**

Compatible with Fmoc/tBu SPPS strategy.  
The immobilization of a diol leads to acetal formation.

**4-Carboxytrityl alcohol**

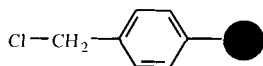
Carrier: Aminofunctionalized PEG-PS

<b>Cleavage:</b> - AcOH/MeOH/DCM or 0.5% TFA or HFIP - 5-50% TFA/1% TIS	<b>Product after cleavage:</b> - Carboxylic acid [40] - Alcohol, Amine, Hydrazide [41]
--	---

**Observations:**

- Before the immobilization of the nucleophile, the alcohol must be converted to the chloride by treatment with  $\text{AcCl}$  or  $\text{SOCl}_2$  [42–44]
- Compatible with Fmoc/tBu SPPS strategy.

**Chloromethyl-PS/DVB**  
Merrifield resin



Cleavage:	Product after cleavage:
- HF/Thioanisole	- Carboxylic acid [45]
- DIBAH	- Alcohol [46]
- SnCl <sub>4</sub>	- Alcohol [47]

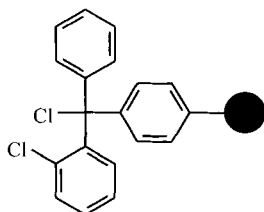
**Observations:**

- Carboxylic acids are immobilized as their cesium salt.
- See also base-labile linkers.
- Stable also to hydration with NaCNBH<sub>3</sub>/MeOH/HCl.
- Stable against TFA.
- In order to obtain alcohols a resin-attached ether bond is cleaved by SnCl<sub>4</sub>.

**Chemical reaction:**

Bischler–Napiralski, Dieckmann cyclization [15], Suzuki reaction [48], Wittig reaction, ozonolysis, condensation, esterification, nucleophilic substitution [49], Henry reaction, 1,3-dipolar cyclo-addition, electrophilic addition [50], oxidation: chloride → aldehyde [50], sulfide → sulfone [51], alcohol → ketone, Arbuzov reaction (phosphine-phosphoroxide) [52], reduction: hydration [45], ester → alcohol [49, 53]

**2-Chlorotriylchloride-PS/DVB**



- Carrier: - PS/DVB (1%) [54, 59, 60]  
- PS/DVB (2%) [43, 61, 62]

Cleavage:	Product after cleavage:
- AcOH/TFE/DCM 1–2 h, 0.5% TFA, HFIP	- Carboxylic acid [54]
- 5–50% TFA/5% TIS (or 30% HFIP in DCM)	- Carboxylic acid - Alcohol [55], Phenol, Amine [41, 56]

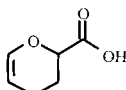
**Observations:**

- Commercially available resins are preloaded (*e.g.* with amino acids, alcohols, phenols, diamines).
- Compatible with Fmoc/tBu SPPS strategy (Boc-N-protective groups also remain after cleavage)

**Chemical reaction:**

Mitsunobu [57], Mannich [58] reaction, and many more.

**1,2-Dihydropyrane carboxylic acid**



Carrier: Chloromethyl-PS/DVB

Cleavage:	Product after cleavage:
3% HCl	Tetrazole [63]

**Observations:**

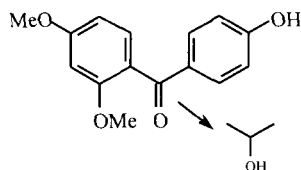
Stable against LiOH and diazomethane.

**Chemical reactions:**

Suzuki reaction.

**2,4-Dimethoxy-4'-hydroxybenzophenone**

Rink-Acid resin [64]



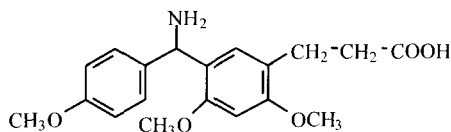
Carrier: Chloromethyl-PS/DVB

**Cleavage:**0.2% TFA in DCM  
3–5 min or AcOH in  
DCM**Product after cleavage:**

Carboxylic acid

**Observations:**

- After coupling to Merrifield resin the ketone is reduced to the alcohol, then the first building block can be immobilized.
- Very acid-labile anchor, even HOBt can partially cleave the product from the solid support (this problem is circumvented by adding DIEA).
- Compatible with Fmoc/tBu SPPS strategy (Boc-N-protective groups also remain after cleavage).

**2,4-Dimethoxy-5-((4-methoxyphenyl)-aminomethyl)-phenylpropionic acid**

Carrier: Aminomethyl-PS/DVB

**Cleavage:**

- Diluted TFA [65]
- Me<sub>3</sub>SiBr/TFA [66]

**Product after cleavage:**

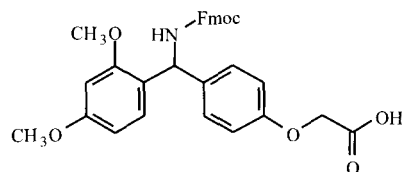
Carboxamide

**Observations:**

- Compatible with Fmoc/tBu SPPS strategy.
- Higher acid sensitivity as 4,4'-dimethoxybenzhydryl linker.

**4-(2',4'-Dimethoxyphenyl-Fmoc-aminomethyl)-phenoxyacetic acid**

Rink-'acetic acid'-amide, Knorr resin

**Cleavage:**

TFA

**Product after cleavage:**

Carboxamide [64]

**Observations:**

Compatible with Fmoc/tBu SPPS strategy.

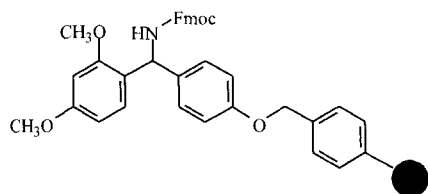
**Chemical reaction:**

Mitsunobu reaction [67], enzyme-catalyzed reactions [67], oxidation: sec. alcohol → ketone, reduction: ketone → alcohol [69].

Reductive cleavage of disulfide bridges (best results with PEG resin [70]), reductive N-alkylation [71], microwave-assisted Suzuki and Stille reaction [72].

- Carrier:
- With norleucine-loaded aminomethyl-polystyrene
  - With norleucine-loaded 4-methylbenzhydryl-polystyrene (MBHA resin)
  - Aminofunctionalized-PEG-polystyrene [68]
  - Polyacryl pins [69]

**4-(2,4-Dimethoxyphenyl-Fmoc-aminomethyl)-  
phenoxymethyl-PS/DVB**  
Rink-Amide resin



Carrier: All common resins and Sepharose beads [80]

**Cleavage:**  
TFA

**Product after cleavage:**  
– Carboxamide [64, 73]  
– Sulfonamide [74]

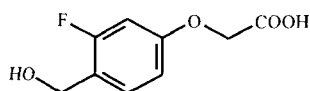
**Observations:**

- Compatible with Fmoc/tBu SPPS strategy.
- Higher stability towards reduction than Rink-Amide resin equipped with a norleucine spacer.
- Stable towards 1.7% conc. HCl in THF [75].

**Chemical reaction:**

[3 + 2]-Cyclo-addition [76] (nitroxide → isoxazole; alkene → isoxazoline), 1,3-dipolar cyclo-addition (pyrrole), Ugi 4-components reaction [75], Aza-Wittig reaction, N-alkylation [77], Stille reaction [78], Heck reaction [74], Pd-catalyzed amination with primary and secondary alkyl- or arylamines [79].

**3-Fluoro-4-(hydroxymethyl)phenoxyacetic  
acid [81]**



Carrier: TentatGel S NH<sub>2</sub>

**Cleavage:**  
TFA:H<sub>2</sub>O (2:1)

**Product after cleavage:**  
Carboxylic acid

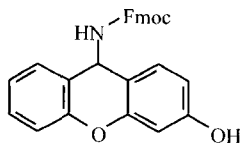
**Observations:**

Fluorinated linker for monitoring SPOS by gel-phase <sup>19</sup>F NMR.

**Chemical reactions:**

N- and O-acylation, N-alkylation, carbanion chemistry.

**9-Fmoc-amino-xanthen-3-ol**



Carrier: – Chloromethyl-Polystyrene (Sieber-Amide resin)  
– Aminofunctionalized-PEG-polystyrene

**Cleavage:**  
TFA (1%)

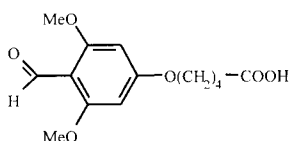
**Product after cleavage:**  
Carboxamide [24, 82–84]

**Observations:**

Compatible with Fmoc/tBu SPPS strategy.



**5-(4-Formyl-3,5-dimethoxy-phenoxy)-valeric acid**  
Backbone Amide Linker (BAL) [85]



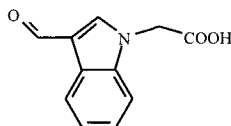
Carrier: PEG-PS/DVB

<b>Cleavage:</b> TFA	<b>Product after cleavage:</b> Amide
-------------------------	---

**Observations:**

- Amine is immobilized by reductive amination. This can be done prior to or after immobilization of the linker to the resin. N-acylation gives resin-bound amide.
- Compatible with Fmoc/tBu SPPS synthesis.
- Diketopiperazine formation after deprotection of the amino function of the second amino acid is possible.
- Stable to short treatments of 2% TFA in H<sub>2</sub>O.
- On-resin cyclization of peptides.

**3-Formyl-indol-1-yl-acetic acid**  
Indol-Resin [86]



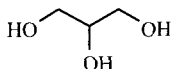
Carrier: Aminofunctionalized PS/DVB

<b>Cleavage:</b> TFA	<b>Product after cleavage:</b> sec. Amide Sulfonamide Urea Guanidine Carbamate Aniline
-------------------------	--

**Observations:**

Amine is immobilized by reductive amination. N-acylation gives resin-bound amide. A wide range of amine modification is accessible.

**Glycol**



Carrier: Chloromethyl-PS/DVB

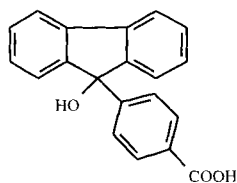
<b>Cleavage:</b> Diluted HCl	<b>Product after cleavage:</b> Aldehyde [87-89], Ketone [90]
---------------------------------	--

**Observations:**

The linker is immobilized as 2,2-dimethyl-1,3-dioxolane-4-methanol (derivative of glycerol).

**Chemical reaction:**

Acetylation, chlorination, reduction (aldehyde → alcohol), Benzoin condensation, Wittig-, Grignard-, Aldol reaction, Suzuki-Miyaura cross-coupling, directed ortho-metalation [91].

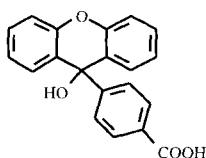
**9-Hydroxy-9-(4-carboxyphenyl)fluorene** [92]

Carrier: Aminomethyl-PS

<b>Cleavage:</b> TFA/TFMSA	<b>Product after cleavage:</b> Amide
-------------------------------	---

**Observations:**

- Stable to TFA.
- After immobilization on the resin, the anchor is chlorinated at position 9 with  $\text{SOCl}_2$ .

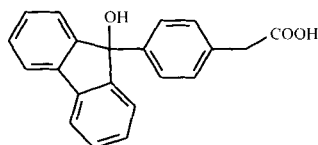
**9-Hydroxy-9-(4-carboxyphenyl) xanthene**

Carrier: Aminofunctionalized PES/PS/DVB

<b>Cleavage:</b> TFA	<b>Product after cleavage:</b> Carboxamide [93]
-------------------------	--

**Observations:**

The hydroxy group of the immobilized linker is transferred to the amine by chlorination and ammonia.

**4-(9-Hydroxyfluorenyl) benzylic acid**  
PhFI Acetic Acid [94]

Carrier: Aminofunctionalized polymers

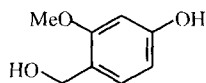
<b>Cleavage:</b> TFA	<b>Product after cleavage:</b> Amine Carboxylic acid Phenol Aniline
-------------------------	---

**Observations:**

- X = RCOO, Ph-O, RNH, Ph-NH.
- More stable to acidic conditions than trityl-resins.

**Chemical reactions:**

Pd-catalyzed deallylation, reduction.

**4-Hydroxy-2-methoxybenzylalcohol**

Carrier: Chloromethyl-PS/DVB (SASRIN Super Acid Sensitive Resin)

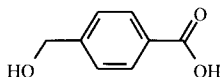
<b>Cleavage:</b> TFA 1 %	<b>Product after cleavage:</b> Carboxylic acid [95, 96]
-----------------------------	--

**Observations:**

- Compatible with Fmoc/tBu SPPS strategy (Boc-N-protective groups also remain after cleavage).
- Immobilization on the resin by esterification of the phenoxy group.
- See also base-labile linkers.

**Chemical reaction:**

Pd-catalyzed deallylation, carboxylic acid activation with Pfp-OH [97, 98]

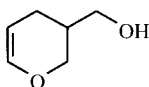
**Hydroxymethylbenzoic acid**  
HMBA

Carrier: – Aminomethyl-PS/DVB  
– Cellulose

<b>Cleavage:</b> TFMSA/TFA	<b>Product after cleavage:</b> Amine [99]
-------------------------------	--

**Observations:**

- Compatible with Fmoc/tBu SPPS strategy.
- See also base-labile linkers.
- By treating the anchor with p-nitrophenylchloroformate, an active carbonate is formed; amines can be immobilized by forming a carbamate bonding.

**3-Hydroxymethyl-2,3-dihydro-4H-pyran**

Carrier: Chloromethyl-PS/DVB

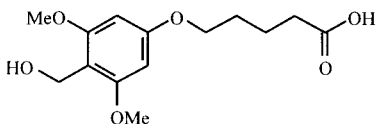
<b>Cleavage:</b> – PPTS/BuOH/DCM (Pyridinium-p-toluene sulfonate) – 10% TFA	<b>Product after cleavage:</b> – Alcohol [100–103] – Substituted indole [104]
--	---

**Observations:**

- Stable to nucleophiles and bases.
- Compatible with Fmoc/tBu SPPS strategy.
- Immobilized aniline resists cold acetic acid and 2 N HCl.

**Chemical reaction:**

Grignard reaction, reduction: Azide → Amine, nucleophilic substitution, N-alkylation.

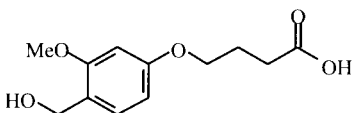
**5-(4-Hydroxymethyl-3,5-dimethoxy-phenoxy)-valeric acid**

Carrier: Aminomethyl-PS/DVB (HAL resin)

<b>Cleavage:</b> TFA 0.1% in DCM 60 min, 25 °C	<b>Product after cleavage:</b> Carboxylic acid [105]
--	---

**Observations:**

- Compatible with Fmoc/tBu SPPS strategy (Boc-N-protective groups also remain after cleavage).

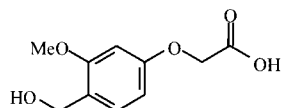
**4-(4-Hydroxymethyl-3-methoxyphenoxy)-butyric acid**

Carrier: – Aminomethyl-polystyrene (HMPB resin)  
– Benzhydrylamine-polystyrene  
– 4-Methoxybenzhydrylamine-polystyrene

<b>Cleavage:</b> – TFA 1–5% in DCM	<b>Product after cleavage:</b> – Carboxylic acid [106–108] – Alcohol [57, 109]
--	--

**Observations:**

- Compatible with Fmoc/tBu SPPS strategy (Boc-N-protective groups also remain after cleavage).
- Alcohols are immobilized under Mitsunobu-conditions.

**4-Hydroxymethyl-3-methoxyphenoxy-acetic acid**

Carrier: Aminomethyl-polystyrene/DVB

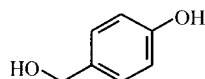
<b>Cleavage:</b> TFA 1%	<b>Product after cleavage:</b> Carboxylic acid [110, 111]
----------------------------	--

**Observations:**

Compatible with Fmoc/tBu SPPS strategy.

**4-Hydroxymethylphenol**

Wang- or HMP- Resin



<b>Cleavage:</b> – TFA – HF or TFA/TFMSA – AlCl <sub>3</sub> /amine – 10% TFA	<b>Product after cleavage:</b> – Carboxylic acid [112, 113] – Amine [114] – Alcohol [115, 116] – Alcohol, phenol [57, 109, 117] – Carboxamide [118] – Guanidine [119]
---	---

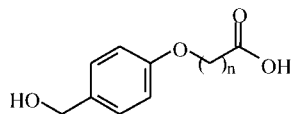
**Observations:**

- Compatible with Fmoc/tBu SPPS strategy.
- More acid-labile than Merrifield resin.
- Stable against bases and reductive conditions.
- Alcohols are immobilized by a Mitsunobu reaction.
- Stable against dilute HCl [75]
- In order to obtain amines the Wang resin is transformed into the p-nitrophenyl or succinimidyl carbonate and treated with amines. The formed carbamate bond is cleaved with TFA.
- In order to obtain alkyl-, aryl- and acyl- guanidines Wang linker is treated with carbonylimidazole and thiourea consecutively. The displacement of sulfur by primary or secondary amines leads to acid-labile carbamate linkage [119].

**Chemical reaction:**

Pd-catalyzed deallylation [98], phosphorylation [120], asymmetric hydration with Rh-Catalysis [121], intramolecular cyclization/Heck reaction [122], Stille reaction [78], Biginelli condensation [123], 1,3-dipolar cyclo-addition, Ugi 4-components reaction [75]; if an alcohol is immobilized, the p-alkoxybenzyl ether bond is stable to enolate chemistry, Mitsunobu chemistry, ozonolysis, Grignard reaction, reduction (aldehyde – alcohol) under sonication, free radical reaction (allylation of esters) [124].

Carrier: – Chloromethyl-PS/DVB (Wang resin, HMP resin)  
– PEG-polystyrene

**4-Hydroxymethylphenoxyacetic- or -propionic acid**

<b>Cleavage:</b> TFA or TFMSA	<b>Product after cleavage:</b> – Carboxylic acid [108, 111] – Alcohol, phenol [57, 109] – Amine [99, 116] – Amidine [125]
----------------------------------	---

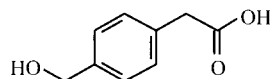
**Observations:**

- Compatible with Fmoc/tBu SPPS strategy.
- $n = 1, 2$  [126].
- More acid-labile than Merrifield resin.
- Stable against bases.
- Alcohols are immobilized by a Mitsunobu reaction.
- By treating the anchor with p-nitrophenylchloroformate or DSC an active carbonate is formed; amines can be immobilized by forming a carbamate bonding.

**Chemical reaction:**

N-Alkylation [127, 128], Stille-reaction [129]

Carrier: – Aminofunctionalized PEG-PS/DVB or PS

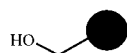
**4-Hydroxymethylphenylacetic acid**

<b>Cleavage:</b> HF or TFMSA	<b>Product after cleavage:</b> Carboxylic acid [130, 131]
---------------------------------	--

**Observations:**

- See also base-labile anchors.
- 100 times more stable than Merrifield linker.
- Compatible with Fmoc/tBu- and Boc/Bnz-SPPS-strategy.

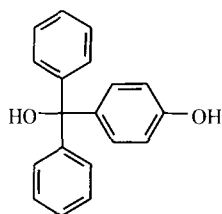
Carrier: Aminofunctionalized-PS/DVB (PAM resin)

**Hydroxymethyl-PS**

<b>Cleavage:</b> HF	<b>Product after cleavage:</b> Alcohol [116]
------------------------	---

**Observations:**

- Resin is treated with DSC to obtain activated succinimidyl carbonate.
- Carbonate bond is partially cleaved by amines (piperidine).

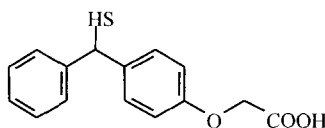
**4-Hydroxytrityl alcohol**

Carrier: Bromofunctionalized PEG-polystyrene

<b>Cleavage:</b> 1–5% TFA in DCM	<b>Product after cleavage:</b> Alcohol [132], Amine, Carboxylic acid, Hydra- zide, Thiol [133]
-------------------------------------	---

**Observations:**

- Before immobilizing the nucleophile, the alcohol must be converted to the chloride by treatment with AcCl or SOCl<sub>2</sub> [42–44].
- Compatible with Fmoc/tBu SPPS strategy.

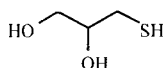
**4-( $\alpha$ -Mercaptobenzyl)-phenoxyacetic acid**

Carrier: Aminofunctionalized-PS/DVB

<b>Cleavage:</b> HF	<b>Product after cleavage:</b> Thiocarboxylic acid [134]
------------------------	---

**Observations:**

Compatible with Boc/Bnz SPPS strategy.

**3-Mercapto-1,2-propanediol**

Carrier: Chloromethyl-PS/DVB

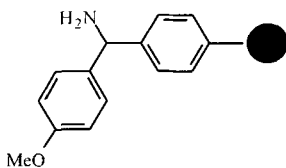
<b>Cleavage:</b> Aqueous dioxan 95 °C, 10 h, toluene sulfonic acid (20 %)	<b>Product after cleavage:</b> Aldehyde, ketone [135]
--	--

**Observations:**

- Immobilization of the linker on the resin at the thiol group.
- Immobilization of the carbonyl group by azeotropic removal of water in refluxing benzene.

**Chemical reaction:**

Reduction, hydroboration, oxidation (vinyl  $\rightarrow$  alcohol).

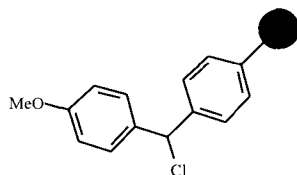
**4-Methoxy-benzhydrylamine-PS/DVB**  
MOBHA resin

<b>Cleavage:</b> HF	<b>Product after cleavage:</b> Carboxamide [32, 136]
------------------------	---

**Observations:**

- Stable against HCl/HOAc.
- Compatible with Boc/Bnz SPPS strategy.
- More efficient cleavage of the product as observed with BHA resin.

**4-Methoxy-benzhydryl chloride-PS/DVB**  
(MAMP, Alpha-Methoxy Phenyl resin)[137]

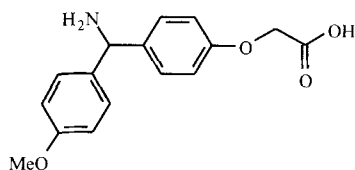


<b>Cleavage:</b> 10% TFA/DCM	<b>Product after cleavage:</b> Amide
---------------------------------	---

**Observations:**

- Amines are immobilized and acylated.
- High loadings (1–2 mmol/g).
- Stable to basic conditions.
- Stable at 80 °C.

**4-((4-Methoxyphenyl)-aminomethyl)-phenoxyacetic acid**



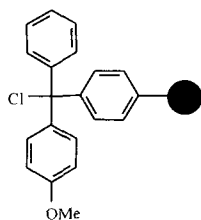
Carrier: Aminomethyl-PS/DVB

<b>Cleavage:</b> TFA	<b>Product after cleavage:</b> Carboxamide [138, 139]
-------------------------	--

**Observations:**

Compatible with Fmoc/tBu SPPS strategy.

**4-Methoxytritylchloride-PS/DVB**

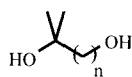


<b>Cleavage:</b> 1–5% TFA + TIS	<b>Product after cleavage:</b> Alcohol, Amine, Carboxylic acid, Hydra- zide, Thiol [41, 56]
------------------------------------	--

**Observations:**

- Disulfides can be obtained by releasing thiols in the presence of electrophilic oxidation reagents (z. B. I<sub>2</sub>, TI<sup>3+</sup>) [133].
- Compatible with Fmoc/tBu SPPS strategy.

**x-Methyl-1,x-alkane diol**



<b>Cleavage:</b> 10% TFA/DCM	<b>Product after cleavage:</b> Primary and secondary amines [140].
---------------------------------	--

**Observations:**

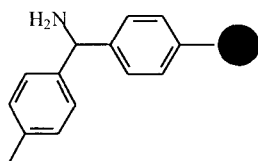
- $x = 3-6$ .
- Prim. alcohol (linker) is immobilized on Merrifield resin.
- Tert. alcohol (linker) is treated with CDI to form the activated carbamate.
- Activation of the immobilized imidazole is increased by methylation with MeOTf.
- The formed (tert. alkoxy-carbonyl)-3-methylimidazolium triflate is used to immobilize amines.

**Chemical reactions:**

- Acylation.

Carrier: Chloromethyl-PS/DVB

**4-Methylbenzhydrylamine-PS/DVB**  
MBHA resin



<b>Cleavage:</b> HF or TFMSA	<b>Product after cleavage:</b> Carboxamide [141]
---------------------------------	---

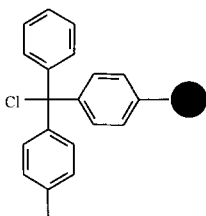
**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- More acid-labile than BHA resin.

**Chemical reaction:**

Reductive methylation [142], N-alkylation, reduction with  $\text{BH}_3/\text{THF}$ , cyclization with thiophosgene and triphosgene [143].

**4-Methyltritylchloride-PS/DVB**

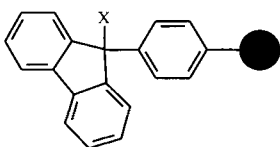


<b>Cleavage:</b> - 1-5 % TFA + 5 % TIS or 30 % HFIP in DCM	<b>Product after cleavage:</b> - Alcohol, Amine [89], Thiol [41] - Carboxylic acid [60]
---	--

**Observations:**

Compatible with Fmoc/tBu SPPS strategy (Boc-N-protectible groups also remain after cleavage).

**9-Phenylfluoren-9-yl – PS/DVB**  
PhFI Resin [144]



<b>Cleavage:</b> TFA	<b>Product after cleavage:</b> Amine Carboxylic acid Phenol Aniline
-------------------------	---

**Observations:**

- $\text{X} = \text{RCOO}, \text{C}_6\text{H}_5\text{O}, \text{RNH}$ .
- More stable to acidic conditions than trityl-resins.

**Chemical reactions:**

Claisen condensation, pyrazole synthesis, Pd-catalyzed deallylation.

**Phosgene**



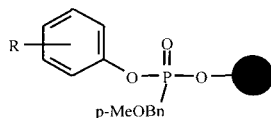
Carrier: Hydroxymethyl-PS/DVB

<b>Cleavage:</b> HBr/TFA	<b>Product after cleavage:</b> Amine [145]
-----------------------------	---

**Observations:**

Inverse peptide synthesis by coupling azides [146].



**Phosphate ester** [147]

Carrier: Wang-PS/DVB

<b>Cleavage:</b> 30% TFA/DCM	<b>Product after cleavage:</b> Phenylphosphate
---------------------------------	---

**Observations:**

Compatible with Fmoc/tBu SPPS strategy

**Chemical reactions:**

Reductive amination, Pd-catalyzed allyl deprotection.

**Piperazine**

Carrier: Chloromethyl PS/DVB

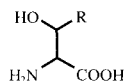
<b>Cleavage:</b> 3% TFA	<b>Product after cleavage:</b> Ketone [148, 149]
----------------------------	---

**Observations:**

- Piperazine is immobilized to Merrifield resin in DMSO at 70 °C.
- Stable against strong bases.
- Propargyltriphenylphosphine bromide is immobilized on the piperazine to form acid-labile enamines.

**Chemical reactions:**

Anion chemistry, Wittig reaction, [4+2] cyclo-addition chemistry.

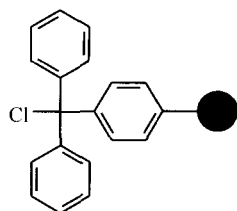
**Serine/Threonine**

Carrier: Methacrylic acid/dimethylacrylamide grafted polyethylene pins

<b>Cleavage:</b> TFA or AcOH	<b>Product after cleavage:</b> Peptide aldehyde [150]
---------------------------------	--

**Observations:**

- Condensation with acyl fluorides.
- A support-bound Ser or Thr is treated with aldehyde to give an imine intermediate, which cyclizes to give a stable (but acid-labile) oxazolidine.

**Trylchloride-PS/DVB**

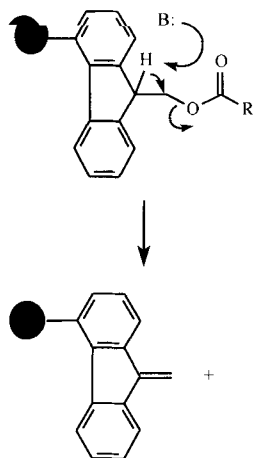
<b>Cleavage:</b> - 4-Toluene sulfonic acid - 2% TFA/DCM - 1 N HCl/THF - 10% TFA	<b>Product after cleavage:</b> - Carboxylic acid [42, 56]  - 4-Hydroxyphenyl-sulfonates [151]
---	--

**Observations:**

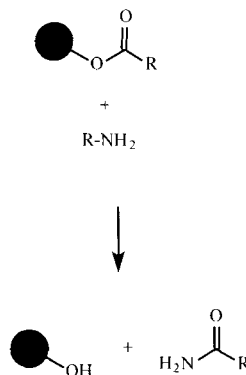
- Compatible with Fmoc/tBu SPPS strategy.
- Compatible with Rh(I)-catalyzed hydroformylation (75 atm of CO/H).

### 5.3 Anchors Cleaved by Nucleophiles

The second synthesis strategy commonly used in peptide synthesis and combinatorial chemistry is based on base-labile anchors. The cleavage mechanism is based either on a  $\beta$ -elimination (Fig. 5.6) or a hydrolysis, re-esterification or aminolysis (Fig. 5.7).



**Figure 5.6.**  $\beta$ -Elimination on fluorenyl linker.



**Figure 5.7.** Aminolysis of ester linker.

The reagents used and the products liberated are depicted in Table 5.1. From a single anchor structure and synthesis, different products can be generated with the use of different cleavage reagents providing a further source of diversity (*e.g.*, with hydroxymethylbenzoic acid as linker  $\rightarrow$  five carboxylic acid derivatives).

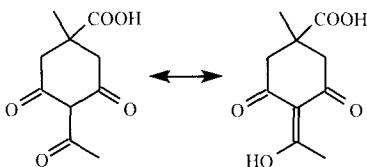
Other cleavage procedures listed in Section 5.3 include: use of a Grignard reagent to generate alcohols, the synthesis of amines by cleavage with hydrazine (ADCC linker), and the generation of tertiary amines by Hoffmann-elimination (acrylic acid linker).

**Table 5.1.** Cleavage reagents and products.

Cleavage reagent	Product
NaOH	Carboxylic acid
Amine/NH <sub>3</sub>	Carboxamide
Hydrazine	Hydrazide
Alcohol	Carboxylic acid ester

**4-Acetyl-3,5-dioxo-1-methyl-cyclohexyl-carboxylic acid**

ADCC linker



Carrier: Aminomethyl-PS/DVB

**Acrylic acid**

REM resin

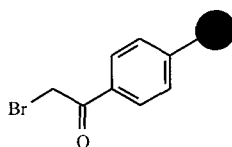
(Regenerated Michael addition)



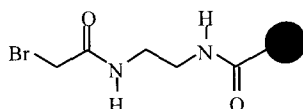
Carrier: Hydroxymethyl resin

**Bromoacetyl-PS/DVB**

[155]

**N-(N-Bromoacetyl)-2-aminoethyl-Polyacrylamide**

[156, 157]



<b>Cleavage:</b> 2% Hydrazine/DMF	<b>Product after cleavage:</b> Amine [152]
--------------------------------------	---

**Observations:**

- Stable against acids (TFA).
- Stable against bases (Piperidine, DBU).

<b>Cleavage:</b> DIEA	<b>Product after cleavage:</b> <i>tert</i> -Amine [153, 154]
--------------------------	---

**Observations:**

- X = Cl, OH.
- Immobilization on the resin through the acid function.
- Immobilization of the first building block by adding primary or secondary amines (Michael-Addition).
- Cleavage by a Hoffmann elimination.
- Compatible with acid- and base-sensitive protecting groups.

**Chemical reaction:**

Reductive alkylation, N-alkylation.

<b>Cleavage:</b> - NaOH <sub>aq.</sub> - MeOH/Imidazole - NH <sub>3</sub> aq. - Hydrazine hydrate	<b>Product after cleavage:</b> - Carboxylic acid - Carboxylic acid ester - Carboxamide - Hydrazide
---	--

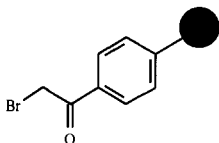
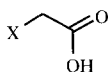
**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- More sensitive to bases than benzyl derivative.

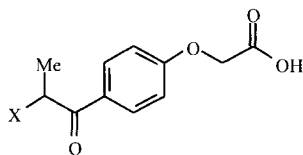
<b>Cleavage:</b> - Alcohol - Ammonia - Trifluoroethanol/ NH <sub>3</sub>	<b>Product after cleavage:</b> - Carboxylic acid ester - Carboxamide - Carboxylic acid
--	---

**Observations:**

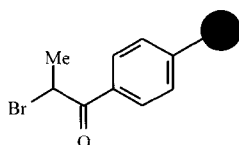
- Compatible with Boc/Bnz SPPS strategy.
- Labile against nucleophiles; nevertheless, the Fmoc/tBu-SPPS-Strategy can be applied [158].

**2-Bromoacetyl-PS/3 % DVB** [159]**2-Bromo- bzw. iodo-acetic acid**

Carrier: - Polyacrylate resin  
- CPG

**4-(2-(Br or Cl)propionyl)phenoxyacetic acid**

Carrier: - Aminomethyl-PS/DVB [162, 163]

**4-(2-Bromopropionyl)phenyl-PS/DVB**

<b>Cleavage:</b> - $(\text{CH}_3)_3\text{SnOH}/$ - $[(\text{C}_4\text{H}_9)_3\text{Sn}]_2\text{O}$ (BBTO)	<b>Product after cleavage:</b> - Carboxylic acid
---	---

**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- Cleavage of the phenylacetyl ester proceeds with complete retention of N-Boc and O-Bnz-protecting groups.
- Stable against acids.

<b>Cleavage:</b> - NaOH - $\text{NH}_2\text{NH}_2$	<b>Product after cleavage:</b> - Carboxylic acid [158] - Hydrazone [160]
--	--

**Observations:**

- $x = \text{I}, \text{Br}$ .
- Stable against HF,  $\text{FSO}_3\text{H}$ , TFA, TFMSA.
- Although saponification is feasible, the anchor is stable against secondary amines.

**Chemical reaction:**

- Oligosaccharide synthesis.
- Enzymatic reactions.

<b>Cleavage:</b> - Saponification - Photolysis (350 nm) - Esterification with alcohol - Hydrazine	<b>Product after cleavage:</b> - Carboxylic acid [161] - Carboxylic acid - Carboxylic acid ester - Hydrazone
---	--

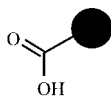
**Observations:**

- $X = \text{Br}, \text{Cl}$ .
- Compatible with Boc/Bnz SPPS strategy.
- See also photolysis-labile anchors.

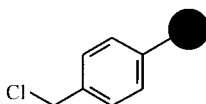
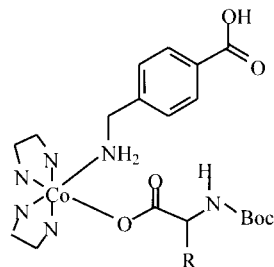
<b>Cleavage:</b> - Saponification - Photolysis - KCN/18-crown-ether-6 in DMF - $\text{NH}_3$ in MeOH - Hydrazinolysis	<b>Product after cleavage:</b> - Carboxylic acid [164] - Carboxylic acid [163] - Carboxylic acid [165] - Carboxamide [155, 166] - Hydrazone [163]
--	--

**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- The formation of diketopiperazine and Schiff bases as by-product in SPPS is reported [165, 167]

**Carboxy-PS/DVB****Chloromethyl-PS/DVB**

Merrifield resin  
[170–172]

**Cobalt(III)-complex**

Carrier: Aminomethyl-PS/DVB

<b>Cleavage:</b> – Methyl-Mg-Br [168] – Bu <sub>4</sub> NOH/THF [169]	<b>Product after cleavage:</b> Alcohol [168, 169]
---	--

**Observations:**

By treating the anchor with thionylchloride the activated acid chloride is obtained.

**Chemical reaction:**

Oxidation (lead tetra-acetate, iodosobenzenebisacetate), formation of dehydrobenzene, Diels–Alder reaction, inverse Wittig reaction.

<b>Cleavage:</b> – K <sub>2</sub> CO <sub>3</sub> /(nBu) <sub>4</sub> NHSO <sub>4</sub> , – Et <sub>3</sub> N/ROH [173, 174], – KCN/ROH [175, 176], – 2-Dimethyl-aminoethanol – DMAE [170]	<b>Product after cleavage:</b> – Carboxylic acid – Carboxylic acid ester
---	--

**Observations:**

- See also acid-labile anchors.
- Stable against TFA.

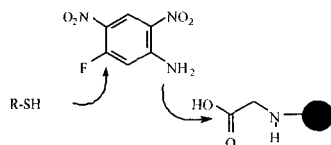
**Chemical reaction:**

- Photochemical reactions on the resin.
- Formation of amides applying chromocarbon [177]
- Pictet–Spengler reaction [178]

<b>Cleavage:</b> Mercaptoethanol in DMF	<b>Product after cleavage:</b> Carboxylic acid [179]
---	---

**Observations:**

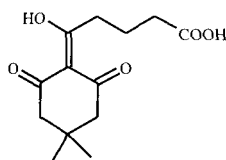
- Compatible with Fmoc/tBu- and Boc/Bnz-SPPS-strategy
- High stability against TFA (50% TFA in DCM, 72 h leads to 3.1% cleavage)
- The first building block is attached to the linker in a solution-phase reaction.

**3-Cysteinyl-4,6-dinitroanilide of N-carboxymethyl-aminomethyl-PS/DVB**

<b>Cleavage:</b> Thiolysis	<b>Product after cleavage:</b> R1-Cys-R2 [180]
-------------------------------	---

**Observations:**

- The linker is immobilized through an anilide formation. N-Boc-cysteine is coupled to the anchor by a nucleophilic substitution. Further reactions can be performed either on the COOH or NH<sub>2</sub> of cysteine.
- Compatible with Boc/Bnz-SPPS strategy.

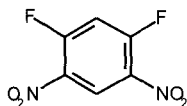
**Dde-based primary amine linker [181]**

Carrier: Aminofunctionalized PEG-PS/DVB

<b>Cleavage:</b> 2% Hydrazine hydrate or transamination	<b>Product after cleavage:</b> Primary amine
--	---

**Observations:**

- Primary amines are immobilized as first building blocks.
- Stable to acid and base conditions.
- Compatible with Fmoc/tBu SPPS strategy.

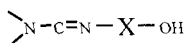
**4,6-Difluoro-1,3-dinitrobenzene**

Carrier: Aminomethyl-PS/DVB

<b>Cleavage:</b> Thiolysis with thiophenol (6 h, rt)	<b>Product after cleavage:</b> Imidazole
---	---

**Observations:**

- Stable against TFA and HBr/HOAc
- Stable against pyridine
- In the first nucleophilic substitution the linker is immobilized on the resin.
- With the second nucleophilic substitution an imidazole derivative (e.g. histidine) is immobilized.

**N-(Hydroxy-(alkyl or aryl)-dimethylformamide)**

X = different aryl and alkyl chains

Carrier: Chloromethyl resin

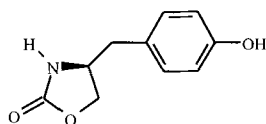
<b>Cleavage:</b> Hydrazine/AcOH, KOH in MeOH, ZnCl <sub>2</sub> in EtOH or LiAlH <sub>4</sub>	<b>Product after cleavage:</b> Secondary amine [182, 183]
---	---

**Observations:**

- Secondary amines are immobilized by an amine exchange to give solid-phase-bound amidines.

**Chemical reactions:**

- Anion alkylation, O-alkylation with benzylic halides (Williamson reaction), oxidation, reduction.

**5-(4-Hydroxybenzyl)-oxazolidone** [184]

Resin: Wang-PS/DVB

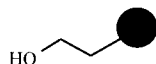
<b>Cleavage:</b> – NaOMe-THF – LiOH, H <sub>2</sub> O <sub>2</sub> , –20°C, THF	<b>Product after cleavage:</b> – Methyl ester – Carboxylic acid
--	---

**Observations:**

- After metalation of the nitrogen the carbamate can be acylated by activated carboxylic acids.
- Attachment of the linker to resin under Mitsunobu conditions.

**Chemical reactions:**

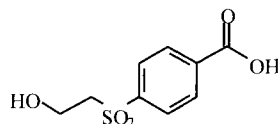
Enantioselective aldol condensation.

**Hydroxyethylene-PS/DVB**

<b>Cleavage:</b> Amine/Al(CH <sub>3</sub> ) <sub>3</sub>	<b>Product after cleavage:</b> Carboxamide [185]
---	---

**Chemical reaction:**

Aldol addition, reductive cyclization, synthesis of quinoline.

**4-(Hydroxyethylsulfonyl)-benzoic acid**

<b>Cleavage:</b> 0.1M NaOH or Ba(OH) <sub>2</sub>	<b>Product after cleavage:</b> Carboxylic acid [186]
--	---

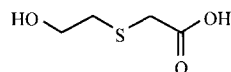
**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- Linkers with different oxidation states of the sulfur (sulfoxide, sulfide) are described.

**Chemical reaction:**

Oligodesoxynucleotide synthesis

Carrier: Ethylamine functionalized polydimethylacrylamide resin

**5-(2-Hydroxyethyl)-mercaptoacetic acid**

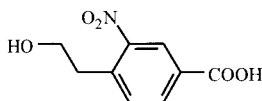
<b>Cleavage:</b> Aqueous NaOH in Dioxane/MeOH	<b>Product after cleavage:</b> Carboxylic acid [187]
---	---

**Observations:**

- Compatible with Fmoc/tBu- and Boc/Bnz-SPPS-strategy.
- After attaching the first building block, the linker is oxidized to a sulfone (no oxidation-sensitive building blocks can be used).

Carrier: – Hydroxymethyl-PS/DVB  
– Aminomethyl-PS/DVB

**4-Hydroxyethyl-3-nitro-benzoic acid**  
NPE resin



Carrier: Aminomethyl-PS/DVB

<b>Cleavage:</b>	<b>Product after cleavage:</b>
<ul style="list-style-type: none"> <li>- 0.1 M DBU in DCM, 2 h</li> <li>- 20% Pip in DMF, 2 h</li> </ul>	Carboxylic acid [188, 189]

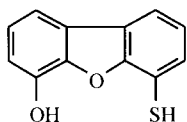
**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- SPPS without the formation of diketopiperazine.

**Chemical reaction:**

Synthesis of oligonucleotides and oligosaccharides.

**8-Hydroxy-1-mercaptodibenzofuran**



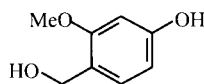
Carrier: Chloromethyl-PS/DVB

<b>Cleavage:</b>	<b>Product after cleavage:</b>
nBu <sub>3</sub> P in Dioxane-H <sub>2</sub> O	Mercapto-dibenzofuranyl ester [190, 191]

**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- After protection of the sulfide, an amino acid is coupled to the hydroxy group; Z-N-Cys(Trt)-OH is immobilized on Merrifield resin; after removal of the trityl side-chain protecting group with methoxycarboxy-sulfonylchloride (SmcCl) the dibenzofuran linker is added, resulting in the formation of a disulfide bond.
- Linker remains completely in the product.

**4-Hydroxy-2-methoxybenzyl alcohol**



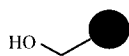
Carrier: Chloromethyl-PS/DVB

<b>Cleavage:</b>	<b>Product after cleavage:</b>
<ul style="list-style-type: none"> <li>- Hydrazine (10–20%)</li> <li>- Amine</li> <li>- NaBH<sub>4</sub>/LiBr</li> <li>- Alcohol + LiBr/DBU</li> </ul>	<ul style="list-style-type: none"> <li>- Hydrazide [192]</li> <li>- Carboxamide [193, 194]</li> <li>- Alcohol [195, 196]</li> <li>- Carboxylic acid ester [175, 176, 197]</li> </ul>

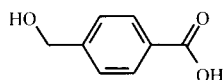
**Observations:**

- Immobilization to the resin through the phenolic group.
- Compatible with Fmoc/tBu SPPS strategy (Boc-N-protective groups remain stable during cleavage).
- See also acid-labile anchors.

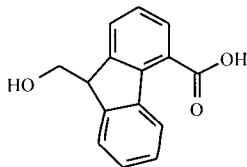


**Hydroxymethyl-PS/DVB****4-Hydroxymethylbenzoic acid**

HMBA  
[111, 202–204]



Carrier: – PEGA  
– 4-Methoxybenzhydrylamino-PS  
– Aminofunctionalized-PEG-PS  
– Cellulose [206]

**1-(9-Hydroxymethylfluorenyl)-formic acid**

Carrier: Aminomethyl-PS/DVB

<b>Cleavage:</b>	<b>Product after cleavage:</b>
– Primary or secondary amine	– Hydantoine [198]
– Alcohol/TMSCl	– Quinazoline [199]
– LiOH, MeOH, H <sub>2</sub> O (reflux).	– Carboxylic acid ester
	– Substituted benzoic acid

**Observations:**

- Cleavage can be performed under cyclization, if an amine is generated on the resin.
- The hydroxy group is activated by p-nitrophenylchloroformate; amines can be immobilized by forming a carbamate bonding.

**Chemical reaction:**

Ozonolysis of alkenes to give carboxylic acids, aldehydes or alcohols [200]; directed ortho metalation, Stille cross-coupling [201].

<b>Cleavage:</b>	<b>Product after cleavage:</b>
– Ammonia	– Carboxamide
– Secondary amine	– Secondary Carboxamide
– Hydrazine	– Hydrazide
– NaOH	– Carboxylic acid
– MeOH (Et <sub>3</sub> N/50 °C)	– Methyl ester

**Observations:**

- Cleavage can be performed under cyclization, if the nucleophile is generated on the resin.
- Stable against AcOH, 70 °C, ZnCl<sub>2</sub>.

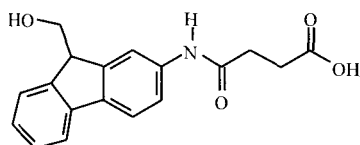
**Chemical reaction:**

Fischer indole synthesis [205]

<b>Cleavage:</b>	<b>Product after cleavage:</b>
15 % Piperidine 5 min	Carboxylic acid [207]

**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- Prolonged treatment with DIEA leads to partial cleavage.

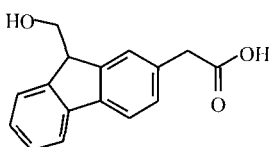
**3-(9-Hydroxymethylfluorenyl)-amido-butanedioic acid**

Carrier: Aminomethyl-PS/DVB

<b>Cleavage:</b> 20 % Morpholine	<b>Product after cleavage:</b> Carboxylic acid [208]
-------------------------------------	---

**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- Stable in DIEA/DMF-solutions.

**3-(9-Hydroxymethylfluorenyl)-acetic acid**

Carrier: Benzhydrylamino-PS/DVB

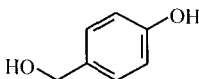
<b>Cleavage:</b> 15 % Piperidine (5 min)	<b>Product after cleavage:</b> Carboxylic acid [209]
---	---

**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- More stable against nucleophilic components than 1-(9-hydroxymethylfluorenyl)-formic acid; nevertheless it is recommended to neutralize the solution during the synthesis by a weak acid (e.g. HOBt).

**4-Hydroxymethylphenol**

Wang resin



Carrier: - Chloromethyl-PS/DVB

<b>Cleavage:</b> - LiAlH <sub>4</sub> - NaOMe or LiOH	<b>Product after cleavage:</b> - Methylamine [210] - Sulfonamide [211]
---	--

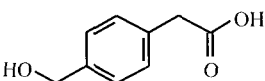
**Observations:**

- Wang resin is transformed into the p-nitrophenyl carbonate and treated with amines. The formed carbamate bond (acid labile) can be reduced and cleaved to the corresponding methylamine.
- In order to obtain the sulfonamides the carbamate bond is treated with RSO<sub>2</sub>Cl.
- See also acid-labile anchors.

**4-Hydroxymethylphenylacetic acid**

[130]

PAM resin; 'Phenylacetamidomethyl resin'

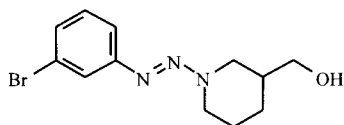


Carrier: Aminofunctionalized-PS/DVB

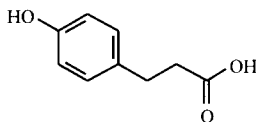
<b>Cleavage:</b> - THF/H <sub>2</sub> O, LiBr/DBU - MeOH/DBU/ LiBr, EtOH/Ti(OEt) <sub>4</sub>	<b>Product after cleavage:</b> - Carboxylic acid - Carboxylic acid ester
--	--

**Observations:**

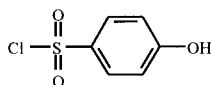
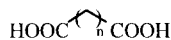
- 100 times more stable than Merrifield resin.
- Compatible with Fmoc/tBu and Boc/Bnz-SPPS-strategy.
- See also acid-labile anchors.

**3-[(3-Hydroxymethyl)-piperidylazo]-bromobenzene**

Carrier: Chloromethyl-PS/DVB

**3-(4-Hydroxyphenyl)-propionic acid**

Carrier: BHA-PS/DVB

**4-Hydroxyphenylsulfonylchloride** [214, 215]**Resin:** Chloromethyl-PS/DVB**Malonic acid/Succinic acid/Glutaric acid**

Carrier: – Aminomethyl-PS/DVB, -PEG/PS  
 – Hydroxy-PS/DVB, -PEG/PS  
 – Silica gel  
 – CPG

<b>Cleavage:</b> CH <sub>3</sub> I	<b>Product after cleavage:</b> Iodo-substituted phenyl derivatives [212]
---------------------------------------	---

**Observations:**

Triazen linker is immobilized on the resin through the hydroxy group.

**Chemical reaction:**

Pd-catalyzed substitution of acetylene.

<b>Cleavage:</b> – Saponification – Re-esterification with 2-dimethylamino-ethanol (DMAE) – Ammonia	<b>Product after cleavage:</b> – Carboxylic acid [213] – DMAE-Ester [170]  – Carboxamide
--	--

**Observations:**

– DMAE-Ester can be saponified by aqueous DMF.  
 – The side chains of Asn and Glu are also saponified.

<b>Cleavage:</b> Amine Imidazole Thiolate	<b>Product after cleavage:</b> Substituted amine Substituted imidazole Thioether
--	---

**Observations:**

Alcohols are immobilized.

**Chemical reactions:**

Grignard reaction, Wittig reaction, reduction (NaBH<sub>4</sub>), reductive amination, Suzuki coupling.

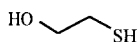
<b>Cleavage:</b> – NaOH or MeOH/ K <sub>2</sub> CO <sub>3</sub> , NH <sub>4</sub> OH <sub>conc.</sub> , 50 °C – 0.5 M LiCl, NH <sub>4</sub> OH, 55 °C, 16 h	<b>Product after cleavage:</b> – Carboxylic acid [216]  – 3'-modified oligonucleotides [217]
--	---

**Observations:**

– n = 1, 2, 3.

**Chemical reaction:**

Oxidation: phenol → biphenyl derivatives with VOCl<sub>3</sub> [218], synthesis of oligonucleotide [219], Mitsunobu reaction [220].

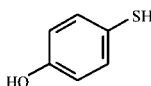
**2-Mercaptoethanol**

Carrier: Chloromethyl-PS/DVB

<b>Cleavage:</b> NaOH in Dioxane/ MeOH; 3 min	<b>Product after cleavage:</b> Carboxylic acid [221, 222]
---	---

**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- The Na<sup>+</sup>-salt of 2-mercaptoethanol in NH<sub>3</sub>,<sub>li</sub> is immobilized on Merrifield resin. The oxidation gives the 2-hydroxyethylsulfone-PS/DVB.

**4-Mercatophenol** [223]

Carrier: Chloromethyl-PS/DVB

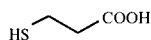
<b>Cleavage:</b> Primary amine	<b>Product after cleavage:</b> Amide
-----------------------------------	---

**Observations:**

- The linker is immobilized through the thiol group to form 4-hydroxyphenyl-thiomethyl-PS/DVB.
- Stable to acids.
- Stable at 80 °C.
- Cleavage can also be achieved via intramolecular cyclization.
- See also safety-catch linker.

**Chemical reactions:**

Picet-Spengler synthesis.

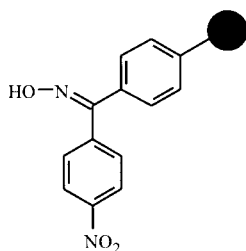
**3-Mercaptopropionic acid**  
[224]

Carrier: Aminomethyl-PS

<b>Cleavage:</b> Amine + AgNO <sub>3</sub>	<b>Product after cleavage:</b> Amide
---	---

**Observations:**

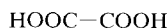
- Compatible with Boc/Bnz strategy.

**p-(4-Nitrobenzophenoxime) resin**  
Kaiser oxime resin [225-228]Carrier: - PS/DVB  
- Polyamide [229, 230]

<b>Cleavage:</b> - Hydrazinolysis - Aminolysis with amino acid ester - Hydroxypiperidine	<b>Product after cleavage:</b> - Hydrazone - Amide - Piperidine ester (+ Zn/HOAc gives Carboxylic acid)
---	---

**Observations:**

- Compatible with N-Boc/Bnz/cHex-SPPS-strategy.
- Labile against nucleophile.
- Stable against acids (25% TFA).
- Diketopiperazine formation in peptide synthesis.

**Oxalic acid**

Carrier: CPG

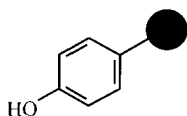
<b>Cleavage:</b> – 5% $\text{NH}_4\text{OH}/\text{MeOH}$	<b>Product after cleavage:</b> – Alcohol [231]
---	---

**Observations:**

- Oxalic acid linker is more base-stable than succinic acid linker. The linker remains on the resin.

**Chemical reactions:**

Synthesis of oligodeoxyribonucleoside methylphosphate.

**Phenol-PS/DVB/Phenol-PEG**  
[161]

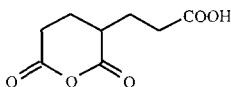
<b>Cleavage:</b> – Saponification – Re-esterification – Hydrazinolysis	<b>Product after cleavage:</b> – Carboxylic acid – Carboxylic acid ester – Hydrazide
---	---

**Observations:**

- Phenylester (compared to benzyl esters) are more easily attacked by nucleophiles and are more stable against acids.
- Saponification is accelerated with peroxides [232].
- TFA-stable.
- Compatible with Boc/Bnz SPPS strategy.

**Chemical reaction:**

1,3-Dipolar cyclo-addition, Diels–Alder reaction [233], metalorganic alkylation (Pd, Mn) [234], Mitsunobu reaction [57].

**Pentane-1,3,5-tricarboxylic acid**  
**1,3-anhydride** [235]

Carrier: Aminofunctionalized CPG or PEG-PS/DVB

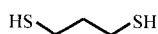
<b>Cleavage:</b> 40% $\text{MeNH}_2$	<b>Product after cleavage:</b> Amine
---	---

**Observations:**

The anhydride and the amino function of the first building block form an imide.

**Chemical reactions:**

Synthesis of modified oligonucleotides.

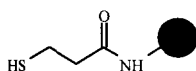
**Propane-1,3-dithiol**  
oxidative cleavable linker

Carrier: Chloromethyl resin

<b>Cleavage:</b> DMTST (Dimethylmethylthio-sulfonium-trifluoromethylsulfonate)	<b>Product after cleavage:</b> Methoxysaccharides [236]
--	--

**Observations:**

- Saccharides are immobilized as their trichloroacetimidates.
- Stable against NaOMe.

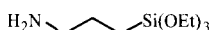
**Thiols** [237]

Carrier: PEGA

<b>Cleavage:</b>	<b>Product after cleavage:</b>
- Amine	Amide
- Transesterification	

**Observations:**

- Compatible with Boc/Bnz SPPS strategy
- Boc-amino acids are immobilized on the resin giving  $\alpha$ -thiol esters.
- The  $\alpha$ -thiol ester attachment reacts in a chemical ligation with the N-terminal of peptides to form amide bonds.
- In an Arbuzow reaction phosphite triesters can be immobilized to obtain phosphorothiolates. Iodine-mediated transesterification of phosphorothiolates leads to cyclic phosphate of oligonucleotides [238].

**3-(Triethoxysilyl)-propylamine**

Carrier: Silica gel

<b>Cleavage:</b>	<b>Product after cleavage:</b>
- $\text{NH}_4\text{OH}_{\text{conc}}/50^\circ\text{C}/15\text{ h}$	Alcohol [219]
- $(\text{Et}_3\text{N})\text{HCO}_3/\text{RT}/12\text{ h}$	

**Observations:**

The linker is immobilized through Si (other hydroxy groups are capped with TMS).

**Chemical reaction:**

Deoxynucleotide synthesis (Phosphoramidite-strategy).

**Triphenylphosphine-PS/DVB** [239]

<b>Cleavage:</b>	<b>Product after cleavage:</b>
- NaOMe	- Toluene
- Inter- and intramolecular Wittig reaction	- Alkene

**Observations:**

- Benzylbromides are immobilized on the polymer-bound phosphine.

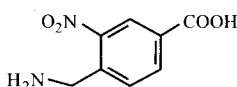
**Chemical reactions:**

Reduction: Nitrate  $\rightarrow$  Amine, condensation.

## 5.4 Photolysis-Labile Anchor

An additional orthogonal strategy to the acid- and base-labile structures previously described, involves the use of photolytically labile anchors. These anchors were first introduced for solid-phase chemistry by Rich and Zehavi (1973). Products with functional groups such as carboxylic acids, carboxamides, amidines or hydroxy can be liberated by irradiation with light of wavelength 320–365 nm. A very useful advantage of this technique is that the cleavage can often be performed in aqueous solution, and the liberated products can then be directly employed in biological assays.

### 4-Aminomethyl-3-nitrobenzoic acid



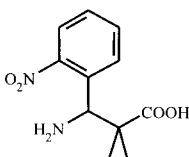
Carrier: Aminomethyl-PS/DVB (NBA resin)

<b>Cleavage:</b> Photolysis (350 nm)	<b>Product after cleavage:</b> Carboxamide [240–242]
---	---

**Observations:**

- Compatible with Fmoc/tBu and Boc/Bnz-SPPS strategy.
- The time requirement for cleavage is prolonged due to the transformation of the linker to the nitroso aldehyde during irradiation.

### 3-Amino-3-(2'-nitrophenyl)-2,3-dimethylpropionic acid [243]



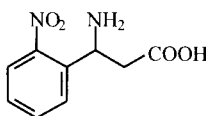
Carrier: TentaGel S NH<sub>2</sub>

<b>Cleavage:</b> 365 nm	<b>Product after cleavage:</b> Carboxamide
----------------------------	---

**Observations:**

- Stable to acetic and basic conditions.
- Stable in the presence of Lewis acids.
- Compared to the 3-amino-3-(2-nitrophenyl)propionic acid more stable to basic and Lewis acid conditions.

### 3-Amino-3-(2-nitrophenyl)propionic acid



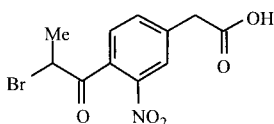
Carrier: Aminoethyl-PEG/PS (ANP resin)

<b>Cleavage:</b> Photolysis (365 nm)	<b>Product after cleavage:</b> Carboxamide [244]
---	---

**Observations:**

- Compatible with Fmoc/tBu SPPS strategy.
- Stable against acids and bases.
- Compared to the NBA resin simplified synthesis and cleavage.

### 4-(2-Bromopropionyl)-3-nitrophenylacetic acid

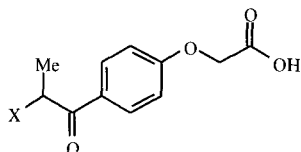


Carrier: Aminomethyl-PS

<b>Cleavage:</b> Photolysis (350 nm)	<b>Product after cleavage:</b> Carboxylic acid [245]
---	---

**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- Formation of small amounts of Schiff base and diketopiperazine is observed in peptide synthesis.

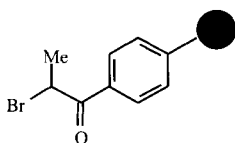
**4-(2-(Br or Cl)propionyl)phenoxyacetic acid**

Carrier: - Aminomethyl-PS/DVB  
- BHA resin [162, 163]

<b>Cleavage:</b>	<b>Product after cleavage:</b>
- Photolysis (350 nm)	- Carboxylic acid [161]
- Saponification	- Carboxylic acid
- Re-esterification	- Carboxylic acid ester
- Hydrazine	- Hydrazide

**Observations:**

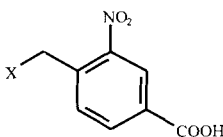
- X = Br, Cl
- Compatible with Boc/Bnz SPPS strategy
- See also base-labile anchors

**4-(2-Bromopropionyl)-phenyl-PS/DVB**

<b>Cleavage:</b>	<b>Product after cleavage:</b>
Photolysis (350 nm)	Carboxylic acid [161, 162, 167, 246]

**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- Formation of small amounts of Schiff base and diketopiperazine is observed in peptide synthesis.
- Loss of 3% product per Fmoc-deprotection with piperidine/DMF or piperidine/DCM.

**4-Bromo-(or 4-Hydroxy)methyl-3-nitrobenzoic acid**

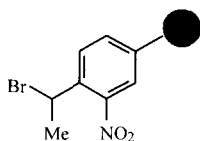
Carrier: - Aminomethyl-PS [116, 240, 247]  
- Aminobenzhydryl-PS (Nbb resin) [115, 248, 249]  
- PEG-PS [250, 251]

<b>Cleavage:</b>	<b>Product after cleavage:</b>
- Photolysis (350 nm)	- Carboxylic acid
- NaOH (dioxan/MeOH)	- alcohol carboxamide
- Hydrazinolysis	- Carboxylic acid
- Amine	- Hydrazide
- MeOH	- Carboxamide
- HF or TFMSA/TFA	- Carboxylic acid ester
	- Alcohol

**Observations:**

- X = Br, OH.
- Synthesis of side chain-protected peptides.
- Compatible with Boc/Bzl-SPPS-strategy
- Stable against acids and thiolysis..
- Labile against secondary amines (e.g. piperidine).
- Formation of diketopiperazine during peptide synthesis.
- See also base- and acid-labile anchors.

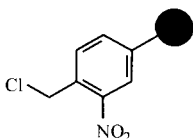


**4-(1-Bromoethyl)-3-nitrophenyl-PS/DVB**

<b>Cleavage:</b> Photolysis (350 nm)	<b>Product after cleavage:</b> Carboxylic acid [252]
---	---

**Observations:**

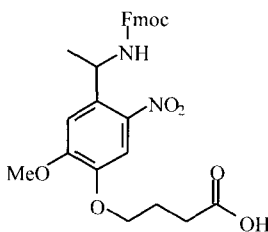
- Compatible with Boc/Bnz SPPS strategy
- More efficient cleavage than 4-bromomethyl-3-nitrobenzoic acid-linker if larger amounts of resin is used (> 500 mg).
- Low yields for peptides with more than five amino acid residues.

**4-Chloromethyl-3-nitrobenzyl-PS/DVB**

<b>Cleavage:</b> - Photolysis (350) - Hydrogenolysis	<b>Product after cleavage:</b> - Carboxylic acid [253] - Carboxylic acid [254]
--	--

**Observations:**

- Compatible with Boc/Bnz-SPPS strategy.
- Hydrogenolysis was done in liquid phase with soluble PEG-polymers.

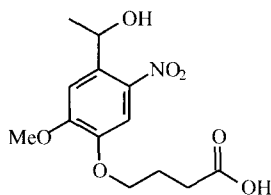
**4-(N-Fmoc-1-aminoethyl)-2-methoxy-5-nitrophenoxybutyric acid**

Carrier: Aminofunctionalized PS/DVB

<b>Cleavage:</b> Photolysis	<b>Product after cleavage:</b> Carboxamide [257–258]
--------------------------------	---

**Observations:**

- Short cleavage time required.
- Compatible with Fmoc/tBu and Boc/Bnz-SPPS-strategy.
- Stable against TFA.
- Cleavage in aqueous solution (e.g. PBS) possible.

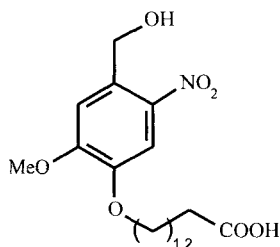
**4-((1-Hydroxyethyl)-2-methoxy-5-nitro)-phenoxybutyric acid**

Carrier: Aminofunctionalized PS/DVB

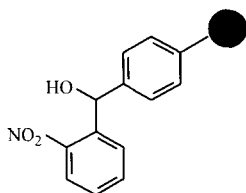
<b>Cleavage:</b> Photolysis	<b>Product after cleavage:</b> Carboxylic acid [256–259]
--------------------------------	---

**Observations:**

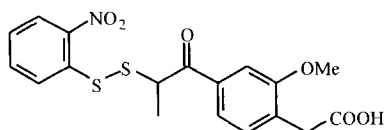
- Short cleavage time.
- Compatible with Fmoc/tBu and Boc/Bnz-SPPS-strategy.
- Stable against TFA.
- Cleavage in aqueous solution possible (e.g. PBS).

**4-Hydroxymethyl-2-methoxy-5-nitro-phenoxybutyric/propionic acid**

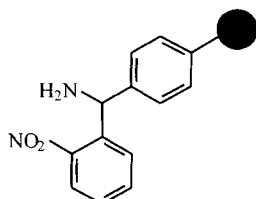
Carrier: Aminofunctionalized PS/DVB

**Hydroxy-(2'-nitro)-benzhydryl-PS/DVB****2-Methoxy-5-(2-(2-nitrophenyl)dithio)-1-oxopropylphenylacetic acid**

Npss resin



Carrier: Aminofunctionalized PEG/PS

**2-Nitro-benzhydrylamine-PS/DVB**

Carrier: PS/DVB (NBHA resin)

<b>Cleavage:</b> Photolysis	<b>Product after cleavage:</b> – Carboxylic acid [256–258] – Amidine [125] – Oligonucleotide [260]
--------------------------------	---

**Observations:**

- Short cleavage time required.
- Compatible with Fmoc/tBu and Boc/Bnz-SPPS-strategy.
- Stable against TFA
- Cleavage in aqueous solution (e.g. PBS) possible.
- The hydroxy group of nucleotides is immobilized by a carbonate bond.

<b>Cleavage:</b> Photolysis (350 nm)	<b>Product after cleavage:</b> Carboxylic acid [261]
---	---

**Observations:**

- Compatible with Boc/Bnz SPPS strategy
- Compared to the BHA resin, more stable against acids.

<b>Cleavage:</b> Photolysis 350 nm in ACN	<b>Product after cleavage:</b> Substituted toluene [262]
---	---

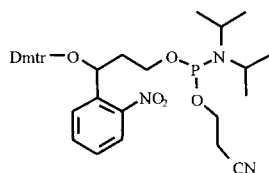
**Observations:**

- Following the cleavage of the disulfide group, benzene derivatives are immobilized through S-alkylation.
- Benzaldehydes are formed as by-products.

<b>Cleavage:</b> Photolysis (350 nm)	<b>Product after cleavage:</b> Carboxamide [242, 263]
---	--

**Observations:**

- Compatible with Boc/Bnz SPPS strategy.

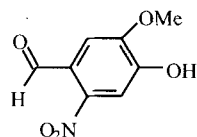
**2-Nitrophenyl-1,3-propanediol**

**Resin:** LCAA-CPG (long-chain alkyl amine-controlled pore glass)

<b>Cleavage:</b> 400 nm band-pass-filtered output of a high-pressure Hg/Xe lamp (1–2 h)	<b>Product after cleavage:</b> Oligonucleotide (up to 20 mers) [264]
--	--

**Observations:**

- Immobilization through a stable phosphoramidate diester link.
- Synthesis of base-sensitive oligonucleotides, according to cyanoethylphosphoramidate chemistry on a DNA synthesizer.

**6-Nitrovanillin**

**Carrier:** Chloromethyl-PS/DVB

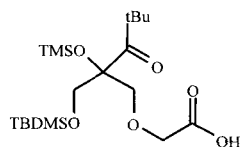
<b>Cleavage:</b> Photolysis (320 nm)	<b>Product after cleavage:</b> Alcohol (Saccharide) [265]
---	---

**Observations:**

The linker is immobilized through the hydroxy group. For the immobilization of the first building block, the aldehyde is reduced to the alcohol.

**Chemical reaction:**

Oligosaccharide synthesis

**Pivaloylglycol**  
[266]

**Carrier:** Aminofunctionalized PEG/PS

<b>Cleavage:</b> 320 nm	<b>Product after cleavage:</b> Carboxylic acid
----------------------------	---

**Observations:**

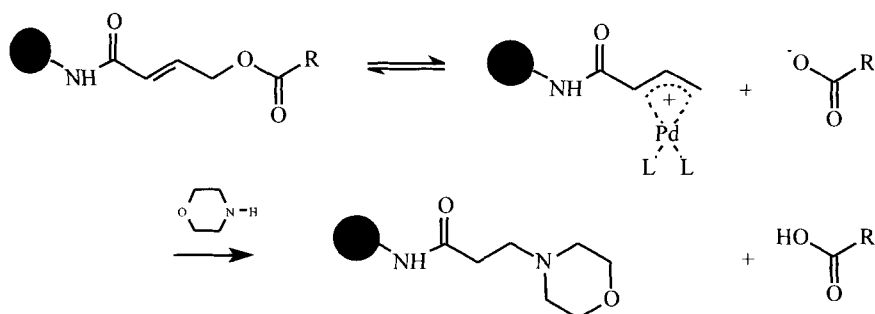
- The product is released by photolytic generation of a radical center and subsequent spontaneous  $\beta$ -C-O bond scission.
- Stable to acids and bases.

**Chemical reactions:**

Epoxidation, palladium-catalyzed cross-coupling (Stille, Suzuki reaction), peptide synthesis.

## 5.5 Allyl-Functionalized Anchors

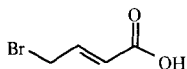
A further orthogonal strategy compatible with peptide and many other organic synthesis is offered by allyl-functionalized anchors. The mechanism of release is depicted in Figure 5.8, and involves the use of an intermediate Pd-complex with the double bond [267, 268]. Recent developments use Pd-catalyzed metathesis techniques to cleave unsaturated anchors in inter- and intramolecular reactions. An unsaturated tin-anchor has been used to achieve cleavage with an intramolecular Suzuki reaction.



**Figure 5.8.** Palladium-catalyzed cleavage of allyl esters.

The cleavage of the double bond with ozone was first reported in 1971 by Fréchet for the release of aldehydes from the solid-phase carrier. These anchors are described under the heading ‘Miscellaneous anchors cleaved by oxidation’.

#### 4-Bromocrotonic acid



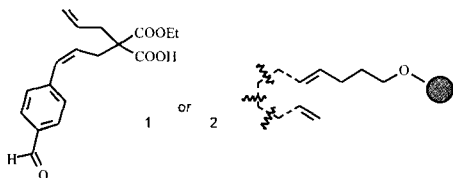
Carrier: Aminomethyl-PS/DVB  
(Hycram resin) [267]

<b>Cleavage:</b> 0.1 eq (Ph <sub>3</sub> P) <sub>4</sub> Pd + nucleophile (HOBt, morpholine, ...) in DMF	<b>Product after cleavage:</b> Carboxylic acid
---	---

#### Observations:

- Compatible with Boc/Bnz- and Fmoc/tBu-SPPS-strategy.
- Acid- and base-labile protecting groups, as well as glycosidic bondings are stable during cleavage.
- Synthesis of phosphopeptides [98].

#### Diolefinic linkers



Carrier: Chloromethyl resin

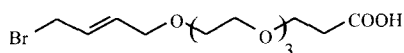
<b>Cleavage:</b> Ring closing olefin metathesis with Bis(tricyclohexyl phosphine benzylidene ruthenium dichloride)	<b>Product after cleavage:</b> Alkene [269, 270]
---	---

#### Observations:

Linker 1 is immobilized through the carboxy function. Linker 2 is generated on solid support by a Wittig reaction. A second alkene is generated during or at the end of the synthesis to give the second olefin for metathesis.

#### Chemical reactions:

Reduction, Mitsunobu, condensation reaction.  
In principle, all reactions that do not interact with olefinic structures.

**Hycron**Carrier:  $\beta$ Ala-Aminomethyl-PS/DVB

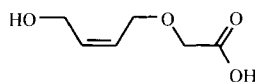
<b>Cleavage:</b> Pd(0)/N-Methylaniline	<b>Product after cleavage:</b> Carboxylic acid [271, 272]
---	--

**Observations:**

- Compatible with Fmoc/tBu and Boc/Bnz SPPS strategy.
- More stable against amines than Hycram linker (4-bromocrotoic acid).
- Stable against bases (pyridine, NaOMe) and acids (TFA).

**Chemical reaction:**

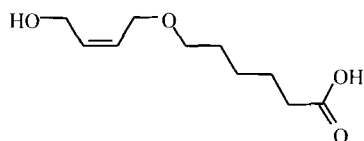
Glycosylation of peptides.

**4-Hydroxybut-2-en-oxy-acetic acid**Carrier: Chloromethyl-PS/DVB  
H-Ala-MBHA-PS/DVB

<b>Cleavage:</b> - PdCl <sub>2</sub> + nBu <sub>3</sub> SnH + H <sup>+</sup> -donator - (Ph <sub>3</sub> P) <sub>2</sub> PdCl <sub>2</sub>	<b>Product after cleavage:</b> - Carboxylic acid [268, 273] - Amine [274]
---	---

**Observations:**

- Compatible with Fmoc/tBu- and Boc/Bnz-SPPS-strategies.
- Stable against acids and bases.

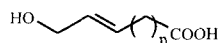
**4-Hydroxybut-2-en-oxy-caproic acid**

Carrier: Aminofunctionalized cellulose [275]

<b>Cleavage:</b> 2 eq. (Ph <sub>3</sub> P) <sub>4</sub> Pd + 3 eq. HOBT	<b>Product after cleavage:</b> Carboxylic acid
---	---

**Observations:**

- Compatible with Fmoc/tBu SPPS strategy.

 **$\omega$ -(3-Hydroxypropen-1-yl)-alkanoic acid [276]**

Carrier: LCAA (long-chain alkyl-amine) CPG

<b>Cleavage:</b> Pd (0)/tetrabutyl-ammonium format	<b>Product after cleavage:</b> Oligonucleotides
---	--

**Observations:**

- Compatible with oligonucleotide synthesis.

## 5.6 Safety-Catch Anchor

For safety-catch anchors, two independent, separate reactions are required in order to liberate the product from the solid-phase carrier. The first reaction is like a switch, in that it converts the anchor into a cleavable form; the second results in the release of the product.

### Arylsulfide



Carrier: Thiolfunctionalized PEG/PS

<b>Cleavage:</b> 1. Oxidation 2. Amine	<b>Product after cleavage:</b> Aromatic amine [277]
--	--

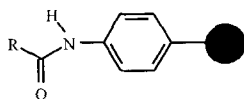
**Observations:**

- Ar = *e.g.* pyrimidine.
- The sulfide is oxidized with mCPBA (in DCM).
- The product is cleaved by nucleophile aromatic substitution.

**Chemical reactions:**

Mitsunobu reaction, reduction.

### Benzamide-PS/DVB [278]



<b>Cleavage:</b> 1) (Boc) <sub>2</sub> O 2) MeONa	<b>Product after cleavage:</b> Methylester
---	---

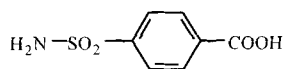
**Observations:**

- Boc-activation of the benzamide carbonyl promotes facile cleavage from the resin with methoxide.

**Chemical reactions:**

Ugi multi-component reaction (MCR).

### 4-Carboxybenzenesulfonamide [279, 280]



Carrier: Aminomethyl-PS/DVB

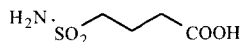
<b>Cleavage:</b> – Hydrolysis – Amine	<b>Product after cleavage:</b> – Carboxylic acid – Carboxamide
---	--

**Observations:**

- Compatible with Fmoc/tBu- and Boc/Bnz-SPPS-strategy.
- Stable against hydrolysis.
- Treatment of the linker with diazomethane or bromoacetonitrile gives base-labile N-alkylsulfonamide. The N-cyanomethyl-sulfonamide linker is also cleaved by sterically hindered amines.

**Chemical reaction:**

Suzuki reaction, alkylation, acetylation.

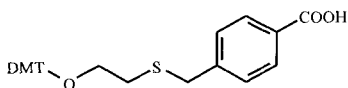
**3-Carboxypropanesulfonamide**  
 [280]


Carrier: Aminomethyl-PS/DVB

<b>Cleavage:</b>	<b>Product after cleavage:</b>
- Hydrolysis	- Carboxylic acid
- Aminolysis	- Carboxamide

**Observations:**

Compared to the 4-carboxybenzene-sulfonamide, these aliphatic sulfonamides are more nucleophilic. Therefore, the activation with bromoacetonitrile is complete and better yields are observed (especially for electrophilic substituents).

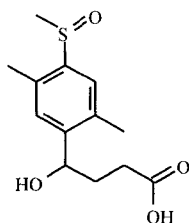
**4-[2-[4,4'-Dimethoxytriphenyl-methoxy]-ethylthiomethyl]-benzoic acid** [281]


Carrier: MBHA-PS/DVB

<b>Cleavage:</b>	<b>Product after cleavage:</b>
1) Oxidation (mCPBA)	Amine
2) NH <sub>4</sub> OH	

**Observations:**

- The first building block (an isocyanate) is attached to the liberated hydroxy group of the linker.
- The cleavage is achieved by oxidation of the sulfide to the sulfone, followed by  $\beta$ -elimination under basic conditions to give free amines.

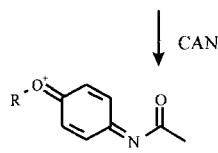
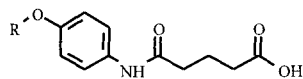
**4-(2,5-Dimethyl-4-methylsulfinylphenyl)-4-hydroxybutyric acid**  
 DSB resin


Carrier: Aminomethyl-PS/DVB

<b>Cleavage:</b>	<b>Product after cleavage:</b>
SiCl <sub>4</sub> , Thioanisole, Anisole, TFA	Carboxylic acid [282]

**Observations:**

- Compatible with Boc/Bnz-strategy.
- Reduction of the sulfoxide to the sulfide gives an acid-labile structure.

**Glutaric acid-mono-4-alkoxyanilide**

Carrier: Aminomethyl-PS

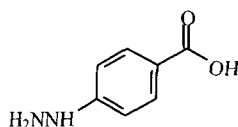
<b>Cleavage:</b> 1. CAN (ceric ammonium nitrate, $\text{Ce}(\text{NH}_4)_2(\text{NO}_3)_6$ , or DQQ 2. $\text{H}_2\text{O}$	<b>Product after cleavage:</b> Alcohol [283]
---	---

**Observations:**

- Stable against acids.
- Synthesis of saccharides.

**4-Hydrazinobenzoic acid**

[284, 285]

Carrier: Chloromethyl-PS/DVB  
TentaGel S  $\text{NH}_2$ 

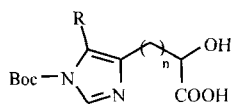
<b>Cleavage:</b> - Hydrolysis, e.g. $\text{H}_2\text{O}$ , alcohol (MeOH, EtOH, iPrOH) + base + Cu(II) acetate - $\text{NH}_3$ or $\text{PrNH}_2$ + Cu (II) acetate	<b>Product after cleavage:</b> - Carboxylic acid  - Carboxamide
---	--

**Observations:**

- Linker is immobilized on the resin through the carboxylic acid function.
- Compatible with Fmoc/tBu SPPS strategy.
- The first building block is immobilized by hydrazide formation.
- In order to remove products from the resin, the hydrazide is oxidized to the base-labile azo derivative.

**3-Imidazol-4-yl-2-hydroxypropionic acid**

[286, 287]

R = H,  $\text{CH}_3$   
n = 0, 1Carrier: -  $\beta$ Ala-Hydroxymethyl-PS/DVB  
- TentaGel S Amine  
- Cellulose

<b>Cleavage:</b> 1) TFA 2) Hydrolysis (buffer pH = 7-7.5)	<b>Product after cleavage:</b> Carboxylic acid
---	---

**Observations:**

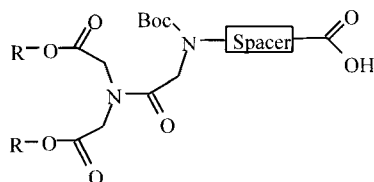
- Compatible with Fmoc/tBu-strategy.
- The linker is immobilized through its carboxy function.
- The linker is activated by the removal of the Boc-protecting group with TFA. At neutral pH the imidazole can catalyze the intramolecular hydrolysis (cleavage) of the ester bond.



**Iminodiacetic acid linker**

IdaDC

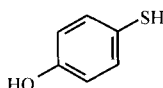
(Iminodiacetic-acid double-cleavable linker)



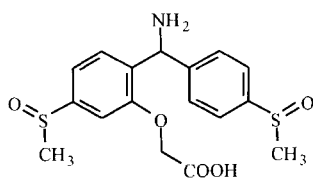
Carrier: Aminofunctionalized PEG/PS

**4-Mercaptophenol**

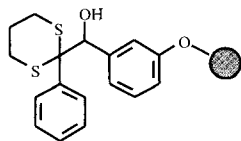
[289]



Carrier: Chloromethyl-PS/DVB

**2-(4-Methylsulfoxy)phenyl)aminomethyl)-5-methylsulfoxyphenoxyacetic acid**  
SCAL linker [290]

Carrier: Aminomethyl-PS/DVB

**2-Phenyl-2', 2''-dithiane-3-hydroxyethyl-phenoxymethyl-PS/DVB** [291]**Cleavage:**

- 1) pH 7-9/NaOH
- 2) pH 12/NaOH

**Product after cleavage:**

Two alcohols [220, 288]

**Observations:**

- The linker is activated by removal of the Boc-protecting group with TFA. Following the formation of a diketopiperazine the first alcohol is released. The second alcohol is obtained by saponification at pH 12.
- Compatible with Fmoc/tBu-strategy.

**Chemical reaction:**

Mitsunobu reaction

**Cleavage:**

- Aminolysis with amino acids

**Product after cleavage:**

- Peptide

**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- The linker is immobilized through the thiol group to form the 4-hydroxybenzene-thiomethyl-PS/DVB. In order to release the product from the resin the sulfide is oxidized with H<sub>2</sub>O<sub>2</sub> to a base-labile sulfone.
- Unsuitable for structures that are sensitive to oxidation.
- See also base-labile linker.

**Cleavage:**Reductive acidolysis  
(EtO)<sub>2</sub>P(S)SH/TFA**Product after cleavage:**

Carboxamide

**Observations:**

- Stable against acids and bases.
- Compatible with Fmoc/tBu- and Boc/Bnz-SPPS strategy
- The reduction of the sulfoxides generates acid-labile sulfides.

**Cleavage:**

- 1) HgClO<sub>4</sub>, H<sub>5</sub>IO<sub>6</sub>
- 2) Photolysis (350 nm)

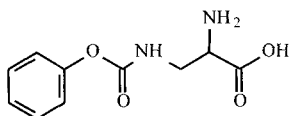
**Product after cleavage:**

- Carboxylic acid
- Alcohol

**Observations:**

The oxidation of the dithiane ring with HgClO<sub>4</sub> or H<sub>5</sub>IO<sub>6</sub> gives the benzoin derivative; the product is released from the resin by photolysis.

**3-N-Phenylloxycarboxy-2,3-diaminopropionic acid**  
Dpr(Phoc) linker (L-2,3-Diaminopropionic acid(phenylloxycarbonyl) [292, 293]



<b>Cleavage:</b> 1) Activation pH = 9 2) Release of product catalyzed by $\text{Ca}^{2+}$	<b>Product after cleavage:</b> – Carboxylic acid – Carboxamide
---	--

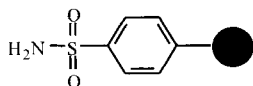
**Observations:**

- Compatible with Fmoc/tBu and Boc/Bnz SPPS strategy.
- By the addition of a base (PhONa or NaOH) a isocyanate is formed in the side chain. This isocyanate reacts with the amide group to the cyclic urea, which is cleaved under basic conditions.

Carrier: Aminofunctionalized-Polyacrylamide (Expansin)  
Aminofunctionalized-PS/PEG/DVB

**Pop- and Pop-Anchor** (see also 'Multifunctional anchors')

**Sulfonamide-PS/DVB**  
[294]

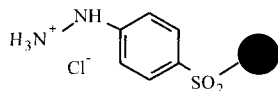


<b>Cleavage:</b> – NaOH/H <sub>2</sub> O – NH <sub>3</sub> /MeOH – Hydrazinolysis/Dioxan	<b>Product after cleavage:</b> – Carboxylic acid – Carboxamide – Hydrazide
---	---

**Observations:**

- Compatible with Fmoc/tBu- and Boc/Bnz-SPPS-strategy.
- Stable against hydrolysis.
- For cleavage, the linker is treated with diazomethane. The generated N-methylamide is cleaved by hydrolysis, aminolysis and hydrazinolysis.

**Sulfonylphenylhydrazine resin**  
[295]



<b>Cleavage:</b> $\text{O}_2 + \text{Cu(Py)}_2 + \text{H}_2\text{O}$	<b>Product after cleavage:</b> Carboxylic acid
---	---

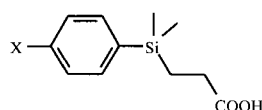
**Observations:**

- Compatible with Fmoc/tBu- and Boc/Bnz-SPPS-strategy.
- Stable against acids, bases and reduction.
- The linker is activated by oxidation of the hydrazide to the phenyldiimide with oxygen and a catalyst.
- Due to the mild conditions for this reaction, even groups sensitive to oxidation (e.g. sulfide) can be used.

Carrier: Aminofunctionalized Poly-2-hydroxyethylmethacrylate

## 5.7 Silicium-Anchor/Traceless-Anchor

Traceless linkers are designed in order to prevent the appearance of additional functional groups (traces) after cleavage. In most cases, fluorides (HF, TBAF, KF) are used as cleavage reagents, since they have a high affinity for Si-atom, thus cleaving the Si-C bond and releasing the product. Further cleavage reagents include acids ( $H^+$ ,  $I^+$ ,  $Br^+$ ) which release the product from the carrier through an *ipso*-substitution (at the Si-carrying aromatic C-atom of the anchor). Recently, the use of triazen anchors for the synthesis of substituted aromatic compounds was described [296]. Anchors cleaved by metathesis also leave no additional functional groups. They are mentioned in the section 'Allylfunctionalized anchors'.

**3-Arylsilyl-propionic acid** [297]

Carrier: Aminofunctionalized PS/DVB

<b>Cleavage:</b> TFA	<b>Product after cleavage:</b> Aryl derivative
-------------------------	---

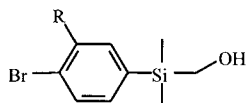
**Observations:**

- X = Br, Me<sub>3</sub>Sn.
- Cleavage is achieved via an anchimerically assisted proteodesilylation.

**Chemical reactions:**

Suzuki coupling.

**2-Bromo-4-(hydroxymethyl(dimethylsilyl))benzonitrile**  
or  
4-(Hydroxymethyl(dimethylsilyl))-1-bromobenzene



Carrier: Hydroxyphenoxymethyl-PS/DVB

<b>Cleavage:</b> CsF or TFA (110 °C)	<b>Product after cleavage:</b> Benzene derivative [298]
---	---

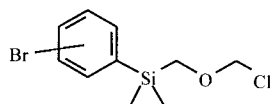
**Observations:**

- R = CN, H.
- Immobilization of the linker by a Mitsunobu reaction.
- The benzene ring remains part of the product.

**Chemical reaction:**

Reduction (imine → amine), N-substitution.

**Chloromethoxymethyl-dimethylsilyl-bromobenzene**  
[299]



Carrier: Wang resin

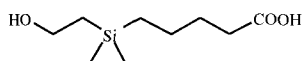
<b>Cleavage:</b> - ICl - Br <sub>2</sub> /pyridine - TFA	<b>Product after cleavage:</b> - Iodobenzene - Bromobenzene - Benzene derivative
---	---

**Observations:**

- Stable against NaOH, NaOMe, KF (80 °C, 24 h, DMF, H<sub>2</sub>O).
- Sometimes an electrophilic substitution on the benzene ring during the cleavage is observed.

**Chemical reaction:**

Suzuki reaction.

**5-[Dimethyl-(2-Hydroxyethyl)-silyl]-valeric acid [300]**

Carrier: Aminofunctionalized PS/DVB

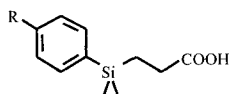
<b>Cleavage:</b> Acetic anhydride + BF <sub>3</sub> OEt <sub>2</sub>	<b>Product after cleavage:</b> Acetyl pyranose
--	---

**Observations:**

Introduction of various acyloxy groups at the anomeric center of pyranoses.

**Chemical reactions:**

Efficient O-benzoylation and O-propionylation of pyranose.

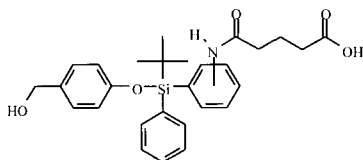
**3-[Dimethyl-(phenyl)-silyl]-propionic acid**

Carrier: BHA-PS/DVB

<b>Cleavage:</b> TFA (neat or vapor)	<b>Product after cleavage:</b> Unsubstituted aryl-moieties [301]
---	--

**Chemical reactions:**

Alkylation, acylation, Mitsunobu reaction.

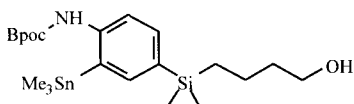
**Glutaric acid-mono-(3- or 4-) ((4-Hydroxymethyl)-phenoxy)-(t-butyl-phenyl)-silyl)-phenyl)-anilide**

Carrier: Aminomethyl-PS/DVB

<b>Cleavage:</b> TBAF in DMF/Thiophenol/ DIPEA	<b>Product after cleavage:</b> Carboxylic acid [302, 303]
---	---

**Observations:**

- Compatible with Fmoc/tBu SPPS strategy.
- Acid-labile.
- Difficult synthesis of the linker.

**4-(4-Hydroxybutyl-dimethylsilyl)-2-trimethylstannyl-N-Bpoc-aniline [304]**

Carrier: Aminomethyl-PS/DVB

<b>Cleavage:</b> TFA or HF	<b>Product after cleavage:</b> Benzene derivative
-------------------------------	--

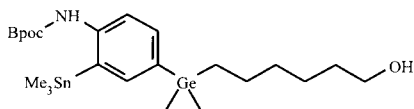
**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- The linker is stable against TFA if the benzene ring carries electronegative substituents. In this case the linker is cleaved with HF.

**Chemical reaction:**

Stille reaction, N-alkylation

**4-(6-Hydroxyhexyl-dimethylgermanyl)-2-trimethylstannyl-N-Bpoc-aniline** [305]



Carrier: Aminomethyl-PS/DVB

<b>Cleavage:</b> – TFA – Br <sub>2</sub>	<b>Product after cleavage:</b> – Benzene derivative – Bromobenzene derivative
--	---

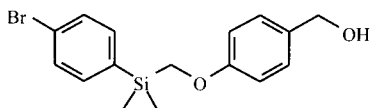
**Observations:**

Compatible with Fmoc/tBu SPPS strategy.

**Chemical reactions:**

Stille reaction, N-alkylation

**4-(4-(Hydroxymethyl)phenoxy)methyl)-dimethylsilyl)-bromobenzene**



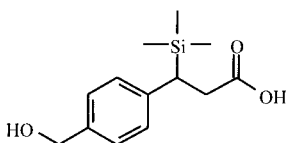
Carrier: Chloromethyl-PS/DVB

<b>Cleavage:</b> – TFA (25 °C) – CsF	<b>Product after cleavage:</b> Benzene derivative [298]
--	--

**Observations:**

- Cleavage of the linker at the Si-atom.
- Immobilization of the hydroxy group under Mitsunobu conditions.
- Benzene ring remains in the product.

**3-(4-Hydroxymethylphenyl)-3-(trimethylsilyl)-propionic acid**



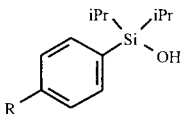
Carrier: Aminomethyl-PS/DVB

<b>Cleavage:</b> TBAF/DMF 5 min	<b>Product after cleavage:</b> Carboxylic acid [306, 307]
------------------------------------	--

**Observations:**

- Compatible with Fmoc/tBu SPPS strategy.
- Acid-labile.

**4-Methoxymethoxy-(or Bromo-)phenyl-diisopropyl-hydroxysilane**



Carrier: Hydroxymethyl-PS/DVB

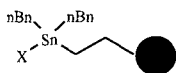
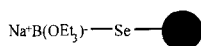
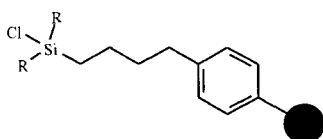
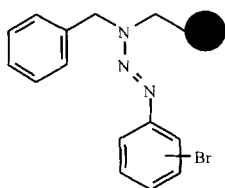
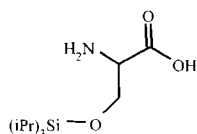
<b>Cleavage:</b> TBAF, DMF, 65 °C, 1 h	<b>Product after cleavage:</b> Benzene derivative [308, 309]
---	---

**Observations:**

- R = –OMOM, –Br.
- Depending on the electronic properties of the linker, the Si-O or the Si-C bond is cleaved.
- Linker is stable against 5% TFA and DBU.

**Chemical reaction:**

Aldol reaction, oxidation (alcohol → ketone), aromatic metalation (nBuLi), Pd(0)-mediated coupling reactions (Heck, Stille, Suzuki).

**PS/DVB-di-n-butyltinhydride**  
PBTH**Selenium-PS**  
[311]**Trialkylsilyl chloride – PS/DVB** [312]**Triazen-PS/DVB****O-Triisopropylsilyl-serine** [313]Carrier: Tentagel-S-NH<sub>2</sub>, PEG-PS/DVB

<b>Cleavage:</b> Intramolecular Suzuki-coupling	<b>Product after cleavage:</b> Substituted alkene [310]
--	--

**Observations:**

– x = H, Cl.

**Chemical reactions:**

Grignard addition, Mitsunobu reaction, oxidation.

<b>Cleavage:</b> HSnBu <sub>3</sub> , AIBN 90 °C	<b>Product after cleavage:</b> Alkane
---	--

**Observations:**

First building block is immobilized by alkylation of the SeCl or SeBr derivatives.

**Chemical reactions:**

Mitsunobu reaction

<b>Cleavage:</b> TFA	<b>Product after cleavage:</b> Alcohol Aromatic compound
-------------------------	--

**Observations:**

– R = Et, Ph, Me, iPr.

– Chemical reactions can be followed by the IR active Si-H stretch (2000–2200 cm<sup>-1</sup>) vibration.

– Aromatic residues are immobilized by their Li-metalated derivatives.

<b>Cleavage:</b> HCl/THF or H <sub>3</sub> PO <sub>2</sub> /Cl <sub>2</sub> HCCO <sub>2</sub> H	<b>Product after cleavage:</b> Substituted aromatic compound [296]
---	---

**Observations:**

– Benzylamine is immobilized on chloromethyl-PS/DVB. The treatment of the secondary amine with diazonium salts yields the triazen anchor.

– After cleavage, the anchor can be recycled.

**Chemical reactions:**

Heck reaction.

<b>Cleavage:</b> TBAF	<b>Product after cleavage:</b> Phenol
--------------------------	--

**Observations:**

– Phenol derivatives are attached to the linker via a carbamate bond.

– Stable to acidic (TFA) and basic conditions.

**Chemical reactions:**

Pictet–Spengler cyclization, acylation, Knoevenagel condensation.

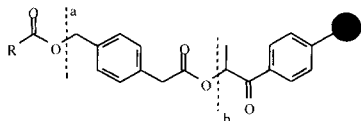
## 5.8 Miscellaneous

## 5.8.2 Multifunctional Anchors

This group of anchors is characterized by the fact that the anchoring to the carrier offers multiple cleavage sites. Therefore, various end-groups can be generated by the use of different cleavage strategies, resulting in diverse final products.

**Acyl-2-[(oxymethyl)phenylacetoxy]-propionyl-PS/DVB**

Pop resin  
[314, 315]



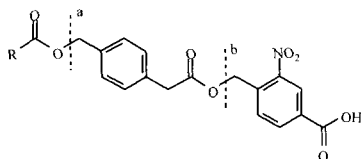
Cleavage:	Product after cleavage:
a) Strong acid or hydrogenolysis	a) Carboxylic acid
a) TMG	a) Carboxylic acid
b) Photolysis or thiophenoxide-ions	b) O-Methylphenylacetic acid ester
b) Photolysis + nBu <sub>4</sub> NCN	b) Carboxylic acid

**Observations:**

– Compatible with Boc/Bnz SPPS strategy.

**[Acyl-4-(4-oxymethyl)phenylacetoxy-methyl]-3-nitrobenzoic acid**

Pon resin  
[314, 315]



Cleavage:	Product after cleavage:
a) Acid or hydrogenolysis	a) Carboxylic acid
b) Photolysis or thiophenoxide-ions	b) O-Methylphenylacetic acid ester
b) Photolysis + Tetra-nBuNH <sub>4</sub> CN	b) Carboxylic acid

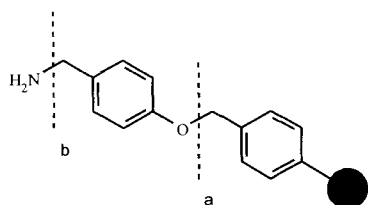
**Observations:**

– Compatible with Boc/Bnz SPPS strategy.

Carrier: Aminomethyl-PS

**p-Benzyloxybenzylamine – PS/DVB**

BOBA [316]



Cleavage:	Product after cleavage:
a) TMSOTf	a) Phenol
b) DQQ	b) Amine

**Observations:**

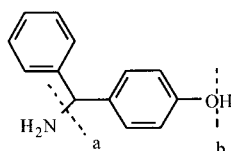
The use of stronger acids (TFA, TfOH) for cleavage leads to decomposition.

**Chemical reactions:**

Imine formation, Mannich reaction.

**p-Hydroxybenzhydrylamine**

[34, 317]

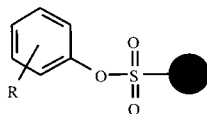


Carrier: – Aminofunctionalized-PS/DVB  
 – Carboxyfunctionalized-PS/DVB

<b>Cleavage:</b> a) HF b) TFA/HF/NH <sub>2</sub> NH <sub>2</sub> b) NaOH	<b>Product after cleavage:</b> a) Carboxylic acid b) p-Hydroxybenzhydrylcarboxamide
---	---

**Observations:**

- Compatible with Boc/Bnz SPPS strategy.
- In order to immobilize the linker, the hydroxy group is either treated directly with a carboxy functionalized resin or is coupled to a dicarboxylic acid spacer on aminomethyl resin.

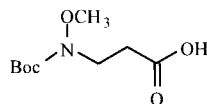
**5.8.2 Anchors Cleaved by Reduction****Arylsulfonates [318]**

Carrier: Dowex 50W ion-exchange resin

<b>Cleavage:</b> Formic acid, Pd(0), 110–140 °C	<b>Product after cleavage:</b> Benzoate ester Benzamides
---	--

**Observations:**

- In contrast to Si and Ge linkages (traceless linkers) of the aryl ring to the resin where electron-rich aromatic ring systems favor rapid cleavage, the reductive cleavage of arylsulfonates works well with electron-poor aromatic rings.
- R = ester or amide.

**N-tert. Butyloxycarbonyl-N-methoxypropionic acid**

Carrier: MBHA-PS/DVB

<b>Cleavage:</b> LiAlH <sub>4</sub>	<b>Product after cleavage:</b> Aldehyde [319]
--	--

**Observations:**

- Stable against HF.
- Compatible with Fmoc/tBu- and Boc/Bnz-SPPS-strategy.

**4-Hydroxy-2-methoxybenzyl alcohol**

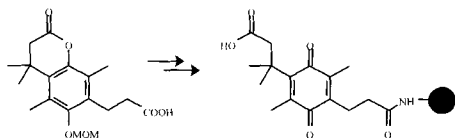
see Section 5.3. Anchors cleaved by nucleophiles.

**4-Hydroxymethylphenol**

Wang resin; see Section 5.3. Anchors cleaved by nucleophiles.



**6-Hydroxy-7-(3-propionic acid)-4,4,5,8-tetramethylhydrocoumarin**  
Redox-Sensitive Linker [320]



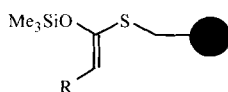
Carrier: Aminofunctionalized PS/DVB

<b>Cleavage:</b> Na <sub>2</sub> S <sub>2</sub> O <sub>4</sub> in H <sub>2</sub> O/THF	<b>Product after cleavage:</b> Alcohol
---	---

**Observations:**

- The immobilized quinone is obtained after NBS oxidation of the coumarin moiety.
- After reduction with Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub>, the product is released by lactonization of the linker.
- Compatible with Boc/Bzl SPPS strategy.

**Trimethylsilylenothioether-PS/DVB**



<b>Cleavage:</b> LiBH <sub>4</sub>	<b>Product after cleavage:</b> Alcohol
---------------------------------------	---

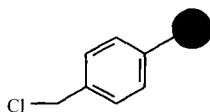
**Observations:**

- Immobilization of imines by forming a thioester bonding [321].
- Reduction of the thioester releases the alcohol.
- The same anchor group is realized on a 5-(4'-chloromethylphenyl)-pentyl-PS/DVB (CMPP resin). R is introduced as R-COSK. Lewis acids can be used during synthesis without decomposition or deactivation (*e.g.* Sc(OTf)<sub>3</sub>-catalyzed imino aldol reaction) [322]

### 5.8.3 Hydrogenolysis-Labile Anchors

This particularly mild and elegant cleavage method is not often used in solid-phase chemistry, although it is an orthogonal method potentially useful with both Fmoc/tBu and Boc/Bnz chemistries. This is due to the amount of work required for a single cleavage and the impurities generated by the catalysis, which have to be removed or washed from the products.

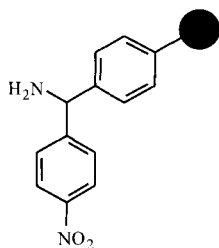
**Chloromethyl-PS/DVB**  
[323–328]



<b>Cleavage:</b> H <sub>2</sub> /Pd(OAc) <sub>2</sub> Ammonium formate/ Pd(OAc) <sub>2</sub>	<b>Product after cleavage:</b> Carboxylic acid
---	---

**Observations:**

- Compatible with Boc/Bnz- and Fmoc/tBu-SPPS-strategy.
- See also acid- and base-labile anchors.

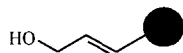
**p-Nitrobenzhydrylamine-PS/DVB**

<b>Cleavage:</b> H <sub>2</sub> or dicyclohexa-1,4-diene/Pd(OAc) <sub>2</sub>	<b>Product after cleavage:</b> Carboxamide [329]
--	---

**Observations:**

- Compatible with Fmoc/tBu SPPS strategy.

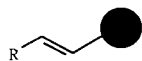
## 5.8.4 Anchors Cleaved by Oxidation

**Hydroxymethylvinyl-PS/DVB**

<b>Cleavage:</b> Ozonolysis	<b>Product after cleavage:</b> Aldehyde [49]
--------------------------------	---

**Chemical reaction:**

- Oligosaccharide synthesis.

**Olefinic linker [330, 331]**

<b>Cleavage:</b> Ozonolysis	<b>Product after cleavage:</b> Aldehyde
--------------------------------	--

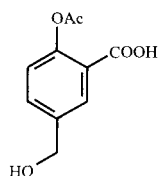
**Observations:**

- Stable to Boc/Bnz and Fmoc/tBu SPPS strategies
- Aldehyde is immobilized in a Wittig or Wittig-Horner reaction (phosphorane or diethylphosphonoacetamide on the polymer support) giving the unsaturated bond.

**Propane-1,3-dithiol** (see Anchors cleaved by Nucleophiles)

## 5.8.5 Enzymatically Cleavable Anchors

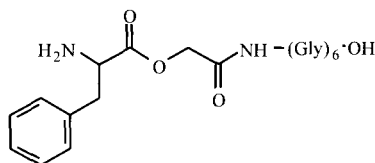
This somewhat exotic class of anchors includes only a small number of examples. Although it also offers an orthogonal cleavage technique, few applications are found in the literature.

**4-Acetyloxybenzyloxy** [332]Carrier: TentaGel S NH<sub>2</sub>

<b>Cleavage:</b> Lipase RB 001-05 pH 5.8, rt	<b>Product after cleavage:</b> Carboxylic acid Amide Alcohol
--	---

**Observations:**

- The enzyme converts the phenolic acetate to the corresponding phenolate, which then fragments.
- Very acid-labile structures can be cleaved selectively.
- Alcohols are attached to the linker as carbonates.

**Phenylalanine ester** [10]

Carrier: Aminopropyl silica gel

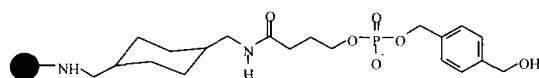
<b>Cleavage:</b> $\alpha$ -Chymotrypsin	<b>Product after cleavage:</b> Carboxylic acid
--	---

**Observations:**

- Due to the surface properties of the carrier, even large biomolecules have an access to the functional groups on the solid support.

**Chemical reaction:**

- Enzymatic cleavage.
- Enzymatic synthesis of oligosaccharides with  $\beta$ -1,4-galactosyl-transferase and  $\alpha$ -2,3-sialyltransferase.

**Phosphoric diester** [333]

Carrier: Polyacrylamide resin

<b>Cleavage:</b> Phosphodiesterase	<b>Product after cleavage:</b> Carboxylic acid
---------------------------------------	---

**Observations:**

- Compatible with Fmoc/tBu-SPPS-strategy.
- Stable against TFA.
- 1-N-(4'-hydroxy)-butane-carbonylamino-methyl-4-aminomethyl-cyclohexane is immobilized on a carrier (Pep-syn K) through the amino group; after phosphorylation, 1,4-bis(hydroxy-methyl)-benzene is coupled to the linker.

## 5.9 Acknowledgements

I would like to thank Professor G. Jung for advice and discussions and Dr. R. Jack for his great help in preparing this manuscript.

## 5.10 References

- [1] R. R. Merrifield, *J. Am. Chem. Soc.* **85** (1963), 2149–2154.
- [2] R. R. Merrifield, *Angew. Chem.* **97** (1985), 801–812; *Angew. Chem. Int. Ed. Engl.* **24** (1985), 799–810.
- [3] M. R. Pavia, T. K. Sawyer, W. H. Moos, *Bioorganic and Medical Chemistry Letters* **3** (1993), 387–396.
- [4] G. Jung, A. G. Beck-Sickinger, *Angew. Chem.* **104** (1992), 375–400; *Angew. Chem. Int. Ed. Engl.* **31** (1992), 367–383.
- [5] G. B. Fields, R. L. Noble, *Int. J. Pept. Protein Res.* **35** (1990), 161–214.
- [6] S. L. Beaucage, R. P. Iyer, *Tetrahedron* **48** (1992), 2223–2311.
- [7] P. W. Davis, D. Hudson, M. Lyttle, N. D. Sinaha, N. Usman, M. Weiss, P. Wright, *Innovations and Perspectives in Solid Phase Synthesis*, Intercept, Andover 1992, p. 63.
- [8] F. X. Montsera, A. Grandas, R. Eritja, E. Pendroso, *Tetrahedron* **50** (1994), 2617–2622.
- [9] S. J. Danishefsky, K. F. McClure, J. T. Randolph, R. B. Ruggeri, *Science* **260** (1993), 1307–1309.
- [10] M. Schuster, P. Wang, J. C. Paulson, C.-H. Wong, *J. Am. Chem. Soc.* **116** (1994), 1135–1136.
- [11] P. Hermkens, H. H. H. C. J. Ottenheijm, D. Rees, *Tetrahedron* **52** (1996), 4527–4554.
- [12] P. H. H. Hermkens, H. C. J. Ottenheijm, D. C. Rees, *Tetrahedron* **53** (1997), 5643–5678.
- [13] S. Booth, P. H. H. Hermkens, H. C. J. Ottenheijm, D. C. Rees, *Tetrahedron* **54** (1998), 15385–15443.
- [14] S. Hobbs DeWitt, J. S. Kiely, C. J. Stankovic, M. C. Schroeder, D. M. Reynolds Cody, M. R. Pavia, *Proc. Natl. Acad. Sci. USA* **90** (1993), 6909–6913.
- [15] J. I. Crowley, H. Rapoport, *J. Am. Chem. Soc.* **92** (1970), 6363–6365.
- [16] M. Cardno, M. Bradley, *Tetrahedron Lett.* **37** (1996), 135–138.
- [17] M. Lebl, M. Pátek, P. Kocis, V. Krchnák, V. J. Hruby, S. E. Salmon, K. S. Lam, *Int. J. Pept. Protein Res.* **41** (1993), 201–203.
- [18] C. K. Jayawickreme, G. F. Graminski, H. M. Quillan, M. R. Lerner, *Proc. Natl. Acad. Sci. USA* **91** (1994), 1614–1618.
- [19] M. Winter, in G. Jung (Ed.): *Combinatorial Peptide and Nonpeptide Libraries*, VCH-Wiley, Weinheim 1996, p. 465–510.
- [20] T. A. Keating, R. W. Armstrong, *J. Am. Chem. Soc.* **118** (1996), 2574–2583.
- [21] R. J. Booth, J. C. Hodges, *Acc. Chem. Res.* **32** (1999), 18–26.
- [22] F. Albericio, G. Barany, *Int. J. Pept. Protein Res.* **30** (1987), 206–216.
- [23] F. Albericio, N. Kneib-Cordonier, S. Biancalana, L. Gera, R. I. Masada, D. Hudson, G. Barany, *J. Org. Chem.* **55** (1990), 3730–3743.
- [24] N. A. Solé, S. A. Kates, F. Albericio, G. Barany, *Solid Phase Peptide Synthesis* (1993), 105–110.
- [25] H. Dürr, A. G. Beck-Sickinger, G. Schnorrenberg, W. Rapp, G. Jung, *Int. J. Pept. Protein Res.* **38** (1991), 146–153.
- [26] S. A. Kates, B. G. de la Torre, R. Eritja, F. Albericio, *Tetrahedron Lett.* **35** (1994), 1033–1034.
- [27] M. Hiroshige, J. R. Hauske, P. Zhou, *Tetrahedron Lett.* **36** (1995), 4567–4570.
- [28] N. S. Gray, S. Kwon, P. G. Schutlz, *Tetrahedron Lett.* **38** (1997), 1161–1164.
- [29] J. Vagner, V. Krchnak, M. Lebl, G. Barany, *Coll. Czech. Chem. Com.* **61** (1996), 1697–1702.
- [30] M. Kempe, G. Barany, *J. Am. Chem. Soc.* **118** (1996), 7083–7093.
- [31] P. Pietta, O. Brenna, *J. Org. Chem.* **40** (1975), 2995–2996.
- [32] P. G. Pietta, G. R. Marshall, *J. Chem. Soc.* (1970), 650–651.
- [33] P. M. Worster, C. R. McArthur, C. C. Leznoff, *Angew. Chem.* **91** (1979), 255–258; *Angew. Chem. Int. Ed. Engl.* **8** (1979), 221–224.
- [34] J. Tam, R. D. DiMarchi, R. B. Merrifield, *Tetrahedron Lett.* **22** (1981), 2851–2854.
- [35] A. Hirao, S. Itsuno, I. Hattori, K. Yamaguchi, S. Nakahama, N. Yamazaki, *J. Chem. Soc., Chem. Commun.* (1983), 25–26.
- [36] D. H. Coy, S. J. Hocart, Y. Sasaki, *Tetrahedron* **44** (1988), 835–841.
- [37] E. R. Marsh, H. K. Smith, *Molecular Diversity* **2** (1997), 165–170.
- [38] K. Ngu, D. V. Patel, *Tetrahedron Lett.* **38** (1997), 973–976.
- [39] Y.-T. Wu, H.-P. Hsieh, C.-Y. Wu, H.-M. Yu, S.-T. Chen, K.-T. Wang, *Tetrahedron Lett.* **39** (1998), 1783–1784.
- [40] E. Bayer, in R. S. Hodges, A. Smith (Eds.): *Peptides Chemistry, Structure and Biology*, ESCOM, Leiden 1994, p. 156–158.

- [41] Novabiochem, *The Combinatorial Chemistry Catalog*, March 1998.
- [42] J. M. J. Fréchet, K. E. Haque, *Tetrahedron Lett.* 35 (1975), 3055–3056.
- [43] T. M. Fyles, C. C. Leznoff, *Can. J. Chem.* 54 (1976), 935–942.
- [44] A. van Vlient, R. H. P. H. Smulders, B. H. Rietman, A.-M. W. Klink, D. T. S. Rijkers, I. F. Eggen, G. Van de Werken, G. I. Tesser, in C. H. Schneider, A. N. Eberle (Eds.): *Peptides 1992*, ESCOM, Leiden 1992, p. 279–280.
- [45] W. D. F. Meutermaans, P. F. Alewood, *Tetrahedron Lett.* 36 (1995), 7709–7712.
- [46] M. J. Kurth, L. A. Ahlberg Randall, C. Chen, C. Melander, R. B. Miller, *J. Org. Chem.* 59 (1994), 5862–5864.
- [47] D. Stones, D. J. Miller, M. W. Beaton, T. J. Rutherford, D. Gani, *Tetrahedron Lett.* 39 (1998), 4875–4878.
- [48] R. Frenette, R. W. Friesen, *Tetrahedron Lett.* 35 (1994), 9177–9180.
- [49] J. M. Fréchet, C. Schuerch, *J. Am. Chem. Soc.* 93 (1971), 492–496.
- [50] X. Beebe, N. E. Schore, M. J. Kurth, *J. Am. Chem. Soc.* 114 (1992), 10061–10062.
- [51] S. Hobbs DeWitt, A. W. Cyarnik, *Acc. Chem. Res.* 29 (1996), 114–122.
- [52] H. M. Relles, R. W. Schlunz, *J. Am. Chem. Soc.* 96 (1974), 6469–6475.
- [53] E. Giralt, J. Rizo, E. Pedroso, *Tetrahedron* 40 (1984), 4141–4152.
- [54] R. Bollhagen, M. Schmiedberger, K. Barlos, E. Grell, *J. Chem. Soc., Chem. Commun.* (1994), 2559–2560.
- [55] H. Wenschuh, L. A. Carpino, F. Albericio, E. Krause, M. Beyermann, M. Bienert, in H. Maia (Ed.): *Peptides 1994*, ESCOM, Leiden 1994, p. 287–288.
- [56] K. Barlos, D. Gatos, I. Kallitsis, D. Papaioannou, P. Sotiriou, *Liebigs Ann. Chem.* (1988), 1079–1081.
- [57] V. Krcňák, Z. Flegelova, A. S. Weichsel, M. Lebl, *Tetrahedron Lett.* 36 (1995), 6193–6396.
- [58] J. J. McNally, M. A. Youngman, S. L. Dax, *Tetrahedron Lett.* 39 (1998), 967–970.
- [59] K. Barlos, D. Gatos, G. Papaphotiu, W. Schäfer, Y. Wenqing, *Tetrahedron Lett.* 30 (1989), 3947–3950.
- [60] K. Barlos, D. Gatos, G. Papaphotiu, P. Sotiriou, Y. Wenqing, W. Schäfer, *Tetrahedron Lett.* 30 (1989), 3943–3946.
- [61] C. Chen, L. A. Ahlberg Randall, R. B. Miller, A. D. Jones, M. J. Kurth, *J. Am. Chem. Soc.* 116 (1994), 2661–2662.
- [62] C. C. Leznoff, T. M. Fyles, *J. Chem. Soc., Chem. Commun.* (1976), 251–252.
- [63] S.-e. Yoo, J.-s. Seo, K.-y. Yi, Y.-d. Gong, *Tetrahedron Lett.* 38 (1997), 1203–1206.
- [64] H. Rink, *Tetrahedron Lett.* 28 (1987), 3787–3790.
- [65] G. Breipohl, J. Knolle, W. Stüber, *Int. J. Pept. Protein Res.* 34 (1989), 262–267.
- [66] S. Funakoshi, E. Murayam, L. Guo, N. Fuji, H. Yajima, *J. Chem. Soc., Chem. Commun.* (1988), 382–384.
- [67] T. A. Rano, K. T. Chapman, *Tetrahedron Lett.* 36 (1995), 3789–3792.
- [68] M. Meldal, F.-I. Auzanneau, O. Hindsgaul, M. M. Palcic, *J. Chem. Soc.* (1994), 1849–1850.
- [69] A. M. Bray, D. S. Chiefari, R. M. Valerio, N. J. Maeji, *Tetrahedron Lett.* 36 (1995), 5081–5084.
- [70] A. A. Virgilio, J. A. Ellman, *J. Am. Chem. Soc.* 116 (1994), 11580–11581.
- [71] A. V. Purandare, M. A. Poss, *Tetrahedron Lett.* 39 (1998), 935–938.
- [72] M. Larhed, G. Lindeberg, A. Hallberg, *Tetrahedron Lett.* 37 (1996), 8219–8222.
- [73] H. Rink, P. Sieber, in G. Jung, E. Bayer (Eds.): *Peptides 1988*, Walter de Gruyter & Company, Berlin 1989, p. 139–141.
- [74] K. A. Beaver, A. C. Siegmund, K. L. Spear, *Tetrahedron Lett.* 37 (1996), 1145–1148.
- [75] A. M. Strocker, T. A. Keating, P. A. Tempest, R. W. Armstrong, *Tetrahedron Lett.* 37 (1996), 1149–1152.
- [76] Y. Pei, W. H. Moos, *Tetrahedron Lett.* 35 (1994), 5825–5828.
- [77] D. A. Goff, R. N. Zuckermann, *J. Org. Chem.* 60 (1995), 5744–5745.
- [78] M. S. Deshpande, *Tetrahedron Lett.* 35 (1994), 5613–5614.
- [79] Y. D. Ward, V. Farina, *Tetrahedron Lett.* 37 (1996), 6993–6996.
- [80] W. Tegge, R. Frank, *J. Pept. Res.* 4 (1997), 355–362.
- [81] A. Svensson, K. E. Bergquist, T. Fex, J. Kihlberg, *Tetrahedron Lett.* 39 (1998), 7193–7196.
- [82] P. Sieber, *Tetrahedron Lett.* 28 (1987), 2107–2110.
- [83] W. C. Chan, P. D. White, J. Beythien, R. Steinauer, *J. Chem. Soc., Chem. Commun.* (1995), 589–592.
- [84] C. Somlai, P. Hegyes, L. Nyerges, B. Penke, W. Voelter, in C. H. Schneider, A. N. Eberle (Eds.): *Peptides 1992*, ESCOM, Leiden 1993, p. 198–199.

- [85] K. J. Jensen, J. Alsina, M. F. Songster, J. Vágner, F. Albericio, G. Barany, *J. Am. Chem. Soc.* **120** (1998), 5441–5452.
- [86] K. G. Estep, C. E. Neipp, L. M. Stephens Stramiello, M. D. Adam, M. P. Allen, S. Robinson, E. J. Roskamp, *J. Org. Chem.* **63** (1998), 5300–5301.
- [87] C. C. Leznoff, J. Y. Wong, *Can. J. Chem.* **51** (1973), 3756–3764.
- [88] J. Y. Wong, C. Manning, C. C. Leznoff, *Angew. Chem.* **86** (1974), 743–744; *Angew. Chem. Int. Ed. Engl.* **1974**, **13**, 735–736.
- [89] C. C. Leznoff, *Acc. Chem. Res.* **11** (1978), 327–333.
- [90] Z.-H. Xu, C. R. McArthur, C. C. Leznoff, *Can. J. Chem.* **61** (1983), 1405–1409.
- [91] S. Chamoin, S. Houldsworth, C. G. Kruse, W. I. Bakker, V. Snieckus, *Tetrahedron Lett.* **39** (1998), 4179–4172.
- [92] B. Henkel, E. Bayer, *Tetrahedron Lett.* **39** (1998), 9401–9402.
- [93] B. Henkel, W. Zeng, E. Bayer, *Tetrahedron Lett.* **38** (1997), 3511–3512.
- [94] K. H. Bleicher, J. R. Warcing, *Tetrahedron Lett.* **39** (1998), 4591–4594.
- [95] M. Mergler, R. Nyfeler, R. Tanner, J. Gosteli, P. Grogg, *Tetrahedron Lett.* **29** (1988), 4009–4012.
- [96] M. Mergler, R. Tanner, J. Gosteli, P. Grogg, *Tetrahedron Lett.* **29** (1988), 4005–4008.
- [97] K. Vetter, D. Tumelty, S. K. Singh, M. A. Gallop, *Angew. Chem.* **107** (1995), 94–97; *Angew. Chem. Int. Ed. Engl.* **34** (1996), 60–63.
- [98] G. Shapiro, D. Buechler, *Tetrahedron Lett.* **35** (1994), 5421–5424.
- [99] I. R. Marsh, H. Smith, M. Bradley, *J. Chem. Soc.* (1996), 941–942.
- [100] L. A. Thompson, J. A. Ellman, *Tetrahedron Lett.* **35** (1994), 9333–9336.
- [101] G. Liu, J. A. Ellman, *J. Org. Chem.* **60** (1995), 7712–7713.
- [102] E. K. Kick, J. A. Ellman, *J. Med. Chem.* **38** (1995), 1427–1430.
- [103] G. Wess, K. Bock, H. Kleine, M. Kurz, W. Guba, H. Hemmerle, E. Lopez-Calle, K.-H. Baringhaus, H. Glombik, a. Enhsen, W. Kramer, *Angew. Chem.* **108** (1996), 2363–2366; *Angew. Chem. Int. Ed. Engl.* **35** (1996), 2222–2224.
- [104] A. L. Smith, G. I. Stevenson, C. J. Swain, J. L. Castro, *Tetrahedron Lett.* **39** (1998), 8317–8320.
- [105] F. Albericio, G. Barany, *Tetrahedron Lett.* **32** (1991), 1015–1018.
- [106] A. Flörsheimer, B. Riniker, in E. Giralt, D. Andreu (Eds.): *Peptides 1990*, ESCOM, Leiden 1991, p. 131–133.
- [107] B. Kamber, B. Riniker, in J. A. Smith, J. E. Rivier (Eds.): *Peptides Chemistry Structure and Biology*, ESCOM, Leiden 1992, p. 525–526.
- [108] P. Sieber, *Tetrahedron Lett.* **28** (1987), 61476150.
- [109] L. S. Richter, T. R. Gadek, *Tetrahedron Lett.* **35** (1994), 4705–4706.
- [110] R. C. Sheppard, B. J. Williams, *J. Chem. Soc., Chem. Commun.* (1982), 587–589.
- [111] R. Sheppard, C., B. J. Williams, *Int. J. Pept. Protein Res.* **20** (1982), 451–454.
- [112] S.-S. Wang, *J. Am. Chem. Soc.* **95** (1973), 1328–1333.
- [113] G.-S. Lu, S. Mojsov, J. P. Ram, R. B. Merrifield, *J. Org. Chem.* **46** (1981), 3433–3436.
- [114] J. R. Hauske, P. Dorff, *Tetrahedron Lett.* **36** (1995), 1589–1592.
- [115] J. Alsina, C. Chiva, M. Ortiz, F. Rabanal, E. Giralt, F. Albericio, *Tetrahedron Lett.* **38** (1997), 883–886.
- [116] J. Alsina, F. Rabanal, C. Chiva, E. Giralt, F. Albericio, *Tetrahedron* **54** (1998), 10125–10152.
- [117] S. Hanessian, F. Xie, *Tetrahedron Lett.* **8** (1998), 733–740.
- [118] D. R. Barn, J. R. Morphy, D. C. Rees, *Tetrahedron Lett.* **37** (1996), 3213–3216.
- [119] J. A. Josey, C. A. Tarlton, C. E. Payne, *Tetrahedron Lett.* **39** (1998), 5899–5902.
- [120] G. Shapiro, R. Swoboda, U. Stauss, *Tetrahedron Lett.* **35** (1994), 869–872.
- [121] I. Ojima, C.-Y. Tsai, Z. Zhang, *Tetrahedron Lett.* **35** (1994), 5785–5788.
- [122] M. Hiroshige, J. R. Hauske, P. Zhou, *J. Am. Chem. Soc.* **117** (1995), 11590–11591.
- [123] P. Wipf, A. Cunningham, *Tetrahedron Lett.* **36** (1995), 7819–7822.
- [124] M. P. Sibi, S. V. Chandramouli, *Tetrahedron Lett.* **38** (1997), 8929–8932.
- [125] P. Roussel, M. Bradley, *Tetrahedron Lett.* **38** (1997), 4861–4864.
- [126] F. Albericio, G. Barany, *Int. J. Pept. Protein Res.* **23** (1984), 342–349.
- [127] B. A. Bunin, M. J. Plunkett, J. A. Ellman, *Proc. Natl. Acad. Sci. USA* **91** (1994), 4708–4712.
- [128] B. A. Bunin, J. A. Ellman, *J. Am. Chem. Soc.* **114** (1992), 10997–10998.
- [129] M. F. Plunkett, J. A. Ellman, *J. Am. Chem. Soc.* **117** (1995), 3306–3307.
- [130] A. R. Mitchell, B. W. Erickson, M. N. Ryabtsev, R. S. Hodges, R. B. Merrifield, *J. Am. Chem. Soc.* **98** (1976), 7357–7362.

- [131] A. R. Mitchell, S. B. H. Kent, M. Engelhard, R. B. Merrifield, *J. Org. Chem.* **43** (1978), 2845–2852.
- [132] A. van Vlient, D. T. S. Rijkers, G. I. Tesser, in H. Maia (Ed.): *Peptides 1994*, ESCOM, Leiden 1994, p. 267–268.
- [133] A. van Vlient, B. H. Rietman, S. C. F. Karkdijk, P. J. H. M. Adams, G. I. Tesser, in R. S. Hodges, J. A. Smith (Eds.): *Peptides Chemistry, Structure and Biology*, ESCOM, Leiden 1994, p. 151–152.
- [134] L. E. Canne, S. M. Walker, S. B. H. Kent, *Tetrahedron Lett.* **36** (1995), 1217–1220.
- [135] P. Hodge, J. Waterhouse, *J. Chem. Soc., Perkin Trans 1* (1983), 2319–2323.
- [136] R. C. Orlowski, R. Walter, D. Winkler, *J. Org. Chem.* **41** (1976), 3701–3705.
- [137] D. S. Brown, J. C. Revill, S. R. E., *Tetrahedron Lett.* **39** (1998), 8533–8536.
- [138] G. Breipohl, J. Knolle, R. Geiger, W. Stüber, in G. Jung, E. Bayer (Eds.): *Peptides 1988*, Walter de Gruyter & Company, Berlin 1989, p. 136–138.
- [139] G. Breipohl, J. Knolle, W. Stüber, *Tetrahedron Lett.* **28** (1987), 5651–5654.
- [140] A. S. Hernandez, J. C. Hodges, *J. Org. Chem.* **62** (1997), 3153–3157.
- [141] G. R. Matsueda, J. M. Stewart, *Peptides* **2** (1981), 45–50.
- [142] K. Kaljuste, A. Undén, *Int. J. Pept. Protein Res.* **42** (1993), 118–124.
- [143] A. Nefzi, J. M. Ostrech, H.-P. Meyer, R. A. Houghten, *Tetrahedron Lett.* **38** (1997), 931–934.
- [144] K. H. Bleicher, J. R. Wareing, *Tetrahedron Lett.* **39** (1998), 4587–4590.
- [145] R. L. Letsinger, M. J. Kornet, V. Mahadevan, D. M. Jerina, *J. Am. Chem. Soc.* **86** (1964), 5163–5165.
- [146] A. M. Felix, R. B. Merrifield, *J. Am. Chem. Soc.* **92** (1970), 1385–1391.
- [147] C. A. Metcalf III, C. B. Vu, R. Sundaramoorthi, V. A. Jacobsen, E. A. Laborde, J. Green, Y. Green, K. J. Macek, T. J. Merry, S. G. Pradeepan, M. Uesugi, V. M. Varkhedkar, D. A. Holt, *Tetrahedron Lett.* **39** (1998), 3435–3438.
- [148] N. W. Hird, K. Irie, K. Nagai, *Tetrahedron Lett.* **38** (1997), 7115–7118.
- [149] N. W. Hird, K. Irie, K. Nagai, *Tetrahedron Lett.* **38** (1997), 7111–7114.
- [150] N. J. Ede, A. M. Bray, *Tetrahedron Lett.* **38** (1997), 7119–7122.
- [151] T. Takahashi, S. Ebata, T. Doi, *Tetrahedron Lett.* **39** (1998), 1369–1372.
- [152] W. Bannwarth, J. Huebscher, R. Barner, *Bioorg. Med. Chem. Lett.* **6** (1996), 1525–1528.
- [153] J. R. Morphy, Z. Rankovic, D. C. Rees, *Tetrahedron Lett.* **37** (1996), 3209–3212.
- [154] A. R. Brown, D. C. Rees, Z. Rankovic, J. R. Morphy, *J. Am. Chem. Soc.* **119** (1997), 3288–3295.
- [155] T. Mizoguchi, K. Shigezane, N. Takamura, *Chem. Pharm. Bull.* **18** (1970), 1465–1474.
- [156] F. Baleux, J. Daunis, R. Jaquier, *Tetrahedron Lett.* **25** (1984), 5893–5896.
- [157] B. Calas, J. Mery, J. Parello, A. Cave, *Tetrahedron* **41** (1985), 5331–5339.
- [158] F. Baleux, B. Calas, J. Mery, *Int. J. Pept. Protein Res.* **28** (1986), 22–28.
- [159] R. E. Furlán, R. G. Mata, O. A. Mascaretti, *J. Chem. Soc., Perkin Trans 1* (1998), 355–358.
- [160] R. L. Halcomb, H. Huang, C.-H. Wong, *J. Am. Chem. Soc.* **116** (1994), 11315–11322.
- [161] D. Bellof, M. Mutter, *Chimia* **39** (1985), 317–320.
- [162] N. A. Abraham, G. Fazal, J. M. Ferland, S. Rakhit, J. Gauthier, *Tetrahedron Lett.* **32** (1991), 577–580.
- [163] S. S. Wang, *J. Org. Chem.* **41** (1976), 3258–3261.
- [164] C. Birr, M. Wengert-Müller, A. Buku, in M. Goodman, J. Meienhofer (Eds.): *Peptides Chemistry, Structure and Biology*, Halstead Press, New York 1977, p. 510–513.
- [165] J. P. Tam, W. F. Cunningham-Rundles, B. W. Erickson, R. B. Merrifield, *Tetrahedron Lett.* (1977), 4001–4004.
- [166] T. Mizoguchi, K. Shigezane, N. Takamura, *Chem. Pharm. Bull.* **17** (1969), 411–412.
- [167] F. S. Tjoeng, J. P. Tam, R. B. Merrifield, *Int. J. Pept. Protein Res.* **14** (1979), 262–274.
- [168] S. Mazur, P. Jayalekshmy, *J. Am. Chem. Soc.* **101** (1979), 677–682.
- [169] T. M. Fyles, C. C. Leznoff, J. Weatherston, *Can. J. Chem.* **56** (1978), 1031–1041.
- [170] M. A. Barton, R. U. Lemieux, J. Y. Savoie, *J. Am. Chem. Soc.* **95** (1973), 4501–4506.
- [171] G. R. Marshall, R. B. Merrifield, *Biochemistry* **4** (1965), 2394–2401.
- [172] M. K. Anwer, A. F. Spatola, *Tetrahedron Lett.* **33** (1992), 3121–3124.
- [173] H. C. Beyermann, H. Hindriks, E. W. B. de Leer, *J. Chem. Soc., Chem. Commun.* (1968), 1688.
- [174] G. L. Reddy, E. Bikshapathy, R. Nagaraj, *Tetrahedron Lett.* **26** (1985), 4257–4260.
- [175] G. J. Moore, D. McMaster, *Int. J. Pept. Protein Res.* **11** (1978), 140–148.
- [176] G. J. Moore, Y. C. Kwok, *Can. J. Biochem.* **58** (1980), 641–643.
- [177] S. R. Pulley, L. S. Hegedus, *J. Am. Chem. Soc.* **115** (1993), 9037–9047.
- [178] L. Yang, L. Guo, *Tetrahedron Lett.* **37** (1996), 5041–5044.
- [179] N. Mensi, S. S. Isied, *J. Am. Chem. Soc.* **109** (1987), 7882–7884.

- [180] J. D. Glass, A. Talansky, Z. Grzonka, I. L. Scharz, R. Walter, *J. Am. Chem. Soc.* 96 (1974), 6476–6480.
- [181] C. Y. Cho, E. J. Moran, S. R. Cherry, J. C. Stephans, S. P. A. Fodor, C. L. Adams, A. Sundaram, J. W. Jacobs, P. G. Schultz, *Science* 261 (1993), 1303–1305.
- [182] P. S. Furth, M. S. Reitman, A. F. Cook, *Tetrahedron Lett.* 38 (1997), 5403–5406.
- [183] P. S. Furth, M. S. Reitman, R. Gentles, A. F. Cook, *Tetrahedron Lett.* 38 (1997), 6643–6646.
- [184] A. V. Purandare, S. Natarajan, *Tetrahedron Lett.* 38 (1997), 8777–8780.
- [185] T. Ruhland, H. Künzer, *Tetrahedron Lett.* 37 (1996), 2757–2760.
- [186] R. Schwyzer, E. Felder, P. Failli, *Helv. Chim. Acta* 67 (1984), 1316–1327.
- [187] S. B. Katti, P. K. Misra, W. Haq, K. B. Mathur, *J. Chem. Soc., Chem. Commun.* (1992), 843–844.
- [188] F. Albericio, E. Giralt, R. Eritja, *Tetrahedron Lett.* 32 (1991), 1515–1518.
- [189] F. Albericio, F. Robles, D. Fernandez-Forner, Y. Palom, C. Celma, E. Pedroso, E. Giralt, R. Eritja, in E. Giralt, D. Andreu (Eds.): *Peptides 1990*, ESCOM, Leiden 1991, p. 134–138.
- [190] D. S. Kemp, N. G. Galakatos, B. Bowen, K. Tan, *J. Org. Chem.* 51 (1986), 1829–1838.
- [191] D. S. Kemp, N. G. Galakatos, *J. Org. Chem.* 51 (1986), 1821–1829.
- [192] M. Mergler, R. Nyfeler, *Peptides Chemistry, Structure and Biology* 1992, p. 551.
- [193] R. G. Almquist, S. R. Kadambi, D. M. Yasuda, F. L. Weitl, W. E. Polgar, L. R. Toll, *Int. J. Pept. Protein Res.* 34 (1989), 455–462.
- [194] A. M. Bray, N. J. Maeji, A. G. Jhingran, R. M. Valerio, *Tetrahedron Lett.* 32 (1991), 6163–6166.
- [195] M. Mergler, R. Nyfeler, in C. H. Schneider, A. N. Eberle (Eds.): *Peptides 1992*, ESCOM, Leiden 1993, p. 177–178.
- [196] H. C. Brown, Y. M. Choi, S. Narasimhan, *Inorganic Chemistry* 20 (1981), 4454–4457.
- [197] D. Seebach, A. Thaler, D. Blaser, S. Y. Ko, *Helv. Chim. Acta* 74 (1991), 1102–1118.
- [198] B. A. Dressman, L. A. Spangle, S. W. Kaldor, *Tetrahedron Lett.* 37 (1996), 937–940.
- [199] L. Gouilleux, J.-A. Fehrentz, F. Winternitz, J. Martinez, *Tetrahedron Lett.* 37 (1996), 7031–7034.
- [200] C. Sylvain, A. Wagner, C. Mioskowski, *Tetrahedron Lett.* 38 (1997), 1043–1044.
- [201] S. Chamoin, S. Houldsworth, V. Snieckus, *Tetrahedron Lett.* 39 (1998), 4175–4178.
- [202] E. Atherton, R. C. Sheppard, *Solid phase peptide synthesis; a practical approach*, IRL Press 1989, p. 152–154.
- [203] J. M. Stewart, J. D. Young, in Pierce (Ed.): *Solid phase peptide synthesis, 2nd Ed.* 1984, p. 91–92.
- [204] E. Atherton, C. J. Logan, R. C. Sheppard, *J. Chem. Soc., Perkin Trans 1* (1981), 538–546.
- [205] S. M. Hutchins, K. T. Chapman, *Tetrahedron Lett.* 37 (1996), 4869–4872.
- [206] R. Engert-Volkmer, B. Hoffmann, J. Schneider-Mergener, *Tetrahedron Lett.* 38 (1997), 1029–1032.
- [207] M. Mutter, D. Bellhof, *Helv. Chim. Acta* 67 (1984), 2009–2016.
- [208] F. Rabanal, E. Giralt, F. Albericio, *Tetrahedron Lett.* 33 (1992), 1775–1778.
- [209] Y.-Z. Liu, S.-H. Ding, J.-Y. Chu, A. M. Felix, *Int. J. Pept. Protein Res.* 35 (1990), 95–98.
- [210] C. Y. Ho, M. J. Kukla, *Tetrahedron Lett.* 38 (1997), 2799–2802.
- [211] B. Raju, T. P. Kogan, *Tetrahedron Lett.* 38 (1997), 3373–3376.
- [212] J. K. Young, J. C. Nelson, J. S. Moore, *J. Am. Chem. Soc.* 116 (1994), 10841–10842.
- [213] P. Rivaille, J. P. Gautron, B. Castro, G. Milhaud, *Tetrahedron* 36 (1980), 3413–3419.
- [214] E. W. Baxter, J. D. Rueter, S. O. Nortey, A. B. Reitz, *Tetrahedron Lett.* 39 (1998), 979–982.
- [215] J. K. Rueter, S. O. Nortey, E. W. Baxter, G. C. Leo, A. B. Reitz, *Tetrahedron Lett.* 39 (1998), 975–978.
- [216] J. W. Apsimon, D. M. Dixit, *Synth. Commun.* 12 (1982), 113–116.
- [217] P. S. Nelson, S. Muthini, M. A. Kent, T. H. Smith, *Nucleosides & Nucleotides* 16 (1997), 1951–1559.
- [218] J. W. Apsimon, D. M. Dixit, *Can. J. Chem.* 60 (1982), 368–370.
- [219] M. A. Dorman, S. A. Noble, L. J. McBride, M. H. Caruthers, *Tetrahedron* 40 (1984), 95–102.
- [220] V. V. Krchňák, S. Weichsel, D. Cabel, Z. Flegelova, M. Lebl, *Molecular Diversity 1* (1995), 149–164.
- [221] G. I. Tesser, J. T. W. A. R. M. Buis, E. T. M. Wolters, E. G. A. M. Bothé-Helmes, *Tetrahedron* 32 (1976), 1069–1072.
- [222] J. T. W. A. R. M. Buis, G. I. Tesser, R. J. F. Nivard, *Tetrahedron* 32 (1976), 2321–2325.
- [223] P. P. Fantauzzi, K. M. Yager, *Tetrahedron Lett.* 39 (1998), 1291–1294.
- [224] K. Kaljuste, J. P. Tam, *Tetrahedron Lett.* 39 (1998), 9327–9330.
- [225] E. T. Kaiser, H. Milhara, G. A. Laforet, J. W. Kelly, L. Walters, F. A. Findeis, T. Sasaki, *Science* 243 (1989), 187–192.
- [226] G. Losse, A. Barth, K. Schatz, *Justus Liebigs Ann. Chem.* 677 (1964), 185–190.
- [227] G. Losse, K.-H. Hoffman, G. Hetzer, *Justus Liebigs Ann. Chem.* 684 (1965), 236–242.



- [228] W. F. DeGrado, E. T. Kaiser, *J. Org. Chem.* **45** (1980), 1295–1300.
- [229] M. A. Findeis, E. T. Kaiser, *J. Org. Chem.* **54** (1989), 3478–3482.
- [230] E. Atherton, D. L. J. Clive, R. C. Sheppard, *J. Am. Chem. Soc.* **97** (1975), 6584–6585.
- [231] Y. Zhou, P. O. P. Ts'ao, *Nucleic Acids Res.* **24** (1996), 2652–2659.
- [232] G. W. Kenner, *Angew. Chem.* **71** (1959), 741–742.
- [233] V. Yedidia, C. C. Leznoff, *Can. J. Chem.* **58** (1980), 1166–1150.
- [234] C. C. Leznoff, V. Yedidia, *Can. J. Chem.* **58** (1980), 287–290.
- [235] S. R. Waldvogel, W. Pfeleiderer, *Helv. Chim. Acta* **81** (1998), 46–58.
- [236] J. Rademann, R. R. Schmidt, *Tetrahedron Lett.* **37** (1996), 3989–3990.
- [237] J. A. Camarero, G. J. Cotton, A. Adeva, T. W. Muir, *J. Pept. Res.* **51** (1998), 303–316.
- [238] J. S. Vyle, N. H. Williams, J. A. Grasby, *Tetrahedron Lett.* **39** (1998), 7975–7978.
- [239] I. Hughes, *Tetrahedron Lett.* **37** (1996), 7595–7598.
- [240] D. H. Rich, S. K. Gurwara, *J. Am. Chem. Soc.* **97** (1975), 1575–1579.
- [241] R. P. Hammer, F. Albericio, L. Gera, G. Barany, *Int. J. Pept. Protein Res.* **36** (1990), 31–45.
- [242] A. Ajayaghosh, V. N. R. Pillai, *J. Org. Chem.* **55** (1990), 2826–2829.
- [243] S. M. Sternson, S. L. Schreiber, *Tetrahedron Lett.* **39** (1998), 7451–7454.
- [244] B. B. Brown, D. S. Wagner, H. M. Geysen, *Molecular Diversity I* (1995), 4–12.
- [245] F. S. Tjoeng, G. A. Heavner, *J. Org. Chem.* **48** (1983), 355–359.
- [246] J. C. Sheehan, K. Umezawa, *J. Org. Chem.* **38** (1973), 3771–3774.
- [247] W. Giralt, F. Albericio, E. Pedroso, C. Granier, J. van Rietschoten, *Tetrahedron* **38** (1982), 1193–1208.
- [248] E. Giralt, F. Albericio, D. Andreu, R. Eritja, P. Martin, E. Pedroso, *An. Quim.* **77C** (1981), 120–125.
- [249] E. Nicolás, J. Clemente, T. Ferrer, F. Albericio, E. Giralt, *Tetrahedron* **53** (1997), 3179–3194.
- [250] E. Bayer, M. Dengler, B. Hemmasi, *Int. J. Pept. Protein Res.* **25** (1985), 178–186.
- [251] V. N. R. Pillai, M. Renil, V. K. Haridasan, *Indian J. Chem.* **30B** (1991), 205–212.
- [252] A. Ajayaghosh, V. N. R. Pillai, *Tetrahedron* **44** (1988), 6661–6666.
- [253] D. H. Rich, S. K. Gurwara, *J. Chem. Soc., Chem. Commun.* (1973), 610–611.
- [254] F. S. Tjoeng, R. S. Hodges, *Tetrahedron Lett.* **15** (1979), 1273–1276.
- [255] J. Lee, R. E. Barrett, P. R. Bovy, *LIPS* **2** (1995), 253–258.
- [256] C. P. Holmes, D. G. Jones, *J. Org. Chem.* **60** (1995), 2318–2319.
- [257] C. P. Holmes, in P. T. P. Kaumaya, R. S. Hodges (Eds.): *Peptides Chemistry, Structure and Biology*, Mayflower Scientific, Kingswinford 1995, p. 44–45.
- [258] C. P. Holmes, *J. Org. Chem.* **62** (1997), 2370–2380.
- [259] S. J. Teague, *Tetrahedron Lett.* **37** (1996), 5751–5754.
- [260] D. J. Yoo, M. M. Greenberg, *J. Org. Chem.* **60** (1995), 3358.
- [261] A. Ajayaghosh, V. N. R. Pillai, *J. Org. Chem.* **52** (1987), 5714–5717.
- [262] I. Sucholeiki, *Tetrahedron Lett.* **35** (1994), 7307–7310.
- [263] A. Ajayaghosh, V. N. R. Pillai, *Tetrahedron Lett.* **36** (1995), 777–780.
- [264] C. Dellaquilla, J. L. Imbach, B. Rayner, *Tetrahedron Lett.* **38** (1997), 5289–5292.
- [265] U. Zehavi, A. Patchornik, *J. Am. Chem. Soc.* **95** (1973), 5673–5677.
- [266] S. Peukert, B. Biese, *J. Org. Chem.* **63** (1998), 9045–9051.
- [267] H. Kunz, B. Dombo, *Angew. Chem.* **100** (1988), 732–734; *Angew. Chem. Int. Ed. Engl.* **27** (1988), 711–713.
- [268] O. Dangles, F. Guibé, G. Balavoine, *J. Org. Chem.* **52** (1987), 4984–4993.
- [269] J.-U. Peters, S. Blechert, *SYNLETT* **4** (1997), 348–350.
- [270] K. C. Nicolaou, D. Vourloumis, T. Li, J. Pastor, N. Winssinger, Y. He, S. Nikovic, F. Sarabia, H. Vallberg, F. Roschangar, N. P. King, M. R. V. Finlay, P. Giannakakou, P. Verdier-Pinard, E. Hamel, *Angew. Chem.* **109** (1997), 2181–2187; *Angew. Chem. Int. Ed. Engl.* **36** (1997), 2097–2013.
- [271] O. Seitz, H. Kunz, *Angew. Chem.* **107** (1995), 901–904; *Angew. Chem. Int. Ed. Engl.* **34** (1995), 803–806.
- [272] O. Seitz, H. Kunz, *J. Org. Chem.* **62** (1997), 813–826.
- [273] F. Guibé, O. Dangles, G. Balavoine, *Tetrahedron Lett.* **30** (1989), 2641–2644.
- [274] K. Kaljuste, A. Undén, *Tetrahedron Lett.* **37** (1996), 3031–3034.
- [275] B. Blankemeyer-Menge, R. Frank, *Tetrahedron Lett.* **29** (1988), 5871–5874.
- [276] M. M. Greenberg, T. J. Matray, J. D. Kahl, D. J. Yoo, D. L. McMinn, *J. Org. Chem.* **63** (1998), 4062–4068.
- [277] L. M. Gayo, M. J. Suto, *Tetrahedron Lett.* **38** (1997), 211–214.

- [278] C. Hulme, J. Peng, G. Morton, J. M. Salvino, T. Herpin, R. Labaudiniere, *Tetrahedron Lett.* 39 (1998), 7227–7230.
- [279] B. J. Backes, J. A. Ellman, *J. Am. Chem. Soc.* 116 (1994), 11171–11172.
- [280] B. J. Backes, A. A. Virgilio, J. A. Ellman, *J. Am. Chem. Soc.* 118 (1996), 3055–3056.
- [281] C. García-Echeverría, *Tetrahedron Lett.* 38 (1997), 8933–8934.
- [282] Y. Kiso, T. Fukui, S. Tanaka, T. Kumura, K. Akaji, *Tetrahedron Lett.* 35 (1994), 3571–3574.
- [283] K. Fukase, K. Egusa, Y. Nakai, S. Kusumoto, *Molecular Diversity* 2 (1997), 182–188.
- [284] T. Wieland, J. Lewalter, C. Birr, *Justus Liebigs Ann. Chem.* 740 (1970), 31–47.
- [285] C. R. Milligton, R. Quarrell, G. Lowe, *Tetrahedron Lett.* 39 (1998), 7201–7204.
- [286] S. Hoffmann, R. Frank, *Tetrahedron Lett.* 35 (1994), 7763–7766.
- [287] G. Panke, R. Frank, *Tetrahedron Lett.* 38 (1998), 17–18.
- [288] P. Kocis, V. Krchnák, M. Lebl, *Tetrahedron Lett.* 34 (1993), 7251–7252.
- [289] D. L. Marshall, *J. Org. Chem.* 35 (1970), 867–868.
- [290] M. Pátek, M. Lebl, *Tetrahedron Lett.* 32 (1991), 3891–3894.
- [291] A. Routledge, C. Abell, S. Balasubramanian, *Tetrahedron Lett.* 38 (1997), 1227–1230.
- [292] R. Sola, S. Philippe, M.-L. David, R. Pascal, *J. Chem. Soc., Chem. Commun.* (1993), 1786–1788.
- [293] R. Pascal, R. Sola, *Tetrahedron Lett.* 38 (1997), 4549–4552.
- [294] G. W. Kenner, J. R. McDermott, R. C. Sheppard, *J. Chem. Soc., Chem. Commun.* (1971), 636–637.
- [295] A. N. Semenov, K. Y. Gordeev, *Int. J. Pept. Protein Res.* 45 (1995), 303–304.
- [296] S. Bräse, D. Enders, J. Köbberling, F. Avemaria, *Angew. Chem.* 110 (1998), 3614–3616.
- [297] N. D. Hone, S. G. Davies, N. J. Devereux, S. L. Taylor, A. D. Baxter, *Tetrahedron Lett.* 39 (1998), 897–900.
- [298] B. Chenera, J. A. Finkelstein, D. F. Veber, *J. Am. Chem. Soc.* 117 (1995), 11999–12000.
- [299] Y. Han, S. D. Walker, R. N. Young, *Tetrahedron Lett.* 37 (1996), 2703–2706.
- [300] D. Weigelt, G. Magnusson, *Tetrahedron Lett.* 39 (1998), 2839–2842.
- [301] K. A. Newlander, B. Chenera, D. F. Veber, N. X. F. Yim, M. L. Moore, *J. Org. Chem.* 62 (1997), 6726–6732.
- [302] D. G. Mullen, G. Barany, *Tetrahedron Lett.* 28 (1987), 491–494.
- [303] D. G. Mullen, G. Barany, *J. Org. Chem.* 53 (1988), 5240–5248.
- [304] M. J. Plunkett, J. A. Ellman, *J. Org. Chem.* (1995), 6006–6007.
- [305] M. J. Plunkett, J. A. Ellman, *J. Org. Chem.* 62 (1997), 2885–2893.
- [306] R. Ramage, C. A. Barron, S. Bielecki, D. W. Thomas, *Tetrahedron Lett.* 28 (1987), 4105–4108.
- [307] R. Ramage, C. A. Barron, S. Bielecki, R. Holden, D. W. Thomas, *Tetrahedron* 48 (1992), 499–514.
- [308] T. L. Boehm, H. D. H. Showalter, *J. Org. Chem.* 61 (1996), 6498–6499.
- [309] *novabiochem. letters* 2 (1997), 3.
- [310] K. C. Nicolaou, N. Winssinger, J. Pastor, F. Murphy, *Angew. Chem.* 110 (1998), 2677–2680; *Angew. Chem. Int. Ed. Engl.* 37 (1998), 2534–2537.
- [311] R. Ruhland, K. Andersin, H. Pedersen, *J. Org. Chem.* 63 (1998), 9204–9211.
- [312] Y. Hu, J. A. Porco, J. W. Labadie, O. W. Gooding, B. M. Trost, *J. Org. Chem.* 63 (1998), 4518–4521.
- [313] Y.-L. Chou, M. M. Morrissey, R. Mohan, *Tetrahedron Lett.* 39 (1998), 757–760.
- [314] J. P. Tam, F. S. Tjoeng, R. B. Merrifield, *Tetrahedron Lett.* 51 (1979), 4935–4938.
- [315] J. P. Tam, F. S. Tjoeng, R. B. Merrifield, *J. Am. Chem. Soc.* 102 (1980), 6117–6127.
- [316] S. Kobayashi, Y. Aoki, *Tetrahedron Lett.* 39 (1998), 7345–7348.
- [317] J. P. Tam, *J. Org. Chem.* 50 (1985), 5291–5298.
- [318] S. Jin, D. P. Holub, D. J. Wustrow, *Tetrahedron Lett.* 39 (1998), 3651–3654.
- [319] J.-A. Fehrentz, M. Paris, A. Heitz, J. Velek, C.-F. Liu, F. Winternitz, J. Martinez, *Tetrahedron Lett.* 36 (1995), 7871–7874.
- [320] A. Zheng, D. Shan, B. Wang, *J. Org. Chem.* 64 (1999), 156–161.
- [321] S. Kobayashi, I. Hachiya, S. Suzuki, M. Moriwaki, *Tetrahedron Lett.* 37 (1996), 2809–2812.
- [322] S. Kobayashi, M. Moriwaki, *Tetrahedron Lett.* 38 (1997), 4251–4254.
- [323] J. M. Schlatter, R. H. Mazur, *Tetrahedron Lett.* 33 (1977), 2851–2852.
- [324] D. A. Jones Jr., *Tetrahedron Lett.* 33 (1977), 2853–2856.
- [325] S. A. Khan, K. M. Sivanandaiah, *Synthesis* (1978), 750–751.
- [326] A. Ljungquist, K. Folkers, *Z. Naturforschg.* 38b (1983), 1022–1024.
- [327] M. K. Anwer, A. F. Spatola, *Tetrahedron Lett.* 22 (1981), 4369–4372.
- [328] M. K. Anwer, A. F. Spatola, C. D. Bossinger, E. Flanigan, R. C. Liu, D. B. Olsen, D. Stevenson, *J. Org. Chem.* 48 (1983), 3503–3507.

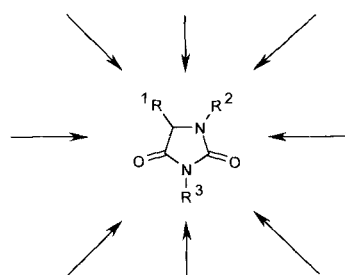
- [329] R. Colombo, *J. Chem. Soc., Chem. Commun.* (1981), 1012–1013.
- [330] B. J. Hall, J. D. Sutherland, *Tetrahedron Lett.* 39 (1998), 6593–6596.
- [331] M. Paris, A. Heitz, V. Guerlavais, M. Cristau, J.-A. Fehrentz, J. Martinez, *Tetrahedron Lett.* 39 (1998), 7287–7290.
- [332] B. Sauerbrei, V. Jungmann, H. Waldmann, *Angew. Chem.* 110 (1998), 1187–1190; *Angew. Chem. Int. Ed. Engl.* 37 (1998), 1143–1146.
- [333] D. T. Elmore, D. J. S. Guthrie, A. D. Wallace, S. R. E. Bates, *J. Chem. Soc., Chem. Commun.* (1992), 1033–1034.

# 6 The Use of Templates in Combinatorial Chemistry for the Solid-Phase Synthesis of Multiple Core Structure Libraries

*Hartmut Richter, Axel W. Trautwein and Günther Jung*

## 6.1 Introduction

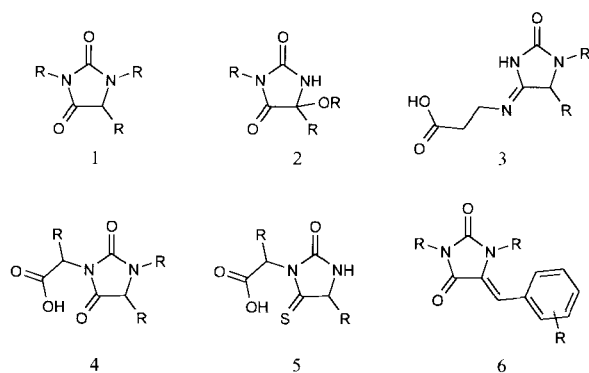
Combinatorial chemistry has become a very important tool in the drug discovery process as it is applied to random screening as well as to lead optimization. The solid-phase synthesis technique is the most widely applied strategy for library construction, and a large number of syntheses have been established on solid support [1, 2]. Incorporating a high level of diversity into a library is a most important issue. Normally, a diverse library is created by synthesizing a single core structure from different building blocks. For introducing a variable substitution pattern into a given core structure, for example a hydantoin, one should use several different pathways with structurally different building blocks. This helps to exploit the space of diversity around a single core structure (Scheme 6.1).



**Scheme 6.1.** Introducing diversity by synthesizing one core structure via different pathways.

For example, several different hydantoin syntheses have been published. Starting from polymer-bound Fmoc-amino acids, De Witt et al. [3] synthesized hydantoin after cleavage of the Fmoc-group via urea synthesis with isocyanates followed by an intramolecular cyclization during cleavage from the resin. Matthews and Rivero [4] and Kim et al. [5] obtained higher diversity by reductive alkylation of the free amino group before urea formation. For the synthesis of 5-alkoxy-hydantoin, Hanessian and Yang [6] used immobilized  $\alpha$ -hydroxy acids which were acylated with  $(\text{CF}_3\text{SO}_2)_2\text{O}$  and subsequently reacted with O-benzyl-hydroxylamine. After urea formation with isocyanates, the resulting substrates were cyclized during cleavage from the resin. Short et al. [7] used polymer-bound isonitriles, which were reacted with aldehydes, amines and potassium cyanate in a Ugi-reaction (Chapter 4) to form iminohydantoin, while Bauser et al. [8] used polymer-bound active carbamates of amino acids which were reacted with methyl esters of amino

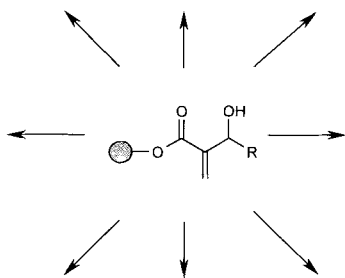
acids. Intramolecular cyclization on the solid support provided polymer-bound hydantoin which could then be modified via alkylation [9]. Bhalay et al. [10] started from dipeptides, the free amino groups being acylated with diphosgene or  $\text{CSCl}_2$  to obtain hydantoin or thiohydantoin after cyclization. Scheme 6.2 shows highly variable hydantoin with their substitution patterns.



**Scheme 6.2.** Hydantoin synthesized via different pathways [3–9].

Another approach for higher diversity is to synthesize different core structures from one template by using different building blocks (Scheme 6.3). This is a promising strategy for lead finding, especially if there is no information on the binding site of the target. Polymer-bound enones,  $\alpha$ -bromoketones and 3-hydroxy-2-methylidene propionic acids have also been shown to be suitable templates for this synthetic strategy.

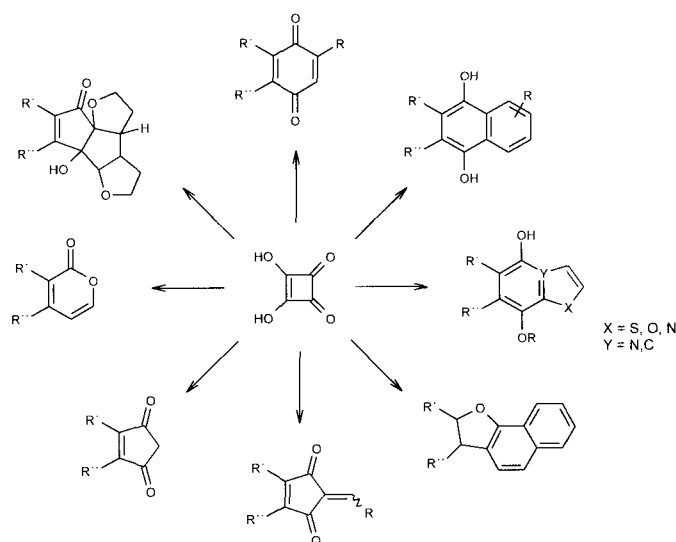
By explaining these templates in detail, we will show their possibilities as key intermediates for the synthesis of a large range of different core structures. Tempest and Armstrong [11] have introduced the term ‘Multiple Core Structure Libraries’ (MCSL) for resin-bound squaric acid derivatives.



**Scheme 6.3.** Synthesis of different core structures by starting from one polymer-bound template (MCSL), for example, 3-hydroxy-2-methylidenepropionic acids.

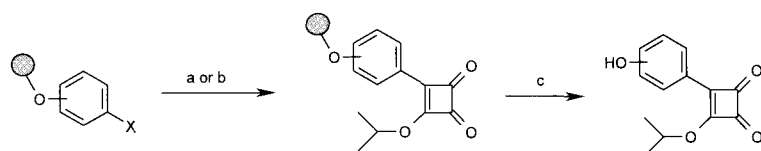
## 6.2 Squaric Acid as Template

Tempest and Armstrong [11] have introduced squaric acid as a template for the generation of MCSLs. Starting from squaric acid derivatives, different carbo- and heterocyclic compounds can be generated (Scheme 6.4).



**Scheme 6.4.** Squaric acid as educt for the synthesis of highly diverse core structures.

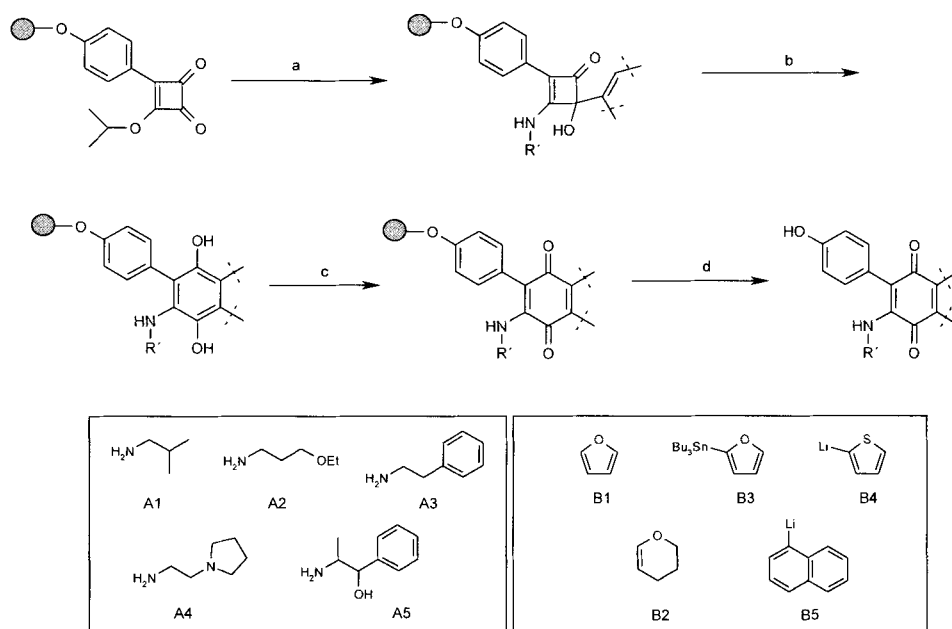
Two different routes were used for the immobilization of squaric acid on Wang resin. Both routes used a halogenated phenol, which was immobilized through a Mitsunobu reaction. In the first pathway, the squaric acid derivative was introduced by coupling tributyltin isopropyl squarate to the polymer-bound phenol via a Stille reaction. In the second pathway, a lithium–halogen-exchange was performed and diisopropyl squarate added. The resulting 1,2-addition product then rearranged under acidic conditions.



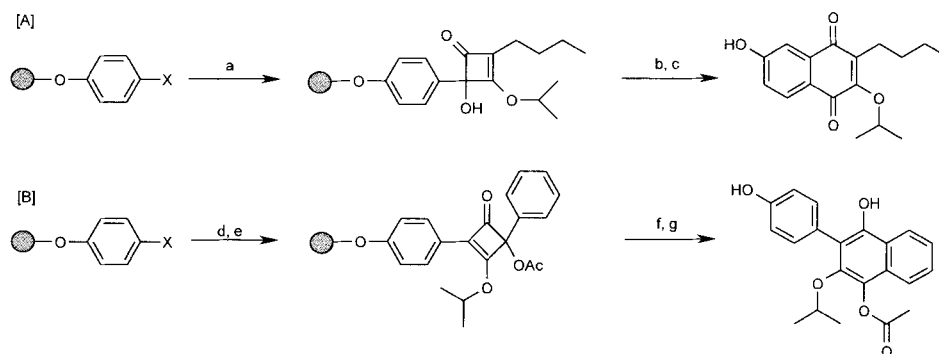
**Figure 6.1.** Key: (a) tributyltin isopropyl squarate, Pd(II), CuI, DMF; (b) 1. *n*-butyl-lithium, THF,  $-78^{\circ}\text{C}$ , 15 min; 2. diisopropyl squarate; 3. cat. 12 N HCl in  $\text{CH}_2\text{Cl}_2$ ; (c) 20% TFA /  $\text{CH}_2\text{Cl}_2$  [10].

This resin-bound substrate was then used for the synthesis of 1,4-quinone derivatives (Fig. 6.2). In the first step, the isopropoxy-group was displaced with primary amines (Fig.

6.2, A1–A5) in a Michael-type reaction. These compounds then reacted with lithiated aromatic or heteroaromatic compounds (Fig. 6.2, B3–B5) or with enoethers (Fig. 6.2, B1–B2). After thermolysis and air oxidation in toluene under reflux, substituted 1,4-quinone derivatives were formed and cleaved from the resin with 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>. Twenty-five single compounds were synthesized in a seven-step synthesis, and the isolated yields ranged from 0% to 53%.



**Figure 6.2.** Key: (a) 1. amine A1–A5, THF, 30 min, 2. B1–B5; (b) toluene, reflux; (c) oxidation; (d) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub> [10].



**Figure 6.3.** Key: [A] (a) *n*-BuLi, THF,  $-78^{\circ}\text{C}$ , then 3-(*n*-butyl) 4-isopropoxy squarate, 15 min; (b) toluene, reflux, air, 20 h; (c) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>; [B] (d) *n*-BuLi, THF,  $-78^{\circ}\text{C}$ , 30 min, then diisopropoxy squarate, cat. HCl, rearrangement, (e) PhLi, THF,  $-78^{\circ}\text{C}$ , then Ac<sub>2</sub>O, (f) toluene, reflux, air, 20 h, (g) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub> [10].

By immobilizing 3-(n-butyl) 4-isopropoxy squarate instead of diisopropyl squarate, naphthalines were accessible. Thermolysis and air oxidation of this resin-bound substrate gave a 1,4-naphthoquinone derivative (Fig. 6.3). If the hydroxy function was acylated first, no air oxidation was possible during thermolysis and 1,4-dihydroxynaphthalines were obtained (Fig. 6.3). Both compounds were cleaved from resin with 20% TFA in  $\text{CH}_2\text{Cl}_2$ .

The synthesis of substituted squaric acid derivatives was shown in two other reactions. In one synthesis, the educt shown in Figure 6.2 was used; the isopropoxy group was displaced with different secondary amines and the products were cleaved from the resin.

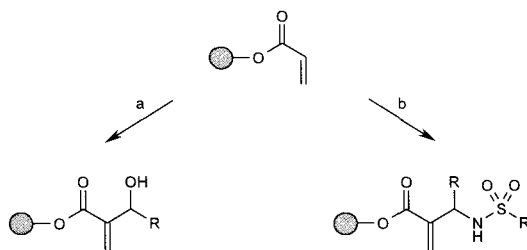
### 6.3 Templates Derived from the Baylis–Hillman Reaction

In Section 6.2, the versatile use of squaric acid as a template was summarized. Only a limited number of diverse squaric acid derivatives is accessible which are suitable for immobilization, and they must be synthesized in solution. Synthesis of a variable template on the resin offers the advantage of introducing higher diversity. Carbon-carbon bond-forming reactions are useful for the construction of such templates. Many of these reactions have been performed on solid phase, such as palladium-catalyzed reactions (Suzuki- [12]; Heck- [13]; and Stille- [14] -reactions), aldol [15] and Knoevenagel [16] condensations, and the Horner–Emmons reaction [17]. However, during these reactions often only one functional group for further transformation is introduced, though two or more functional groups are desirable. By using the Baylis–Hillman reaction for the template synthesis, two functional groups are introduced.

#### 6.3.1 Baylis–Hillman Reaction on the Solid Phase

Prien et al. [18] have synthesized 3-hydroxy-2-methylidene propionic acids on hydroxyethyl resin via a Baylis–Hillman reaction by using aldehydes bearing electron-withdrawing groups, for example nitrobenzaldehyde, trifluoromethylbenzaldehyde and pyridine-carboxaldehyde.

In the solid-phase Baylis–Hillman reaction developed in our group [19] resin-bound acrylic ester reacted with aldehydes to form 3-hydroxy-2-methylidene-propionic acids, or with aldehydes and sulfonamides in a three-component reaction to form 3-aminoaryl-2-methylidene sulfonylpropionic acids [20] (Fig. 6.4).



**Figure 6.4.** Key: (a) aldehyde, DABCO or 3-HQN,  $\text{CHCl}_3/\text{DMSO}$  1:1, rt, 2–10 d; (b) aldehyde, sulfonamide, DABCO, dioxane,  $70^\circ\text{C}$ , 16 h.

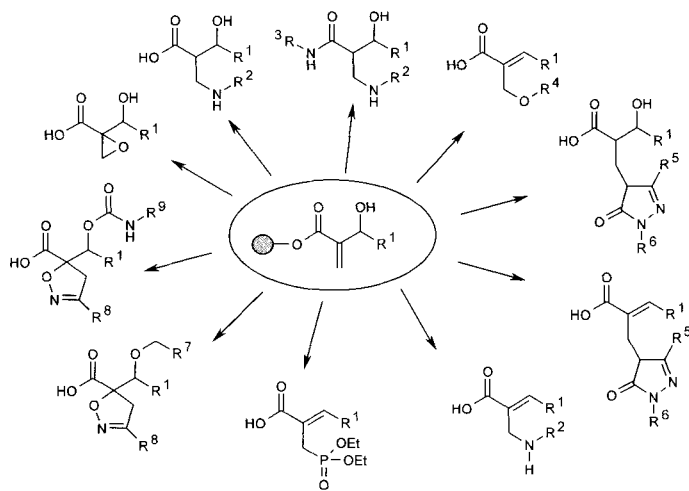


We chose 2-chlorotrityl chloride resin for the attachment of acrylic acid, because in solution-phase chemistry the best results have been obtained by using aryl acrylates or *tert*-butyl acrylates [21]. In addition to DABCO® (1,4-diazabicyclo [2.2.2]octane) – the most common tertiary cyclic amine for this type of reaction – we also used the more reactive 3-quinuclidinol (3-hydroxy-quinuclidine, 3-HQN) for the Baylis–Hillman reaction with aldehydes. We used 26 different aldehydes and obtained good to excellent purities, as determined by analytical HPLC.

The three-component Baylis–Hillman reaction was also performed on 2-chlorotrityl chloride resin by treating polymer-bound acrylic acid with aldehydes and sulfonamides in dioxane at 70 °C for 16 h under DABCO catalysis (Fig. 6.4). Both scaffolds, 3-hydroxy-2-methylidene propionic acids as well as 2-methylidene-3-aminoarylsulfonyl-propionic acids, are precursors for the synthesis of MCSLs.

### 6.3.2 The Use of 3-Hydroxy-2-Methylidene Propionic Acids as Templates

The following syntheses demonstrate the application of resin-bound 3-hydroxy-2-methylidene propionic acids. Within only a few steps, a significant number of different linear and heterocyclic cores can be built up (Scheme 6.5) [22].



**Scheme 6.5.** Polymer-bound 3-hydroxy-2-methylidene propionic acid, a template for the synthesis of various core structures.

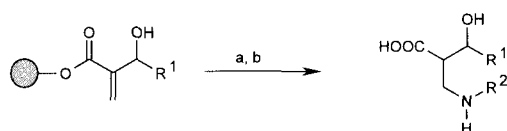
Due to the presence of two functional groups, there is a wide range of possible post-modifications. One possibility is the conversion of the double bond, as in the Michael addition of amines leading to 1,3-amino-alcohols [19], by epoxidation or a Michael addition of a  $\beta$ -ketoester to form pyrazolones [23] after the conversion into hydrazones, followed by an intramolecular cyclization.

Another possibility is to carry out further reactions involving the hydroxy function, for example in a Mitsunobu reaction [22]. In this case, the double bond is rearranged in a manner similar to the synthesis of allylic amines [23].

The final option is to convert both functional groups in one reaction step. During the synthesis of isoxazolines [24] via a 1,3-dipolar cyclo-addition with nitrile oxides [25], the hydroxy function reacts with phenylisocyanate which is used to dehydrate the nitroalkane to form a carbamate.

### 6.3.2.1 Michael Addition of Amines

To investigate further modifications of the resin-bound allylic alcohols, we synthesized 1,3-amino alcohols via a non stereoselective Michael-type addition of amines (Fig. 6.5).



**Figure 6.5.** Key: (a)  $R^2NH_2$ , DMF, rt, 16 h; (b) 5% TFA/ $CH_2Cl_2$ , rt, 30 min.

The Michael-type addition of 2-phenylethylamine was also used to determine the conversion rate in the first step, the Baylis–Hillman reaction. This was done by comparing the amounts of the two resulting addition products derived from addition of 2-phenylethylamine to the Baylis–Hillman product and unreacted acrylic acid. Electron-rich benzaldehydes gave poor results in the Baylis–Hillman reaction, the rate of conversion decreasing from 2-methoxy- to 4-methoxy-benzaldehyde; not surprisingly, with 2,4-dimethoxybenzaldehyde no product was obtained at all [19]. Therefore, only aldehydes which gave a satisfying conversion rate in the Baylis–Hillman reaction were used in the following modifications (Table 6.1).

**Table 6.1.** Variation of the aldehydes in the 1,3-aminoalcohol synthesis by using 2-phenylethyl-amine as constant building block

Pos.	Aldehyde	MW	Purity (%) <sup>a</sup>
1	4-Trifluoromethylbenzaldehyde	367	96
2	4-Fluorobenzaldehyde	317	80
3	Thiophen-2-carboxaldehyde	305	89
4	4-Chlorobenzaldehyde	333	91
5	4-Bromobenzaldehyde	378	96
6	4-Cyanobenzaldehyde	324	93
7	2-Bromobenzaldehyde	378	92
8	3-Trifluoromethylbenzaldehyde	367	91

<sup>a</sup> Purity determined by analytical HPLC ( $\lambda = 214$  nm).

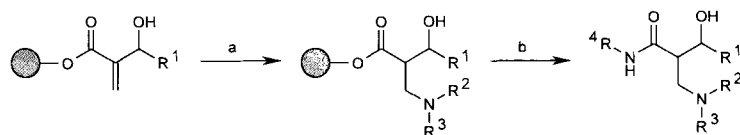
**Table 6.2.** Variation of the amines in the 1,3-amino alcohol synthesis by using 4-trifluoromethyl-benzaldehyde as constant building block

Pos.	Amine	MW	Purity (%) <sup>a</sup>
1	Pentylamine	333	99
2	Cyclopropylamine	303	85
3	4-(2-Aminoethyl)morpholine	376	93
4	4-Aminobutanol	335	95
5	2-(2-Aminoethyl)pyridine	368	90
6	Allylamine	303	96
7	2-Amino-1-phenylethanol	383	71
8	Pyrrolidine	317	94

a Purity determined by analytical HPLC ( $\lambda = 214$  nm).

Finally, the scope and limitations of this solid-phase approach were investigated by varying aldehydes (Table 6.1) and amines (Table 6.2). The compounds were all characterized using ES-MS and analytical HPLC.

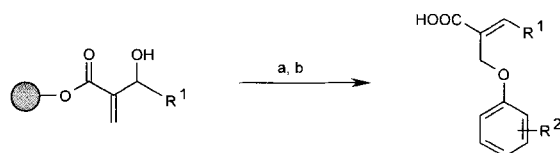
Prien et al. [18] used secondary amines for the formation of 1,3-amino alcohols. Another substituent was introduced by aminolytic cleavage under the formation of C(2), C(3)-disubstituted 3-hydroxypropionic acid amides. Thirty-six compounds were synthesized, with yields ranging from 40 % to 66 % (Fig. 6.6).



**Figure 6.6.** Key: (a)  $R^2R^3NH$ , 22 °C, 18 h, DMF; (b)  $R^4NH_2$ ,  $(CH_3)_3Al$ ,  $CH_2Cl_2$ , toluene, 0–22 °C, 16 h.

### 6.3.2.2 Mitsunobu Reaction

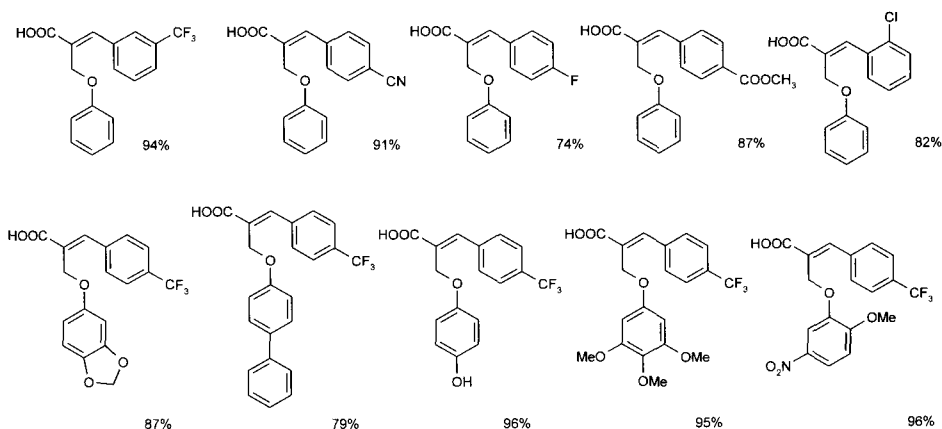
The Mitsunobu reaction [26] with phenols was carried out under standard reaction conditions.



**Figure 6.7.** Key: (a) phenol, DIAD,  $PPh_3$ , THF, rt, 3 h; (b) 5 % TFA/ $CH_2Cl_2$ , rt, 30 min.

This reaction proceeded very well, and only the  $S_N^2$ -reaction product was obtained. To determine the Z-/E-ratio of the final product, NOE-experiments were used, these showing that only the E-product was formed.

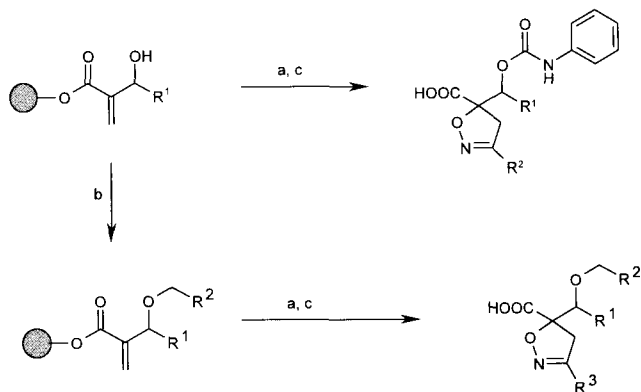
A wide range of aromatic and heteroaromatic aldehydes and phenols were incorporated for the parallel synthesis [22]. Scheme 6.6 shows some representative structures with their HPLC-purities ( $\lambda = 214$  nm).



**Scheme 6.6.** Examples of some substituted aryloxyethers, together with their HPLC-purities ( $\lambda = 214$  nm).

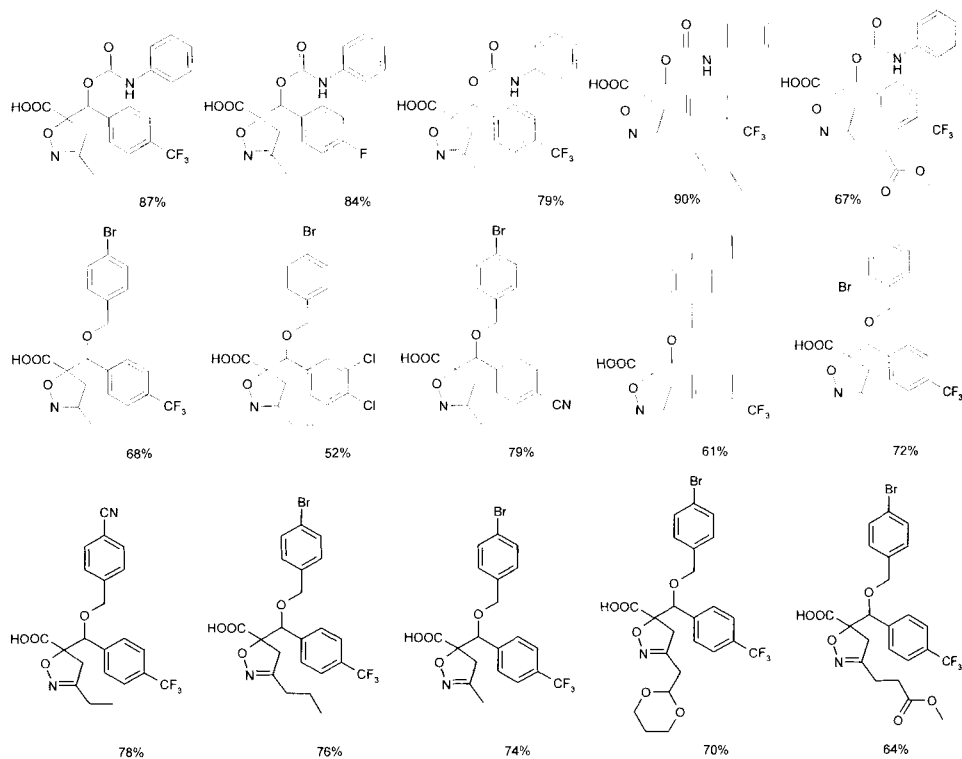
### 6.3.2.3 1,3-Dipolar Cycloaddition with Nitrile Oxides

To investigate further modifications of the resin-bound allylic alcohols, we synthesized isoxazolines [24] via a 1,3-dipolar cycloaddition of nitrile oxides. The nitrile oxides were created by using primary nitroalkanes, phenylisocyanate and triethylamine. Two different strategies were followed (Fig. 6.8). When the 3-hydroxy-2-methylidene propionic acids were treated with nitrile oxides, the hydroxy function reacted with phenylisocyanate under carbamate formation. This by-product formation was prevented by alkylating the hydroxy function with benzyl bromides prior to the 1,3-cycloaddition step.



**Figure 6.8.** Key: (a)  $R^2CH_2NO_2$ , phenylisocyanate,  $Et_3N$ , toluene, rt, 4 h, 3  $\times$ ; (b) benzyl bromide, BEMP, DMF, 60  $^\circ C$ , 5 h; (c) 5% TFA/ $CH_2Cl_2$ , rt, 30 min.

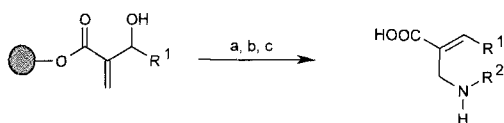
After having optimized the reaction conditions, aldehydes, nitroalkanes and benzyl bromides were varied systematically. In total, 40 isoxazolines were synthesized in parallel, with purities ranging from 52% to 90%. Only one regioisomer was formed as a mixture of diastereomers during the reaction; this was shown by 2D-NMR. Scheme 6.7 shows some representative structures, with the corresponding HPLC-purities ( $\lambda = 214$  nm).



**Scheme 6.7.** Examples of the synthesized isoxazolines, together with their HPLC-purities ( $\lambda = 214$  nm).

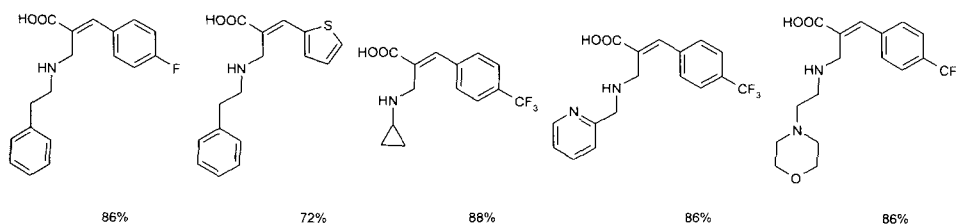
#### 6.3.2.4 Synthesis of 2-Methylidene- $\beta$ -Alanines

To allow for a diverse multi-step synthesis, we transformed the 3-hydroxy-2-methylidene propionic acids (Fig. 6.9) into polymer-bound allylic amines, which can be considered as unusual  $\beta$ -amino acid derivatives. The polymer-bound allylic alcohols were first treated with acetyl chloride and DIEA in  $\text{CH}_2\text{Cl}_2$  to form the ester, which was reacted with primary amines in an addition elimination step to form allylic amines.



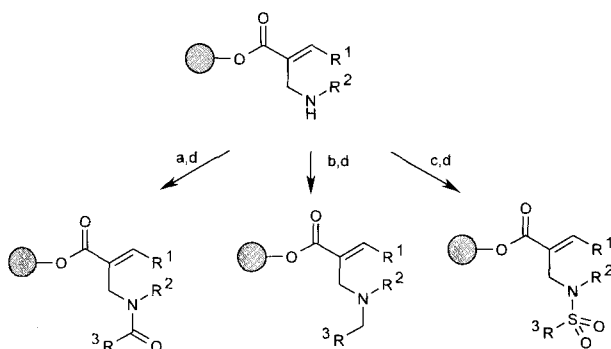
**Figure 6.9.** Key: (a) acetyl chloride, DIEA,  $\text{CH}_2\text{Cl}_2$ , rt, 16 h; (b)  $\text{R}^2\text{NH}_2$ , BEMP, DMF, rt, 6 h; (c) 5% TFA/ $\text{CH}_2\text{Cl}_2$ , rt, 30 min.

Similar to the Mitsunobu reaction, only the E-product of the  $\text{S}_{\text{N}}'1$ -reaction was obtained. This was determined by using NOE-experiments. The systematic variation of 10 amines and 10 aldehydes provided compounds with HPLC-purities of 75–92% ( $\lambda = 214$  nm). Some representative structures of the synthesized 2-methylidene- $\beta$ -alanines, together with their HPLC-purities are shown in Scheme 6.8.



**Scheme 6.8.** Structures of some 2-methylidene- $\beta$ -alanines, together with their HPLC-purities.

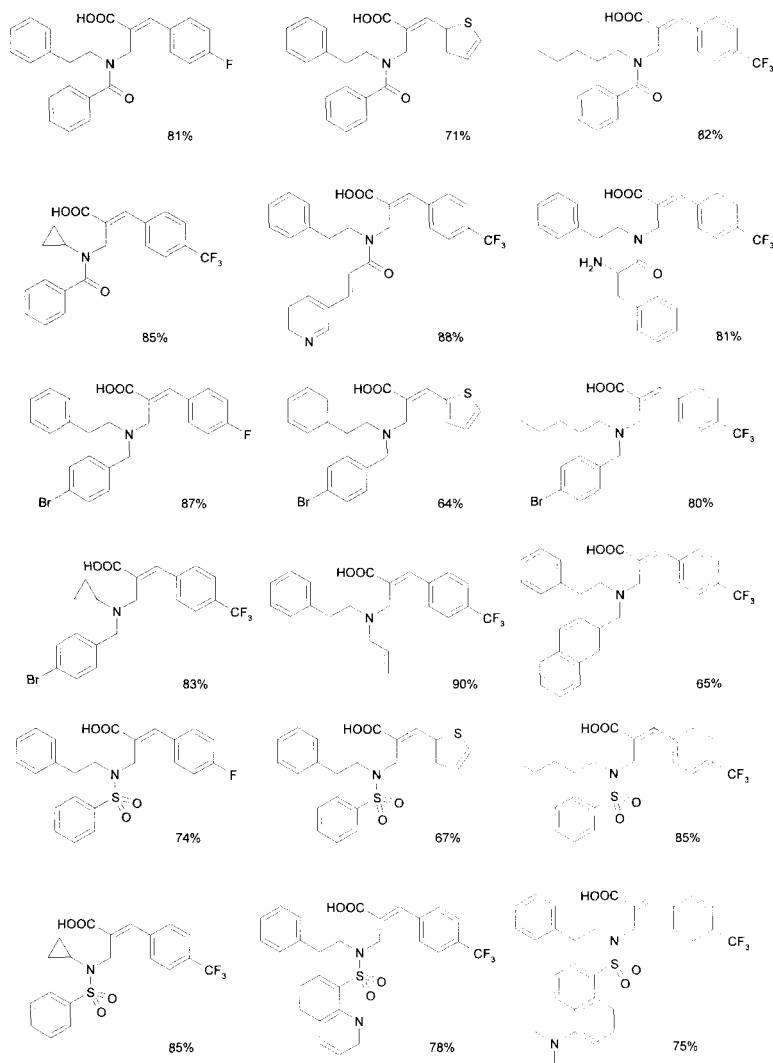
The polymer-bound allylic amines can be employed in several reactions. Figure 6.10 shows acylation, alkylation and sulfonation as possible modifications.



**Figure 6.10.** Key: (a) carboxylic acid, PyBroP, DIEA,  $\text{CH}_2\text{Cl}_2/\text{DMF}$  1:1, rt, 4 h, 2  $\times$ ; (b) benzyl bromide, BEMP, DMF, rt, 16 h; (c) sulfonylchloride, DIEA,  $\text{CH}_2\text{Cl}_2$ , rt, 5 h; (d) 5% TFA/ $\text{CH}_2\text{Cl}_2$ , 30 min.

For the determination of scope and limitations of these reactions, we varied aldehydes, amines, acids, sulfonylchlorides and alkylating reagents by using 10 examples of each building block [22]. HPLC-purities ranged from 60% to 92% ( $\lambda = 214$  nm). Scheme 6.9

shows some representative structures of the acylated, sulfonated and alkylated 2-methylidene- $\beta$ -alanines, together with their HPLC-purities.

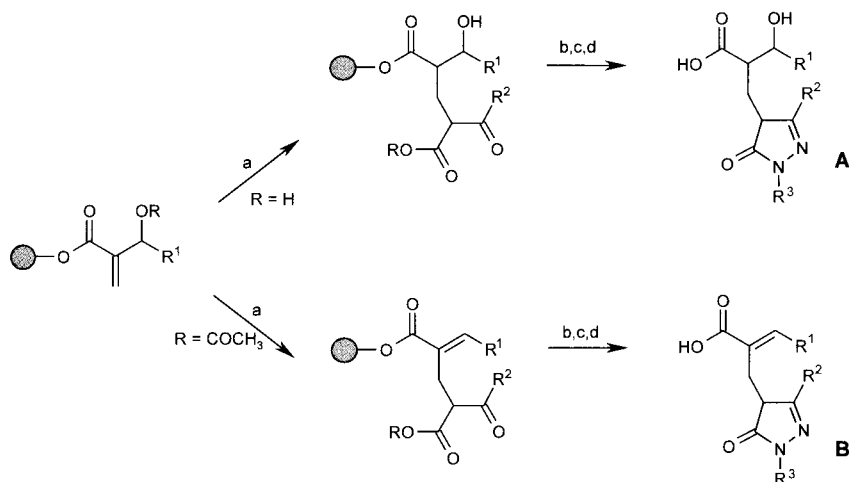


**Scheme 6.9.** Examples of acylated, alkylated and sulfonated 2-methylidene- $\beta$ -alanines, together with their HPLC-purities ( $\lambda = 214$  nm).

### 6.3.2.5 Synthesis of Pyrazolones

To demonstrate the possibilities of Michael additions, we chose the synthesis of pyrazolones [27] as an example (Fig. 6.11). The addition of  $\beta$ -ketoesters in the presence of BEMP as base provided resin-bound  $\beta$ -ketoesters, which were transformed into the cor-

responding hydrazones upon treatment with hydrazine hydrochlorides, DIEA and TMOF. During cleavage with 5% TFA in  $\text{CH}_2\text{Cl}_2$ , the hydrazone cyclized under pyrazolone formation. The use of polymer-bound 3-hydroxy-2-methylidene propionic acids as educts resulted in pyrazolones A, and starting from the acetylated polymer-bound 3-hydroxy-2-methylidene propionic acids we obtained pyrazolones B [22], the dehydrated forms of A.



**Figure 6.11.** Key: (a)  $\beta$ -ketoester, BEMP, THF, rt, 4 h; (b) phenylhydrazine hydrochloride, DIEA, TMOF, THF, rt, 8 h; (c) 5% TFA/ $\text{CH}_2\text{Cl}_2$ , rt, 30 min.

For systematic variation of the aldehydes, ethyl acetoacetate and phenylhydrazine were used as constant building blocks (Table 6.3). 4-Trifluoromethylbenzaldehyde and phenylhydrazine were used in the variation of the  $\beta$ -ketoester (Table 6.4), while 4-trifluoromethylbenzaldehyde and ethyl acetoacetate were applied in the aryhydrazine variation (Table 6.5). All compounds were analyzed by ES-MS and HPLC ( $\lambda = 214$  nm).

**Table 6.3.** Variation of the aldehyde in the pyrazolone synthesis by using ethyl acetoacetate and phenylhydrazine as constant building blocks

Pos.	Aldehyde	Compound A		Compound B	
		MW	Purity (%) <sup>a</sup>	MW	Purity (%) <sup>a</sup>
1	4-Trifluoromethylbenzaldehyde	420	79	402	72
2	4-Chlorobenzaldehyde	386	75	368	66
3	Methyl-4-formylbenzoate	410	70	392	64
4	4-Bromobenzaldehyde	431	76	413	70
5	4-Fluorobenzaldehyde	370	75	352	74
6	Thiophene-2-carbaldehyde	358	53	340	49
7	4-Cyanobenzaldehyde	377	80	359	72

<sup>a</sup> Purity determined by analytical HPLC ( $\lambda = 214$  nm), including all isomers.



**Table 6.4.** Variation of the  $\beta$ -ketoester in the pyrazolone synthesis by using 4-trifluoromethylbenzaldehyde and phenylhydrazine as constant building blocks

Pos.	$\beta$ -Ketoester	Compound A		Compound B	
		MW	Purity (%) <sup>a</sup>	MW	Purity (%) <sup>a</sup>
1	Diethyl-3-oxoheptanedecarboxylate	520	48	502	42
2	Dimethyl-1,3-acetonedecarboxylate	478	68	460	62
3	Ethylbutyrylacetate	448	64	430	85
4	Methyl-4-methoxyacetoacetate	450	85	432	60
5	Ethylpropionylacetate	434	67	416	64

a Purity determined by analytical HPLC ( $\lambda = 214$  nm), including all isomers.

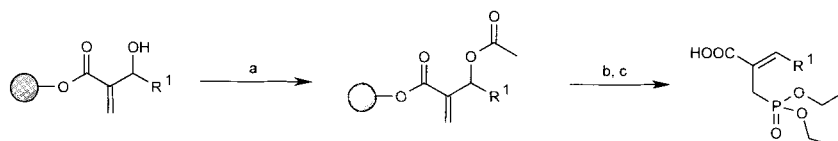
**Table 6.5.** Variation of the arylhydrazine in the pyrazolone synthesis by using 4-trifluoromethylbenzaldehyde and ethyl acetoacetate as constant building blocks

Pos.	Hydrazine	Compound A		Compound B	
		MW	Purity (%) <sup>a</sup>	MW	Purity (%) <sup>a</sup>
1	4-Chlorophenylhydrazine	454	68	436	63
2	3-Nitrophenylhydrazine	465	55	447	52
3	4-Methoxyphenylhydrazine	450	52	432	44
4	Tolylhydrazine	434	74	416	75
5	4-Bromophenylhydrazine	499	65	481	77
6	4-Fluorophenylhydrazine	438	63	410	61

a Purity determined by analytical HPLC ( $\lambda = 214$  nm), including all isomers.

### 6.3.2.6 Synthesis of 2-Diethoxy-Phosphorylmethyl Acrylic Acids

There are two possibilities for converting polymer-bound 3-hydroxy-2-methylidene propionic acids into 2-diethoxy-phosphorylmethyl acrylic acids. One is the treatment of these substrates with diethyl chlorophosphite and  $\text{Et}_3\text{N}$  followed by an Arbuzow rearrangement. The other method, which resulted in higher purities [28], is the reaction of acetylated 3-hydroxy-2-methylidene propionic acids with triethylphosphite in DMF for 5 h at  $60^\circ\text{C}$  (Fig. 6.12). Table 6.6 shows the variation of the aldehydes in this reaction. All compounds were analyzed by ES-MS and analytical HPLC.

**Figure 6.12.** Key: (a) acetyl chloride, DIEA,  $\text{CH}_2\text{Cl}_2$ , rt, 16 h; (b) diethylphosphite, DMF,  $50^\circ\text{C}$ , 6 h; (c) 5% TFA/ $\text{CH}_2\text{Cl}_2$ , rt, 30 min.

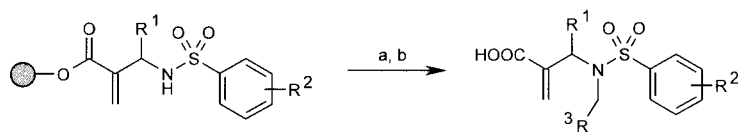
**Table 6.6.** Variation of the aldehyde in the synthesis of 2-diethoxy-phosphorylmethyl acrylic acids

Pos.	Aldehyde	MW	Purity (%) <sup>a</sup>
1	4-Fluorobenzaldehyde	316	50
2	4-Chlorobenzaldehyde	332	36
3	4-Bromobenzaldehyde	377	42
4	4-Trifluoromethylbenzaldehyde	366	70
5	Methyl-4-formylbenzoate	356	50
6	3,4-Dichlorobenzaldehyde	366	42

<sup>a</sup> Purity determined by analytical HPLC ( $\lambda = 214$  nm).

### 6.3.3 Alkylation of 2-Arylsulfonylaminomethyl Acrylic Acids

The polymer-bound 2-arylsulfonylaminomethyl acrylic acids derived from the three-component Baylis–Hillman reaction can also be used as templates for the synthesis of different core structures. Here, we present the alkylation of these templates as an example (Fig. 6.13).



**Figure 6.13.** Key: (a) alkylating reagent, BEMP, DMF, rt, 16 h; (b) 5% TFA/CH<sub>2</sub>Cl<sub>2</sub>, rt, 30 min.

Three different building blocks are incorporated into this synthesis, which were varied systematically. *p*-Toluenesulfonamide and 4-bromobenzyl bromide were employed in the variation of the aldehydes (Table 6.7), 4-trifluoromethylbenzaldehyde and 4-bromobenzyl bromide were used as constant building blocks in the sulfonamide variation (Table 6.8), while the variation of the alkylation reagents was performed with 4-trifluoromethylbenzaldehyde and *p*-toluenesulfonamide (Table 6.9). All compounds were analyzed by ES-MS and analytical HPLC.

**Table 6.7.** Variation of the aldehyde in the synthesis of alkylated 2-arylsulfonylaminomethyl acrylic acids by using *p*-toluenesulfonamide and 4-bromobenzyl bromide as constant building blocks

Pos.	Aldehyde	MW	Purity (%) <sup>a</sup>
1	4-Fluorobenzaldehyde	518	86
2	4-Chlorobenzaldehyde	534	96
3	4-Bromobenzaldehyde	580	80
4	4-Trifluoromethylbenzaldehyde	568	79
5	Methyl-4-formylbenzoate	558	79
6	Thiophene-2-carbaldehyde	506	56
7	3,4-Dichlorobenzaldehyde	568	72

<sup>a</sup> Purity determined by analytical HPLC ( $\lambda = 214$  nm).

**Table 6.8.** Variation of the sulfonamides in the synthesis of alkylated 2-arylsulfonylaminomethyl acrylic acids by using 4-trifluoromethylbenzaldehyde and 4-bromobenzylbromide as constant building blocks

Pos.	Sulfonamide	MW	purity (%) <sup>a</sup>
1	4-Nitrobenzenesulfonamide	599	49
2	Dansylamide	648	84
3	Benzenesulfonamide	554	85
4	4-Methoxybenzenesulfonamide	584	87
5	2-Chlorobenzenesulfonamide	588	90
6	4-Bromobenzenesulfonamide	634	82
7	3-Chlorobenzenesulfonamide	588	71

a Purity determined by analytical HPLC ( $\lambda = 214$  nm).

**Table 6.9.** Variation of the alkylation reagents in the synthesis of alkylated 2-arylsulfonylamino-methyl acrylic acids by using 4-trifluoromethylbenzaldehyde and toluenesulfonamide as constant building blocks

Pos.	Alkylation reagents	MW	Purity (%) <sup>a</sup>
1	4-Fluorobenzylbromide	507	83
2	$\alpha$ -Bromo-p-xylene	503	93
3	2-Bromomethylnaphthalene	538	88
4	N-Bromomethylphthalimide	557	70
5	Allylbromide	439	90
6	Iodomethane	413	84
7	1-Bromopropane	441	78

a Purity determined by analytical HPLC ( $\lambda = 214$  nm).

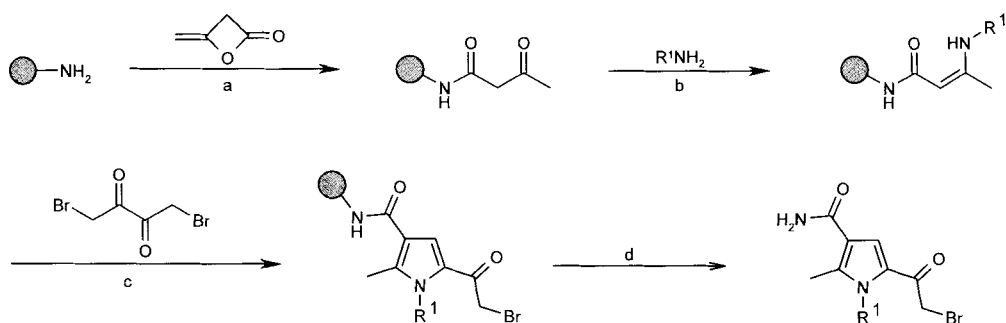
## 6.4 5-(2-Bromoacetyl)pyrroles as Templates

### 6.4.1 Synthesis of Polymer-Bound 5-(2-Bromoacetyl)pyrroles

The  $\alpha$ -bromoketone moiety represents a very versatile functionality in organic synthesis due to its bielectrophilic structure. In our group [29], we have transferred this structural element to the solid phase in the form of polymer-bound 5-(2-bromoacetyl)pyrroles. The synthetic strategy is outlined in Figure 6.14.

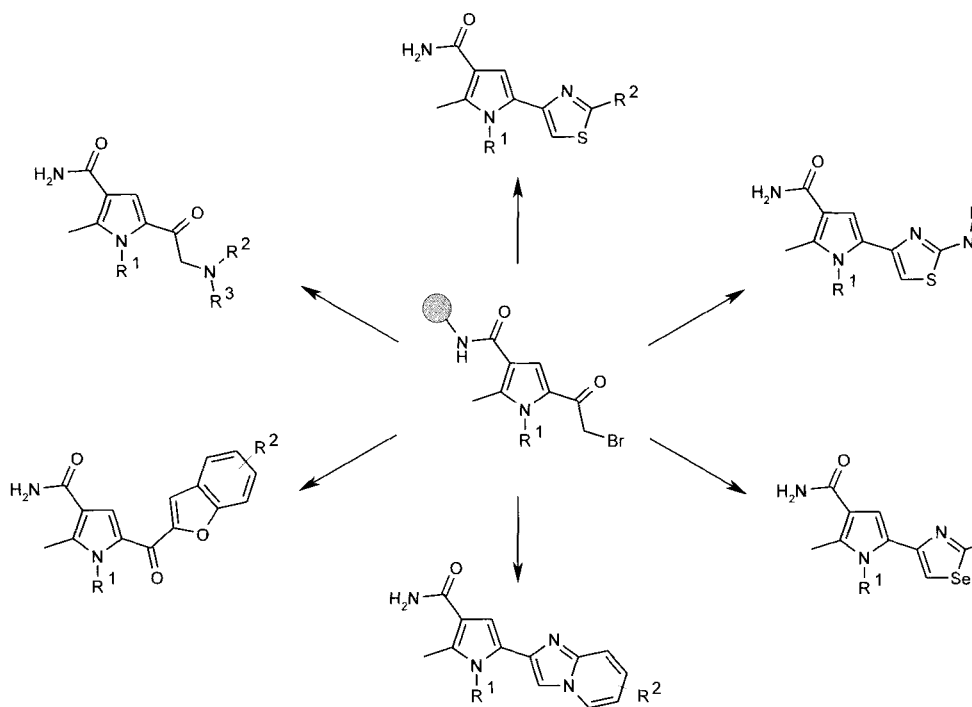
In the first step, Rink amide AM PS Resin was acetoacetylated with diketene. Treatment with primary amines resulted in polymer-bound enaminones which then underwent a Hantzsch reaction [30] with 1,4-dibromo-2,3-butanedione under formation of 5-(2-bromoacetyl)pyrroles which could be cleaved from the resin with 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>.

A wide range of amines can be used in the synthesis of these polymer-bound substrates, except for amines with nucleophilic side chains. This is due to the fact that 1,4-dibromo-2,3-butanedione is a strong alkylating agent. Amines with hydroxy functionalized side chains can be incorporated into the synthesis, whereas amines with tertiary amino groups in the side chain, for example N-(2-aminoethyl)morpholine, are not suitable due to by-product formation.



**Figure 6.14.** Key: (a) diketene,  $\text{CH}_2\text{Cl}_2$ ,  $-15^\circ\text{C} \rightarrow \text{rt}$ , 2.5 h; (b)  $\text{R}'\text{NH}_2$ , DMF, TMOF, rt,  $2 \times 24$  h; (c) 1,4-dibromo-2,3-butanedione, DMF, 2,6-di-*tert*-butylpyridine, rt, 1.5 h; (d) 20% TFA/ $\text{CH}_2\text{Cl}_2$ , rt, 30 min.

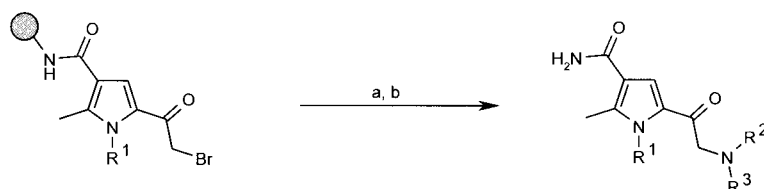
The polymer-bound 5-(2-bromoacetyl)pyrroles offer a wide range of possible transformations, for example nucleophilic substitutions or heterocyclization reactions. Scheme 6.10 shows the structures we have so far derived from these templates. The corresponding syntheses are presented in more detail in Sections 6.4.2–6.4.5.



**Scheme 6.10.** Structures derived from polymer-bound 5-(2-bromoacetyl)pyrroles.

### 6.4.2 Synthesis of 5-(2-Aminoacetyl)pyrroles

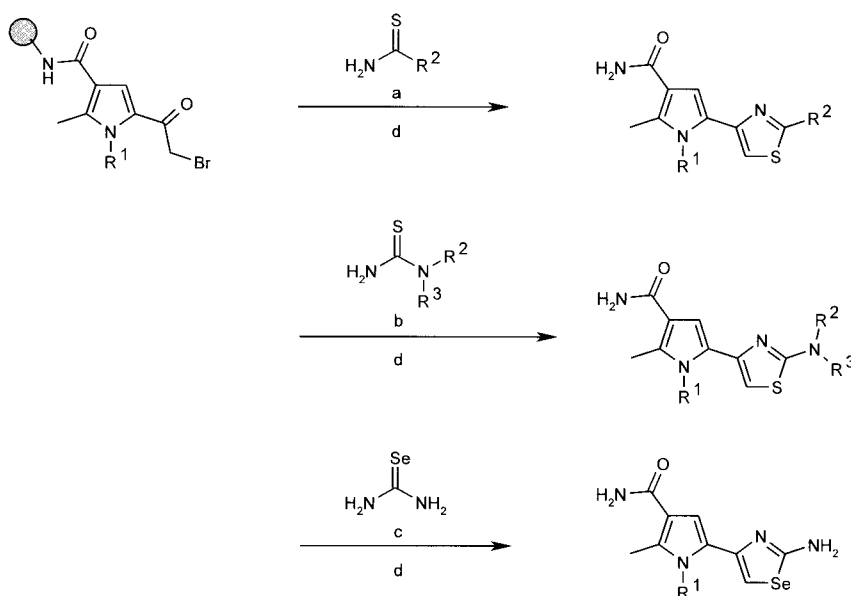
Treatment of polymer-bound 5-(2-bromoacetyl)pyrroles with secondary amines in DMF at room temperature provided 5-(2-aminoacetyl)pyrroles which were cleaved from the resin with 20% TFA/CH<sub>2</sub>Cl<sub>2</sub> (Fig. 6.15). HPLC-purities of the crude products ranged from 74% to 92%.



**Figure 6.15.** Key: (a) R<sup>1</sup>R<sup>2</sup>NH, DMF, rt, 15 h; (b) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>, rt, 30 min.

### 6.4.3 Synthesis of Thiazolylpyrroles, Aminothiazolylpyrroles and Selenazolylpyrroles

Polymer-bound 5-(2-bromoacetyl)pyrroles could be converted into thiazolylpyrroles, aminothiazolylpyrroles and selenazolylpyrroles through Hantzsch-type reaction with thioamides, thioureas or selenourea respectively. The reaction conditions that were optimized for the three different types of building blocks are shown in Figure 6.16.



**Figure 6.16.** Key: (a) R<sup>2</sup>CSNH<sub>2</sub>, DMF, rt, 17 h; (b) R<sup>2</sup>R<sup>3</sup>NCSNH<sub>2</sub>, DMF/AcOH 95:5, rt, 17 h; (c) H<sub>2</sub>NCSeNH<sub>2</sub>, DMF, rt, 1.5 h; (d) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>, rt, 30 min.

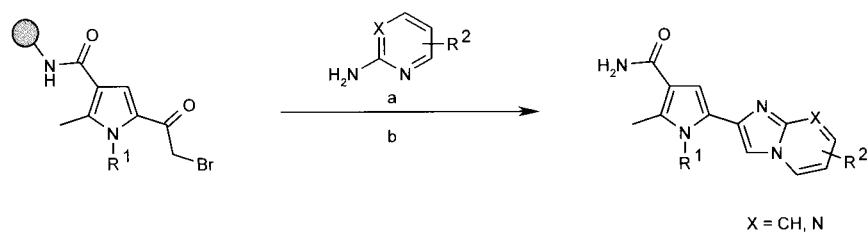
Thiazolylpyrroles were formed in the reaction of thioamides with polymer-bound 5-(2-bromoacetyl)pyrroles in DMF at room temperature (Fig. 6.16a, d). The HPLC purities of the crude products which were obtained in the variation of the amines and the thioamide component ranged from 64% to 95%, the majority of the products having purities >85%.

When these reaction conditions were applied to thioureas as heterocyclization components, the purities of the resulting aminothiazolylpyrroles were poor. However, this drawback was overcome by changing to DMF/AcOH 95:5 as solvent (Fig. 6.16b, d). The variation of amines and thioureas provided crude products with HPLC-purities of 50% to 92%, the majority having purities >80%.

The reaction of polymer-bound 5-(2-bromoacetyl)pyrroles with selenourea in DMF led to selenazolylpyrroles (Fig. 6.16c, d) with HPLC-purities ranging from 0% to 85%. No product was obtained when 3-amino-1-propanol was used as the amino component; purities in all other cases were >70%.

#### 6.4.4 Synthesis of Imidazo[1,2-a]pyri(mi)dylpyrroles

Treatment of polymer-bound 5-(2-bromoacetyl)pyrroles with 2-aminopyridines or 2-aminopyrimidines resulted in the formation of imidazo[1,2-a]pyridylpyrroles or imidazo[1,2-a]pyrimidylpyrroles respectively. Optimal reaction conditions for this heterocyclization were 7 h at 60 °C in DMF/EtOH (1:1) (Fig. 6.17).

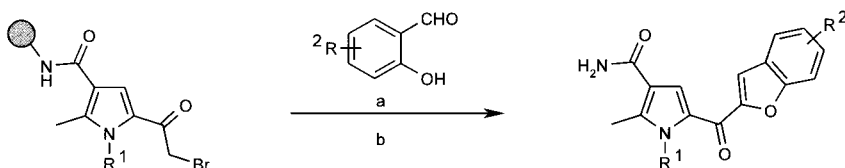


**Figure 6.17.** (a) 2-aminopyri(mi)dine, DMF/EtOH 1:1, 60 °C, 7 h; (b) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>, rt, 30 min.

For the variation of the 2-aminopyri(mi)dine component 2-phenylethylamine was used as constant building block. HPLC-purities ranged from 18% to 85%. A purity of 18% was obtained with 2-amino-5-nitropyridine; this was due to the decreased nucleophilicity of the ring N-atom and the amino group in this building block. 2-Aminopyridine was used as constant building block in the variation of the amino component; in this case, HPLC-purities ranged from 55% to 90%.

### 6.4.5 Synthesis of Benzofurylpyrrolylketones

Polymer-bound 5-(2-bromoacetyl)pyrroles were converted into benzofurylpyrrolylketones through the addition of *o*-hydroxybenzaldehydes under basic conditions. Several bases were tested for this purpose, for example potassium *tert*-butoxide, BEMP (2-*tert*-butylimino-2-diethylamino-1,3-dimethyl-perhydro-1,3,2-diazaphosphorine), sodium ethoxide, and DBU. The best results were obtained with BEMP in NMP at 60 °C (Fig. 6.18), the products being cleaved from the resin with 20 % TFA/CH<sub>2</sub>Cl<sub>2</sub>.

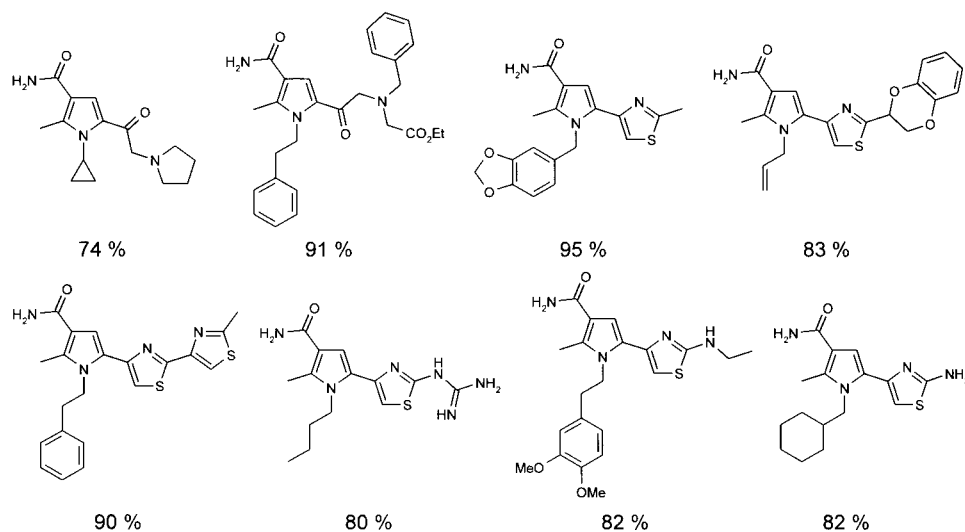


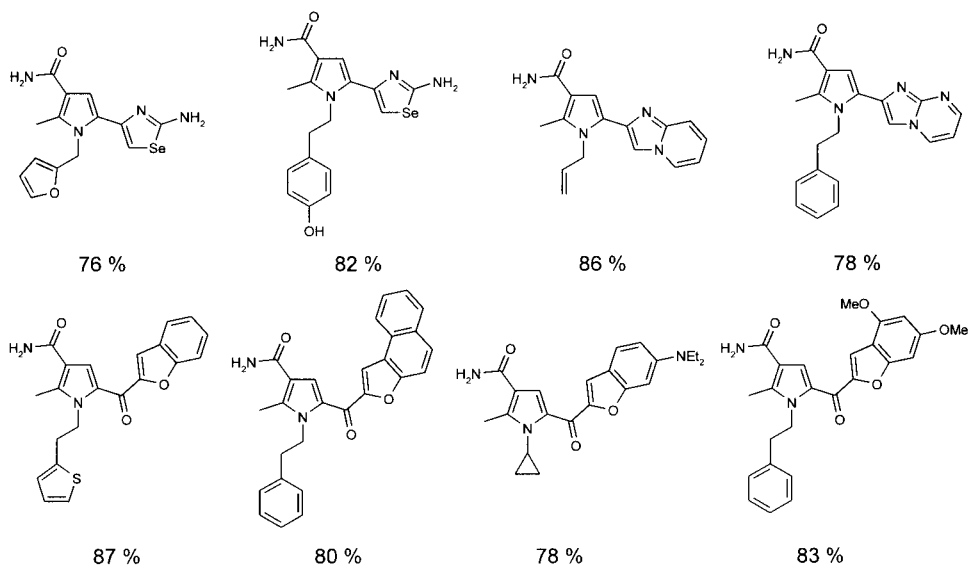
**Figure 6.18.** Key: (a) 2-hydroxybenzaldehyde, NMP, BEMP, 60 °C, 2 h; b) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>, rt, 30 min.

The HPLC-purities of the crude products ranged from 0% to 90%. Little or no product was formed when 2-hydroxy-5-nitrobenzaldehyde or sterically hindered aldehydes such as 3,5-dichloro-2-hydroxybenzaldehyde were used.

### 6.4.6 Examples

Scheme 6.11 shows some structures that were synthesized from polymer-bound 5-(2-bromoacetyl)pyrroles via the strategies shown in Section 6.4.1–6.4.5, together with their corresponding HPLC-purities (crude products,  $\lambda = 214$  nm).

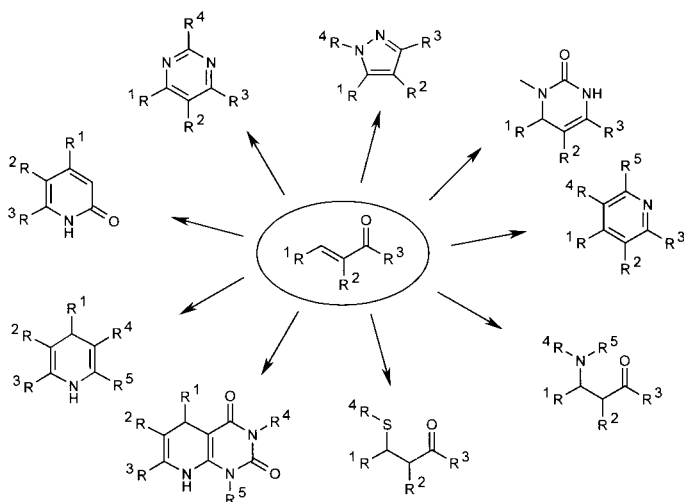




**Scheme 6.11.** Examples of structures derived from polymer-bound 5-(2-bromoacetyl)-pyrroles.

## 6.5 The Use of Enones as Templates

Polymer-bound enones also represent interesting templates for the synthesis of a variety of core structures. Scheme 6.12 illustrates some of the synthetic transformations possible with polymer-bound enones.

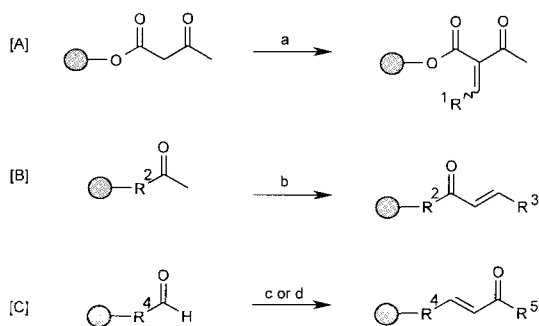


**Scheme 6.12.** Polymer-bound enones as educts for the synthesis of diverse core structures.



### 6.5.1 Synthesis of Polymer-Bound Enones

Several strategies for the synthesis of polymer-bound enones have been described. One way is to start from immobilized  $\beta$ -ketoesters, which can be prepared via transesterification of Wang resin with alkyl  $\beta$ -keto carboxylates [31], or by treatment with diketene [16]. Knoevenagel reactions of these polymer-bound  $\beta$ -ketoesters with aldehydes led to the formation of 2-alkylidene- or arylidene- $\beta$ -ketoesters (Fig. 6.19 (A)).

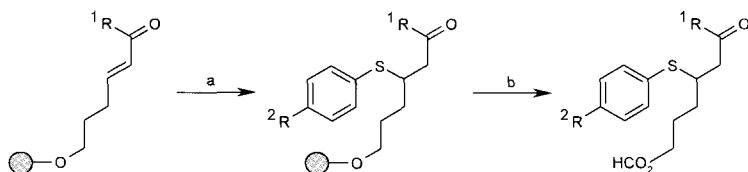


**Figure 6.19.** Key: [1] (a) aldehyde, piperidine cat., isopropanol-benzene, 60 °C; [2] (b) aldehyde, base, THF; [3] (c) acetyl substituted triphenylphosphonates, base, THF, 60 °C, 2 d; (d) acetophenone, LiOH, DME, 16 h, rt.

Another route uses polymer-bound acetophenones, for example hydroxy- or carboxyacetophenones, which are condensed with aldehydes (Fig. 6.19 (B)) [32,33]. A third alternative is to apply immobilized aldehydes in a Wittig reaction, or to condense these aldehydes with acetophenones under base catalysis (Fig. 6.19 (C)) [34, 35].

### 6.5.2 Michael Addition of Aryl Thiolates

The group of M. J. Kurth [36] used polymer-bound enones for the Michael-type addition of aryl thiolates (Fig. 6.20). In the first step 1,4-butanediol was attached to PS-tritylchloride resin. This was followed by an oxidation to the aldehyde, subsequent Wittig reaction and addition of aryl thiolates. Cleavage was performed with formic acid in THF.



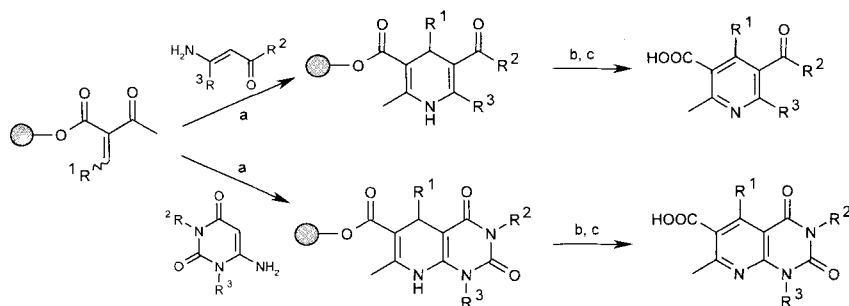
**Figure 6.20.** Key: (a) arylthiol, NaOMe (cat.), THF, rt, 2 d; (b) formic acid, THF, rt, 2 h.

Nine compounds were synthesized via the 'split-mix' method. The compounds were characterized by using GC-MS, HR-MS and by  $^1\text{H}$ - and  $^{13}\text{C}$  NMR, and the isolated yields ranged from 7% to 27%.

### 6.5.3 Synthesis of Pyridines and Pyrido[2,3-d]pyrimidines

Three different pyridine syntheses starting from polymer-bound enones have been reported. The first synthesis starts from 2-alkylidene- or 2-arylidene- $\beta$ -ketoesters immobilized on Wang or Sasrin resin (Fig. 6.21). These substrates reacted with enaminones in a Hantzsch reaction to 1,4-dihydropyridines which could be oxidized to the corresponding pyridines with ceric ammonium nitrate (CAN). Cleavage was performed with TFA/DCM. Sixteen compounds were synthesized, with HPLC purities of between 70% and 99%.

By using 6-aminouracils as enaminones, pyrido[2,3-d]pyrimidines were obtained. Four of these compounds were synthesized, with excellent purities (90–100% in HPLC) (Fig. 6.21).

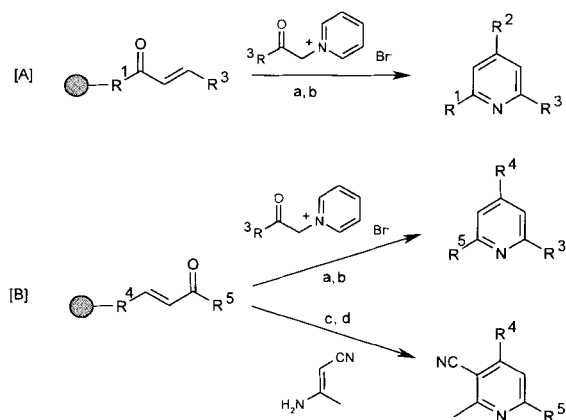


**Figure 6.21.** Key: (a) enaminone, DMF, 80 °C; (b) ceric ammonium nitrate (CAN), DMA, rt, 15 min; (c) 95% aq. TFA or 3% TFA/CH<sub>2</sub>Cl<sub>2</sub> (for Wang or Sasrin resin).

Grosche et al. [33] established the classic Kröhnke pyridine synthesis on the solid phase (Fig. 6.22 (A, B a, c)). The polymer-bound enones used in this reaction were obtained from polymer-bound acetophenones as well as polymer-bound aldehydes, thus providing enones with different substitution patterns (Fig. 6.19 (B and C)).

These enones reacted with substituted phenacyl pyridinium bromides and ammonium acetate under pyridine ring formation. Cleavage was performed with 25% TFA/CH<sub>2</sub>Cl<sub>2</sub>. Twelve compounds were synthesized, with yields of the crude products ranging from 50% to 85% (HPLC-purities: 76–97%).

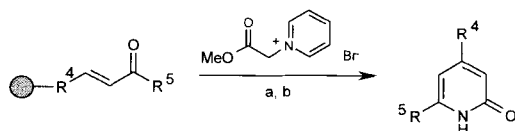
Marzinzik and Felder [34] synthesized a 3-cyanopyridine via the reaction of a polymer-bound enone with 3-amino-crotonic acid nitrile derived in a Thorpe reaction from acetonitrile and potassium *tert*-butoxide (B b, d Fig. 6.22) [26]- The product was obtained in 46% yield, and with a HPLC purity of 78%.



**Figure 6.22.** Key: (a) phenacylpyridinium bromide,  $\text{NH}_4\text{OAc}$ , DMF, HOAc,  $90^\circ\text{C}$ , 24 h; (b) 25% TFA/ $\text{CH}_2\text{Cl}_2$ , rt, 1 h; (c) *tert*-BuOK,  $\text{CH}_3\text{CN}$ , rt, 16 h; (d) 20% TFA/ $\text{CH}_2\text{Cl}_2$ , rt, 15 min.

### 6.5.4 Synthesis of Pyridones

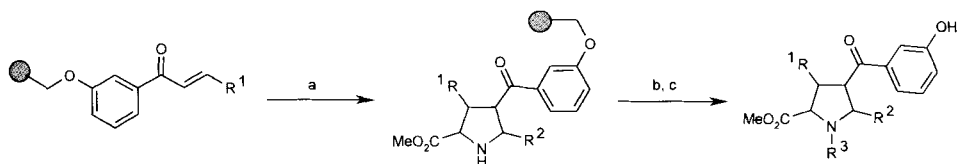
By using 1-[methoxycarbonylmethyl]-pyridinium bromide in the solid-phase Kröhnke reaction, pyridones were obtained [33]. Six compounds were synthesized, with HPLC-purities between 68% and 76% ( $\lambda=214$  nm), and yields of the crude products ranging from 71% to 78% (Fig. 6.23).



**Figure 6.23.** Key: (a) 1-[methoxycarbonylmethyl]pyridinium bromide,  $\text{NH}_4\text{OAc}$ , DMF, HOAc,  $90^\circ\text{C}$ , 24 h; (b) 25% TFA/ $\text{CH}_2\text{Cl}_2$ , rt, 1 h.

### 6.5.5 Synthesis of Pyrrolidines

Hollinshead [32] used polymer-bound enones for the synthesis of highly functionalized pyrrolidines. 3-Hydroxyacetophenone was immobilized on chlorinated Wang resin and transformed into polymer-bound enones upon a Knoevenagel reaction with aldehydes. Pyrrolidines were then formed in the addition of azomethinylides, generated from imines, LiBr and DBU (Fig. 6.24).

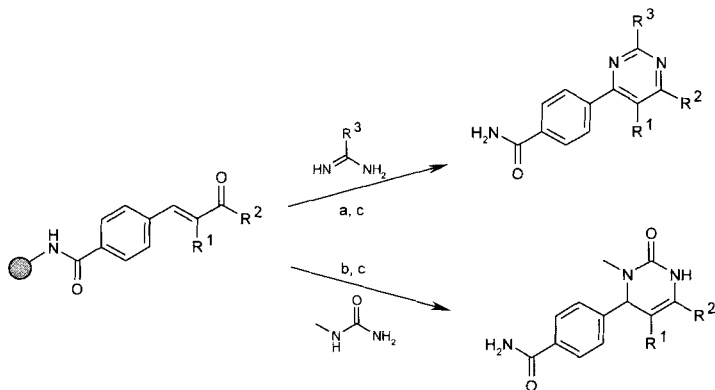


**Figure 6.24.** Key: (a)  $R^2CH=NCH_2CO_2Me$ , LiBr, DBU, THF, rt, 3 d; (b) acetyl chloride or sulfonyl chloride, DMAP, pyridine,  $CH_2Cl_2$ , rt, 20 h; (c) 50% TFA/ $CH_2Cl_2$ , rt, 20 h.

The secondary amino function was modified with acid chlorides and sulphonyl chlorides, and the final products were cleaved from the resin with 50% TFA. Four pyrrolidines were synthesized, with isolated yields between 31% and 68%.

### 6.5.6 Synthesis of Pyrimidines and Pyrimidones

Marzinzik and Felder [34] used polymer-bound enones which were generated from 4-carboxybenzaldehyde immobilized on Rink amide resin for the synthesis of pyrimidines and pyrimidones. The enones were then treated with different amidines in DMA for 16 h at 100 °C (Fig. 6.25).

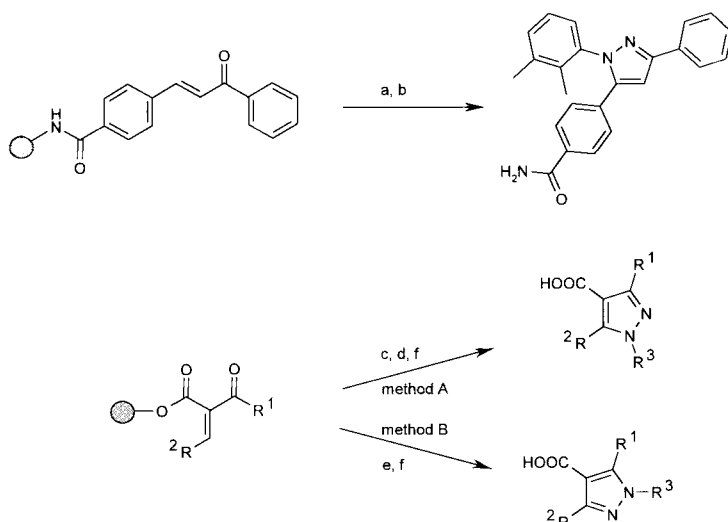


**Figure 6.25.** Key: (a) amidine, DMA, air, 100 °C, 16 h; (b) *N*-methylurea, NaOEt, DMA, rt, 16 h; (c) 20% TFA/ $CH_2Cl_2$ , rt, 15 min.

The isolated yields of the pyrimidines ranged from 58% to 98%, with HPLC purities of >95%. By using *N*-methylurea in the presence of NaOEt, a dihydropyrimidone with a purity of 94% and an isolated yield of 84% was obtained as a single regioisomer (Fig. 6.25b, c).

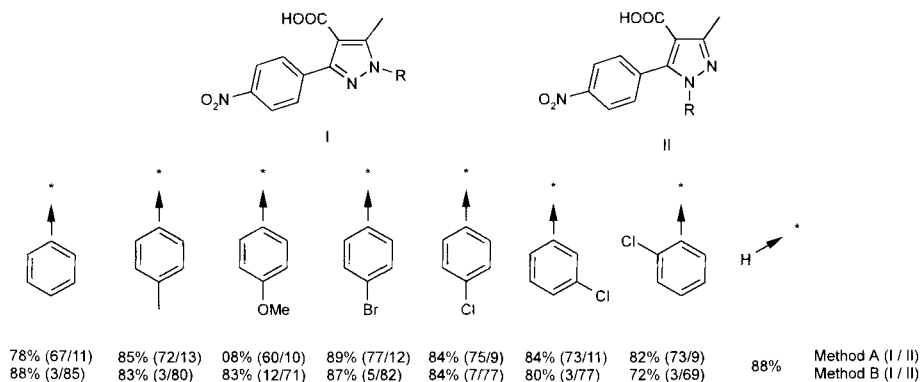
## 6.5.7 Synthesis of Pyrazoles

Marzinik and Felder [34] also performed a pyrazole synthesis as a possible transformation of polymer-bound enones. A single enone was treated with 2,3-dimethylphenylhydrazine under regioselective formation of the *N*-phenylpyrazole shown in Figure 6.26 (a, b). The crude product had a purity of 83%, and the isolated yield was 73%.



**Figure 6.26.** Key: (a) 2,3-dimethylphenylhydrazine, DMSO, 100 °C, 16 h; (b) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>, rt, 15 min; (c) phenylhydrazine hydrochloride, DMF, AcOH, rt; (d) base, DMF, 70 °C, 4 h; (e) phenylhydrazine hydrochloride, base, NMP, 70 °C, 4 h; (f) 25% TFA/CH<sub>2</sub>Cl<sub>2</sub>.

In our group [37], a pyrazole synthesis was performed on acetoacetylated Wang resin. A selective formation of the regioisomers was possible by simply varying the reaction conditions (Fig. 6.26, method A, method B). Scheme 6.13 shows some representative results of this synthesis.



**Scheme 6.13.** Representative results of the pyrazole synthesis (Fig. 6.26, method A, method B), together with HPLC purities ( $\lambda = 214$  nm).

Polymer-bound enones were also used in a hetero Diels–Alder reaction for the synthesis of dihydropyranes [38, 39].

## 6.6 Conclusions

In summary, we have shown that squaric acid, 3-hydroxy-2-methylidene propionic acids, 5-(2-bromoacetyl)pyrroles and enones are useful polymer-bound key intermediates for the synthesis of a large number of different core structures. Squaric acid was used as a fluid template, because cores structures with different ring sizes could be synthesized. All other examples started from linear templates and afforded linear core structures as well as cyclic core structures. The examples shown here demonstrate the advantages of polymer-bound templates for this type of synthesis. This strategy also reduces the optimization time needed for developing the synthetic route for a specific structure because the synthesis of the polymer-bound educt has to be evaluated only once for the variety of core structures derived from this educt.

Several other promising templates suitable for the synthesis of multiple core structure libraries have been reported, but their synthetic potential has not yet been fully investigated.

Further examples are epoxides, from which oxazolidinones [40] and  $\gamma$ - and  $\delta$ -lactones [41] have been synthesized, 1,3-diketones which were used in the syntheses of pyrazoles and isoxazoles [42], 1,2-aminothiols from which thiazolidines [43] were formed.

## 6.7 References

- [1] See recent reviews: Booth, S., Hermkens, P.H.H., Ottenheijm, H.C.J., Rees, D.C. *Tetrahedron*, **1998**, *54*, 15385–15443; Hermkens, P.H.H., Ottenheijm, H.C.J., Rees, D.C. *Tetrahedron*, **1997**, *53*, 5643–5678; Brown, R. *Contemp. Org. Syn.*, **1997**, 216–237; Früchtel, J.S., Jung, G. *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 17–42; Hermkens, P.H.H., Ottenheim, H.C.J., Rees, D. *Tetrahedron*, **1996**, *52*, 4527–4554; Balkenhohl, F., von dem Bussche-Hünnefeld, C, Lansky, A., Zechel, C. *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 2288–2337.
- [2] Jung, G. (ed.), *Combinatorial Peptide and Nonpeptide Libraries: A Handbook*. VCH, Weinheim; New York, **1996**; Bunin, B.A. (ed.), *The Combinatorial Index*. Academic Press, San Diego, **1998**; Gordon, E.M., Kerwin, J.F.J., *Combinatorial Chemistry and Molecular Diversity in Drug Discovery*, John Wiley & Sons, New York, **1998**; Terrett, N.K., *Combinatorial Chemistry*. Oxford University Press, New York, **1998**; Moos, W.H., Pavia, M.R., Kay, B.K., Ellington, A.D., *Annual Reports in Combinatorial Chemistry and Molecular Diversity*. ESCOM, Leiden, **1997**.
- [3] De Witt, S.H., Kiely, J.S., Stankovic, C.J., Schröder, M.C., Cody, D.M.R., Pavia, M.R. *Proc. Natl. Acad. Sci.*, **1993**, *90*, 6909.
- [4] Matthews, J., Rivero, R.A. *J. Org. Chem.*, **1997**, *62*, 6090.
- [5] Kim, S.W., Ahn, S.Y., Koh, J.S., Lee, J.H., Ro, S., Cho, H.Y. *Tetrahedron Lett.*, **1997**, *37*, 7489.
- [6] Hanessian, S., Yang, R.J. *Tetrahedron Lett.*, **1996**, *37*, 5835.
- [7] Short, K.M., Ching, B.W., Mjalli, A.M.M. *Tetrahedron Lett.*, **1996**, *37*, 7489–7492.
- [8] Bauser, M., Winter, M., Valenti, C.A., Wiesmüller, K.H., Jung, G. *Mol. Diversity*, **1998**, *3*, 257–260.
- [9] Bauser, M., Winter, M., Betschinger, J., Jung, G., poster presented at the *International Symposium Speeding up Drug Discovery*, Schliersee, Oct. 21–23, **1998**.
- [10] Bhalay, G., Cowell, D., Hone, N.D., Scobie, M., Baxter, A.D. *Mol. Diversity*, **1998**, *3*, 195–198.
- [11] Tempest, P.A., Armstrong, R.W. *J. Am. Chem. Soc.*, **1997**, *119*, 7607–7608.
- [12] Frenette, R., Friesen, R.W. *Tetrahedron Lett.*, **1994**, *35*, 9177–9180; Brown, S.D., Armstrong, R.W. *J. Am. Chem. Soc.*, **1996**, *118*, 6331–6332.

- [13] Goff, D.A., Zuckermann, R.N. *J. Org. Chem.*, **1995**, *60*, 5748–5749; Yun, W., Mohan R. *Tetrahedron Lett.*, **1996**, *37*, 7189–7192; Koh, J.S., Ellman, J.A. *J. Org. Chem.*, **1996**, *61*, 4494–4495.
- [14] Deshpande, M.S. *Tetrahedron Lett.*, **1994**, *35*, 5613–5614.
- [15] Kurth, M.J., Ahlberg Randall, L.A., Chen, C., Melander, C., Miller, R.B. *J. Org. Chem.*, **1994**, *59*, 5862–5864; Kobayashi, S., Hachiya, I., Suzuki, S., Moriwaki, M. *Tetrahedron Lett.*, **1996**, *37*, 2809–2812.
- [16] Gordeev, M.F., Patel, D.V., Wu, J., Gordon, E.M. *Tetrahedron Lett.*, **1996**, *37*, 4643–4646; Tietze L.F., Steinmetz, A. *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 651–652.
- [17] Chen, C., Ahlberg Randall, L.A., Miller, R.B., Jones, A.D., Kurth, M.J. *J. Am. Chem. Soc.*, **1994**, *116*, 2661–2662.
- [18] Prien, O., Rölfling, K., Thiel, A., Künzer, H. *Synlett*, **1997**, 325–326.
- [19] Richter, H., Jung, G. *Mol. Diversity* **1998**, *3*, 191–194.
- [20] Richter, H., Jung, G. *Tetrahedron Lett.* **1998**, *39*, 2729–2730.
- [21] Perlmutter, P., Puniani, E., Westman G. *Tetrahedron Lett.*, *37* (1996) 1715–1718.
- [22] Richter, H., Walk, T., Hölzel, A., Jung, G. *J. Org. Chem.*, **1999**, *64*, 1362–1365.
- [23] Richter, H., Walk, T., Hölzel, A., Jung, G., poster presented at the ORCHEM 98, Bad Nauheim, Germany, **1998**.
- [24] Richter, H., Hölzel, A., Jung, G. in R. Epton (Ed.), *Innovation and Perspectives in Solid Phase Synthesis and Combinatorial Libraries 1999*, Mayflower Scientific Limited, Birmingham, in press.
- [25] Mukaiyama, T., Hoshino, T. *J. Am. Chem. Soc.* **1960**, *82*, 5339–5340.
- [26] Mitsunobu, O. *Synthesis* **1981**, 1–28.
- [27] Tietze, L.F., Steinmetz, A. *Synlett* **1996**, 667–668.
- [28] Richter, H., Walk, T., Hölzel, A., Jung, G., poster presented at the *International Symposium Speeding up Drug Discovery*, Schliersee, Oct. 21–23, **1998**.
- [29] Trautwein, A.W., Hölzel, A., Jung, G. poster presented at the *International Symposium Speeding up Drug Discovery*, Schliersee, Oct. 21–23, **1998**.
- [30] Trautwein, A.W., Süßmuth, R.D., Jung, G. *Bioorg. Med. Chem. Lett.*, **1998**, *8*, 2381–2384.
- [31] Tietze, L.F., Hippe, T., Steinmetz, A., *Synlett*, **1996**, 1043–1044;
- [32] Hollinshead, S.P., *Tetrahedron Lett.*, **1996**, *51*, 9157–9160.
- [33] Grosche, P., Hölzel, A., Walk, T., Trautwein, A.W., Jung, G. *Synthesis*, **1999**, in press.
- [34] Marzinzik, A.L., Felder, E.R., *J. Org. Chem.*, **1998**, *63*, 723–727.
- [35] Chen, C., Ahlberg Randall, L.A., Miller, R.B., Jones, A.D., Kurth, M.J., *J. Am. Chem. Soc.*, **1994**, *116*, 2661–2662.
- [36] Chen, C.X., Randall, L.A.A., Miller, R.B., Jones, A.D., Kurth, M.J., *J. Am. Chem. Soc.*, **1994**, *116*, 2661–2662; Chen, C.X., Randall, L.A.A., Miller, R.B., Jones, A.D., Kurth, M.J., *Tetrahedron*, **1997**, *53*, 6595–6609.
- [37] Grosche, P., Hölzel, A., Jung, G., poster presented at the ORCHEM 98, Bad Nauheim, Germany, **1998**.
- [38] Tietze, L.F., Hippe, T., Steinmetz, A., *Synlett*, **1996**, 1043–1044
- [39] Graf von Roedern, E. *Mol. Diversity*, **1998**, *3*, 253–256.
- [40] Buchstaller, H.-P. *Tetrahedron*, **1998**, *54*, 3465–3470.
- [41] Le Hetet, C., David, M., Carreaux, F., Carboni, B., Sauleau, A. *Tetrahedron Lett.*, **1997**, *38*, 5153–5156.
- [42] Marzinzik, A.L., Felder, E.R. *Tetrahedron Lett.*, **1996**, *37*, 1003.
- [43] Patek, M., Drake, B., Lebl, M. *Tetrahedron Lett.*, **1995**, *36*, 2227–2230.

# 7 Combinatorial Synthetic Oligomers

*Alberto Bianco*

## 7.1 Introduction

During the recent years, the search for new synthetic oligomers, in particular those defined as peptidomimetics, has increased enormously. Considerable efforts have been made to synthesize sophisticated polymeric frameworks with new chemical, conformational and biological properties, and moreover possessing improved pharmacological features relative to those of peptides [1–3]. Biologically active peptides are highly valuable lead structures for drug discovery. However, although they possess a high ability to bind selectively and either activate or inhibit a large number of biological receptors, their potential use as therapeutic agents is limited by their physico-chemical properties. For example, peptides are characterized by low bioavailability and rapid enzymatic digestion [4]. Therefore, many interesting amide bond modifications have been successfully introduced in the peptide backbone, sometimes leading to active structures where the labile peptide bond has been completely replaced by surrogates. Many different synthetic oligomers have become more easily accessible due to the development of the combinatorial solid-phase procedures [5].

In the design of new peptidomimetics, the physico-chemical nature of the polymer backbone and the side chains are of fundamental importance. Side chain functionalities should guarantee electrostatic, hydrophobic,  $\pi$ -stacking and hydrogen bonding interactions. New degrees of freedom in the conformation of the backbone should be desirable for enhanced ligand–receptor association. From a chemical point of view, the most significant limitation to the overall design is the requirement for an efficient, high-yielding coupling reaction on solid support. This issue has been overcome improving the highly efficient methods developed for the synthesis of peptides [6]. Moreover, the automated and combinatorial methods for the rapid generation of peptide libraries have been extended to the new peptidomimetics in order to create diverse compound collections containing potential lead structures that bind to a specific target such as an enzyme or a receptor [5].

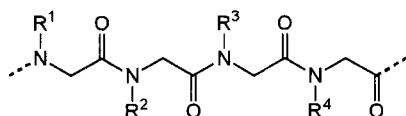
This chapter aims to present an overview on the main peptide bond modifications which have been proposed in the recent years. The solution- and solid-phase synthesis of the new synthetic oligomers, the methods for their characterization, and their biological and folding properties will be discussed. Particular emphasis will be given to those modifications which are suitable for combinatorial chemistry applications and developments.



## 7.2 Peptoids

### 7.2.1 Submonomer Approach to Peptoid Synthesis

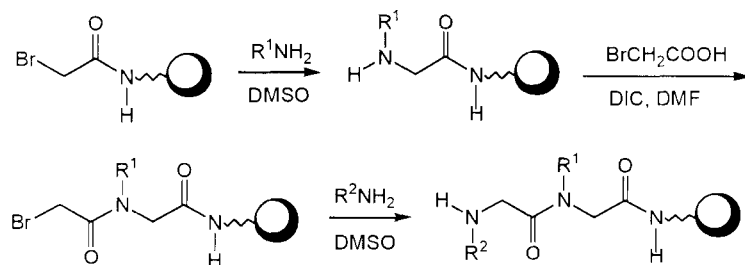
Peptoids are a new class of synthetic oligomers made up of N-substituted glycine (NSG) monomers, which can be assembled on a solid support using fully automated procedures (Fig. 7.1). The side chains of the amino acids are in this case shifted from the C $\alpha$  carbon to the amide nitrogen. Moreover, a large variety of N-alkylations can be introduced, generating a much higher degree of molecular diversity compared to the natural amino acids.



**Figure 7.1.** Molecular structure of a peptoid fragment, which constitutes a N-substituted oligoglycine (NSG).

Although the nature of the peptoid backbone is quite similar to the nature of the peptide backbone, these new unnatural oligomers display some peculiar characteristics: (i) the peptoid backbone is achiral; (ii) peptoids are devoid of amide protons, which decrease their polarity and should increase their oral bioavailability; (iii) they lack the NH-donor bond and, therefore, they do not have the possibility to form intra- or intermolecular H-bonding through the backbone; (iv) they can be more flexible and adopt altered conformations; and (v) they contain only tertiary amide bonds, which display a higher protease stability [7].

Zuckermann et al. [8] developed a successful protocol, defined ‘submonomer approach’, for the solid-phase synthesis of peptoids, starting from the readily available building blocks bromoacetic acid and primary amines (Scheme 7.1).

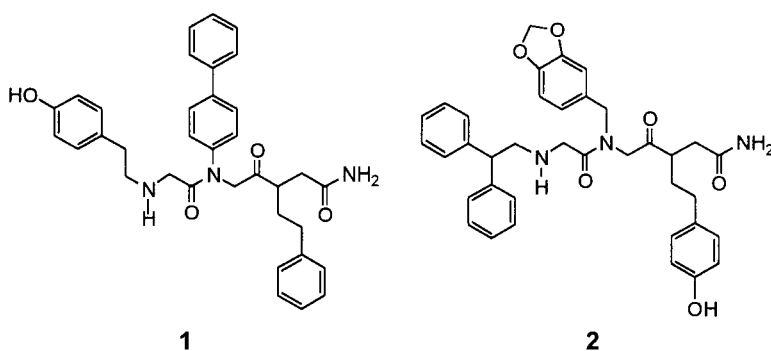


**Scheme 7.1.** Synthesis of a NSG peptoid dimer using the ‘submonomer method’ [8] on a solid support carrying amide linker.

The NSG oligomers derive from an alternated sequence of acylation reactions and S<sub>N</sub>2-reactions which allowed to obtain homo- and heteropentamers in high yield and purity,

using a large variety of commercially available amines. The efficiency of the chemistry and the molecular diversity of the monomers that can be incorporated render the peptoid synthesis ideal for the automated generation of combinatorial libraries.

This highly efficient synthetic route to NSGs was used, for example, to design a biased peptoid library for the binding to 7-transmembrane/G-protein-coupled receptors [9]. Through the screening of a mixture of ca. 5000 dimer and trimer NSGs, a series of nanomolar ligands were discovered (Fig. 7.2). Compound **1** is a potent ligand for the  $\alpha_1$ -adrenergic receptor and it shares few structural elements with the endogenous ligand epinephrine and norepinephrine and with prazosin. Compound **2** inhibits the binding of the two natural ligands morphine and Met-enkephalin to the  $\mu$ -opioid receptor.



**Figure 7.2.** Molecular structures of two peptoid ligands for 7-transmembrane/G-protein-coupled receptors [9].

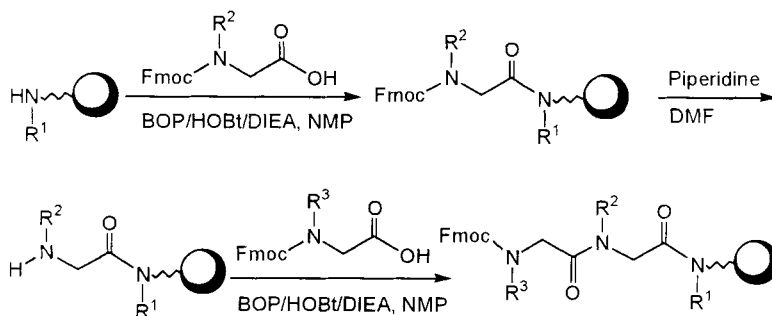
Recently, a combinatorial library of cationic peptoids has been used to discover reagents for gene delivery. A series of NSG oligomers, up to 36 residues in length, were shown to condense plasmid DNA into small peptides, to protect it from nuclease degradation, and to mediate the transfection of several cell lines [10].

These examples show how the combinatorial peptoid approach holds enormous potential for the discovery of novel lead structures for drug development. In this context a post-modification on solid-phase of peptoid side chains, bearing alkene and alkyne moieties, via [3 + 2] cyclo-addition of nitrile oxides, enhances further the molecular diversity and the possibility to find new therapeutic agents [11]. Moreover, polymers of N-substituted glycines containing chiral side chains display interesting conformational preferences [12]. NMR and CD data indicate that the major species adopts in methanol a stable right-handed helical structure with *cis*-amide bonds.

### 7.2.2 Monomer Approach to Peptoid Synthesis

In contrast to the relative simplicity of the 'submonomer method', which can be fully automated using commercially available building blocks, a different approach to generate peptoids is based on the use of preformed Fmoc-N-protected NSG residues. The synthe-

sis of peptoids, using this strategy, was already addressed in 1992 by Simon et al. [13], and recently it has been reinvestigated by the group of Liskamp (Scheme 7.2) [14].



**Scheme 7.2.** Synthesis of a resin-bound NSG peptoid trimer using the ‘monomer approach’ with pre-formed N-substituted N-Fmoc-glycines [13, 14].

The ‘submonomer method’ is suitable when small amounts of peptoids are needed, but it suffers from the drawback that large excesses of reagents have to be used and that each reaction step cannot be monitored. The ‘monomer approach’ can be easily scaled-up to obtain larger amounts of peptoids of reasonable purity. As in the SPPS, the progress of each coupling step can be followed by quantifying the dibenzofulvene adduct during the Fmoc removal. It is also possible to prepare peptomers, peptide–peptoid hybrids (see Section 7.2.3) without modifying the automated peptide synthesis protocols. The NSG monomers can be obtained by alkylation of primary amines with electrophiles, such as haloacetic acids or acrylamides, or by reductive amination of a glycine residue with aldehydes. The following step is to protect the secondary amino function with Fmoc group.

The Fmoc-protected monomers have been used to prepare both the Leu-enkephalin and the substance P peptoid and retropeptoid analogues [14]. It has been found that the substance P peptoid and retropeptoid are agonists, with an activity in micro to submicro concentrations. Moreover, these analogues are resistant to protease degradation. It has also been shown that the peptoids can be easily characterized by mass spectrometry, studying the fragmentation pattern [14, 15], and by sequencing using the Edman degradation method, which permit a direct identification of the resin-bound oligomers [16].

In conclusion, using this ‘monomer approach’ it is in principle possible to translate any peptide into the corresponding peptoid using a fully automated synthesizer, the final goal being the creation of combinatorial oligomeric libraries for structure–activity relationship studies and drug discovery.

### 7.2.3 Peptomers

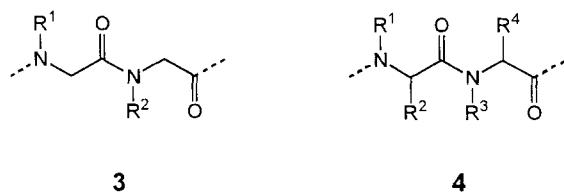
The peptomer approach is based on the preparation of peptide–peptoid hybrids. The method has been already applied to generate in solution this new class of peptidomimetics. Goodman and co-workers recently incorporated peptoid residues (NGS) in collagen-

based sequences [17] and in a selective somatostatin analogue [18]. These new oligomers are able either to exhibit a high triple-helix propensity in the case of the collagen analogue or to increase selectivity for a specific receptor and to inhibit *in vivo* the release of growth hormone without any effect on the inhibition of insulin in the case of somatostatin derivative.

Østergaard and Holm [19] extended the concept of peptomers to their synthesis on solid support, accomplished by simply combining the conventional peptide synthesis with the methodology developed for the preparation of peptoids ('submonomer approach'). A versatile bead library of peptomers has been synthesized and screened. The complexity and diversity of such peptidomimetic bead libraries, derived from the incorporation of  $\alpha$ -amino acids and different amines, can be very high. It should be emphasized that the sequence of the active single bead can be determined by standard Edman degradation sequencing, avoiding the co-synthesis of a coding structure or the introduction of molecular tags [20]. With this combinatorial method, novel ligands to the targets streptavidin and insulin receptor have been identified.

#### 7.2.4 Peralkylated Peptides

A further amide bond modification, which has been introduced some years ago, is the alkylation of the amide nitrogens all along the peptide backbone [21]. This approach permits the construction of oligomeric peptidomimetics comparable to peptomers, in which the N-substituted glycines are replaced by N-substituted  $\alpha$ -amino acids (Fig. 7.3). We could define these oligomers C $^{\alpha}$ -branched peptoids.

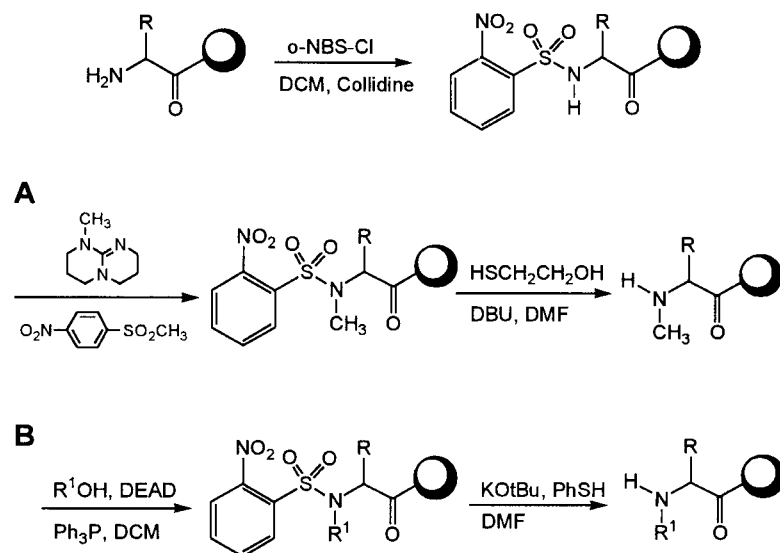


**Figure 7.3.** Comparison between the molecular structures of a peptoid fragment (3) and of a peralkylated peptide fragment (4).

At the beginning, permethylated peptide libraries were derived by a post-modification of peptide libraries attached to the solid-phase support. The amide bond functionalities were alkylated using a solution of sodium hydride in dimethylsulfoxide, followed by addition of neat methyl iodide, therefore generating 'libraries from libraries'. A peptide library that underwent this simple chemical transformation was used to identify compounds with potent antimicrobial activity against Gram-positive bacteria [21].

A further improvement of this approach, to create increased molecular diversity, was applied to the synthesis of a soluble peptidomimetic combinatorial library of about 60 000 compounds [22]. This library generated using the 'split and combine method' is based on a dipeptide scaffold, which incorporates 50 different L-, D- and unnatural amino acids,

and in which each amide NH has been replaced by five different alkyl groups. The two methods of alkylating the backbone NH reported above are non-selective, since it is not possible to modify one specific NH amide bond without involving all the others present in the peptide. Chemoselective N-alkylation has been recently achieved on solid support using two different strategies [23, 24]. One procedure comprises: (i) protection of the free amino group with *o*-NBS-Cl (*o*-nitrobenzenesulfonyl chloride); (ii) selective deprotonation of the sulfonamide with MTBD (7-methyl-1,5,7-triazabicyclo[4.4.0]dec-5-ene) and methylation; and (iii) selective cleavage of the *o*-NBS group with 2-mercaptoethanol and DBU (1,8-diazabicyclo[5.4.0]undec-7-ene) (Scheme 7.3A) [23].



**Scheme 7.3.** Selective N-alkylation of a peptide bond: methylation (A) and alkylation via Mitsunobu reaction (B) of resin-bound *o*-NBS-amino acids [23, 24].

Reichwein and Liskamp [24] developed another route using a Mitsunobu reaction instead of the selective sulfonamide deprotection with a strong base (Scheme 7.3B). To show the versatility of this solid-phase procedure, a N-ethyl scan of Leu-enkephalin has been performed.

Chemoselective N-alkylation on solid support, using these two different synthetic ways, may be incorporated in combinatorial strategies to produce large libraries of NH amide modified peptides, employing a variety of alkylating agents.

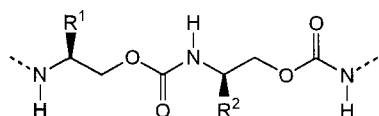
### 7.2.5 $\beta$ -Peptoids

In early 1998, a further class of peptoid oligomers appeared [25], when a solid-phase method was developed to synthesize N-substituted  $\beta$ -aminopropionic acid oligomers, termed  $\beta$ -peptoids. Treatment of Wang's resin with acryloyl chloride, followed by Michael addition of primary amines afforded N-substituted  $\beta$ -alanines. A combinatorial library of

defined mixtures of tri- $\beta$ -peptoids was prepared by mixing equimolar amounts of resins, loaded with different substituted  $\beta$ -alanines on which two sequences of acylation-Michael additions were performed [25].

### 7.3 Oligocarbamates

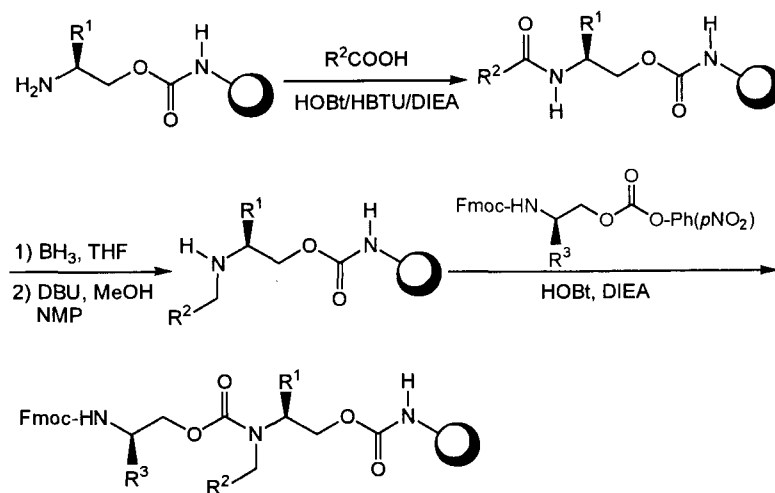
Sequential oligocarbamates were synthesized by Schultz and co-workers [26, 27] via activated Fmoc-protected chiral amino carbonate monomers generating a carbamate backbone (Fig. 7.4).



**Figure 7.4.** Molecular structure of an oligocarbamate fragment.

The N-Fmoc-protected *p*-nitrophenyl carbonate monomers are accessible from the corresponding  $\beta$ -amino alcohols. Solid-phase synthesis of oligocarbamates is very efficient, with a reported coupling yield of 99% per step.

The  $\alpha$ -carbon of the oligocarbamate backbone, like that of peptides, is substituted with side chains containing a variety of functional groups. A further backbone modification and conformational restriction can be incorporated via N-acylation followed by reduction of the N-acyl nitrogen, yielding N-alkyl-oligocarbamates (Scheme 7.4) [28].



**Scheme 7.4.** Solid-phase synthesis of N-alkyl oligocarbamates using activate carbonates of N-Fmoc- $\beta$ -amino alcohols and carboxylic acids as educts [28].

Recently, carbamate derivatives have been studied for their ability to form H-bonds. The *syn*-rotamers (*cis*) have been shown to be stabilized in solution by the formation of H-

bond complexes with carboxylic acids [29]. The N-alkylation of oligocarbamates increases the variety of side chains and removes the possibility of hydrogen bonding formation.

An oligocarbamate was designed to bind TAR RNA (*trans*-activation responsive region). The sequential oligomer derived from the RNA-binding Tat peptide, which is involved in the mechanism of *trans*-activation of human immunodeficiency virus type I (HIV-1), can specifically bind TAR RNA with high affinity, forming a complex stable to proteolytic digestion [30].

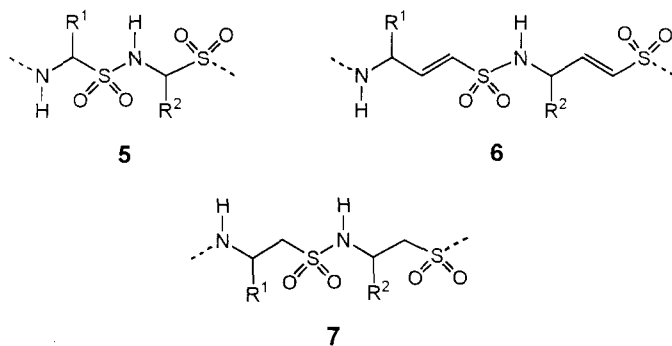
A library of oligocarbamates was screened for its capacity to bind to a monoclonal antibody. Inside this pool, high-affinity ligands ( $IC_{50}$  60–180 nM) have been detected and isolated [26]. These compounds display either a good enzymatic resistance or a higher hydrophobicity than peptide analogues.

Recently, oligocarbamates have been used as new ligands for major histocompatibility complex (MHC) class I molecules. A series of pentamers, based on the structure of the natural T-cell epitope SIINFEKL, have been prepared using a flexible automated synthesis [31]. The ligand binding with the highest affinity shows a good structural similarity with the model octapeptide.

Moreover, the synthesis of cyclic oligocarbamates was successfully achieved [32]. First, the linear oligomers, three to six residues in length, were assembled on a solid support. Then, they were cyclized in solution. The resulting constrained conformation and the lipophilic nature of the backbone make the cyclooligocarbamates very interesting for lead structure search in drug discovery.

## 7.4 Sulfonopeptides and Vinylogous-Sulfonamidopeptides

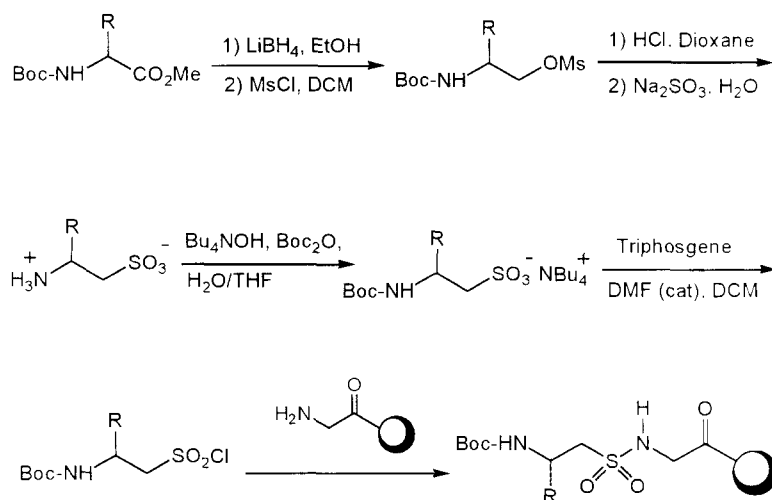
Sulfonamidopeptides can be considered as interesting candidates to find new protease inhibitors and new drugs, since the structure of the sulfonamide bond is similar to the tetrahedral transition state involved in the amide bond hydrolysis by proteolytic enzymes [33]. In the field of sulfonamide-containing peptidomimetics, several scaffolds have been proposed (Fig. 7.5). While the preparation of  $\alpha$ -sulfonopeptides **5** has remained elusive [34, 35], probably due to synthetic problems, such as  $\beta$ -elimination of  $SO_2$ , the other two types of oligomers turned out to be easily obtained.



**Figure 7.5.** Molecular structures of three different fragments of sulfonamide-based oligomers. **5**,  $\alpha$ -sulfonopeptide; **6**, vinylogous sulfonopeptide; **7**,  $\beta$ -sulfonopeptide.

Vinylogous sulfonopeptides **6** (vs-peptides) have been synthesized both in solution and solid phase [34]. They present a secondary structure stabilized by hydrogen bonding in a 12-membered ring in the crystal structure and a 14-membered ring in chloroform solution [36]. Tweezer-like molecular receptors based on vs-peptides bind to an encoded combinatorial tripeptide library with a selectivity compared to that of synthetic receptors built up with  $\alpha$ -amino acids [37, 38]. Combinatorial vs-peptide libraries have been prepared on solid phase by Gennari and co-workers [39] using the 'split-mix' method. The authors were also able to post-modify these oligomers by selective N-alkylation of the acidic sulfonamide proton, in order to increase the molecular diversity.

The third class of sulfonamide-containing peptide is that of  $\beta$ -sulfonopeptides **7**. A new method for their solid-phase synthesis has been reported (Scheme 7.5) [40]. N-Boc-protected sulfonyl chloride monomers, derived from Boc-amino acid methyl esters, must be prepared in solution and these educts are then coupled to resin-bound amino acids.



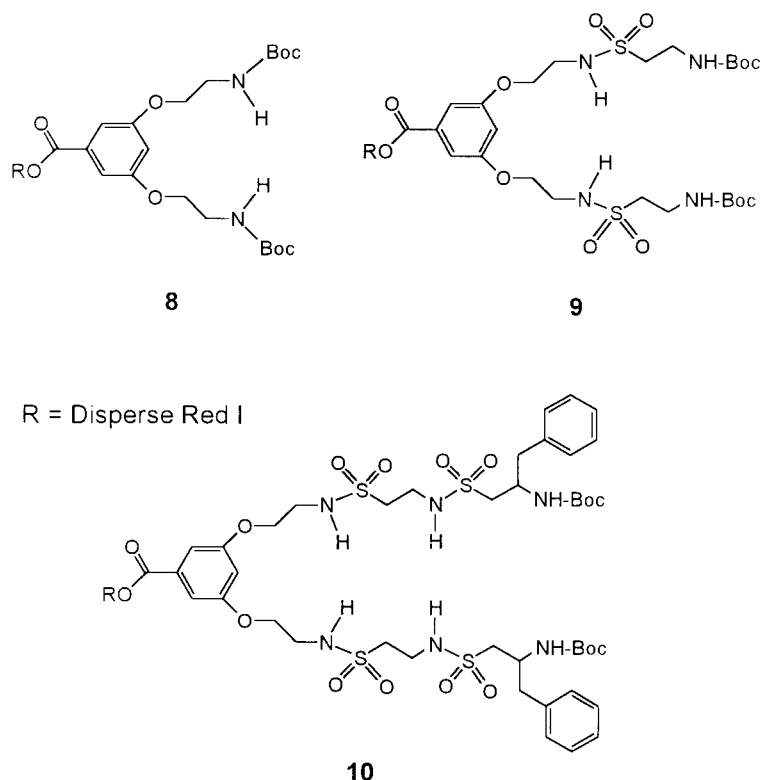
**Scheme 7.5.** Synthesis of a  $\beta$ -sulfonopeptide [40].

$\beta$ -Sulfonopeptides, as the vs-peptides, show a strong tendency to form well-defined folded structures via intramolecular hydrogen bonding. Although many hydrogen bonding possibilities are present, a series of N-terminal carbamate or carboxyamide  $\beta$ -sulfonodipeptide amides showed a strong preference for a 12-membered ring intramolecular hydrogen bond stabilization, involving the N-terminal carbonyl as acceptor [41].

Synthetic  $\beta$ -sulfonopeptide-based receptors **9** and **10** were prepared using the diamino acid derivative **8** (Fig. 7.6) and screened for binding against an encoded library containing ca. 25 000 tripeptides [42].

While the tweezer **9** did not show any interaction, binding against the deprotected peptide library was observed with the tweezer **10**. In order to extend the binding capacity eventually to compounds other than peptides, tweezer-like receptors consisting of  $\beta$ -sulfonopeptides were prepared incorporating less flexible hinges [43]. Varying the size of the



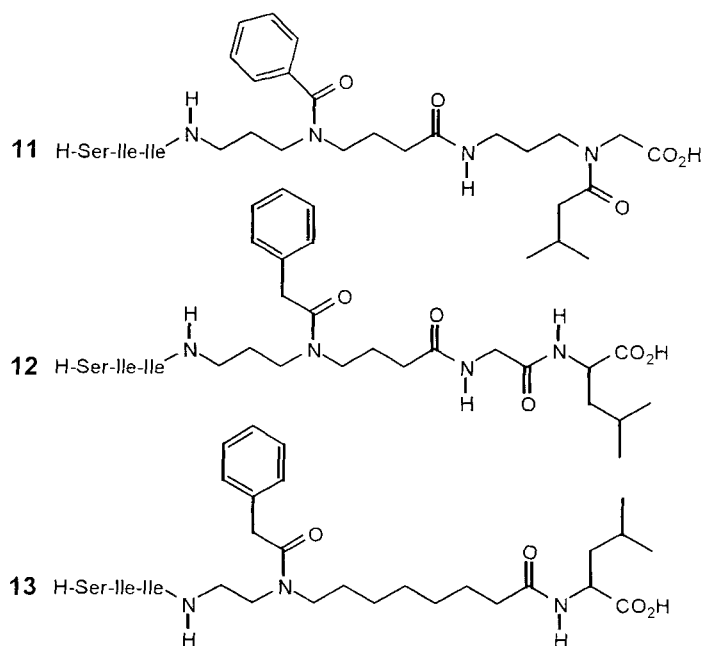


**Figure 7.6.** Molecular structures of the precursor (**8**) and the tweezer peptide library receptors (**9**, **10**) based on  $\beta$ -sulfonopeptides.

hinge part, these receptors display a high binding affinity towards a tripeptide library while maintaining a good selectivity. The concept of tweezer-like synthetic receptors based on peptidosulfonamides can be extended to the preparation of combinatorial libraries of these molecules. This may offer attractive tools to study the structure and interactions of peptidomimetics with peptides, proteins and small molecules. For combinatorial synthetic receptors, see also Chapter 10 in this book.

## 7.5 Poly-N-Acylated Amines

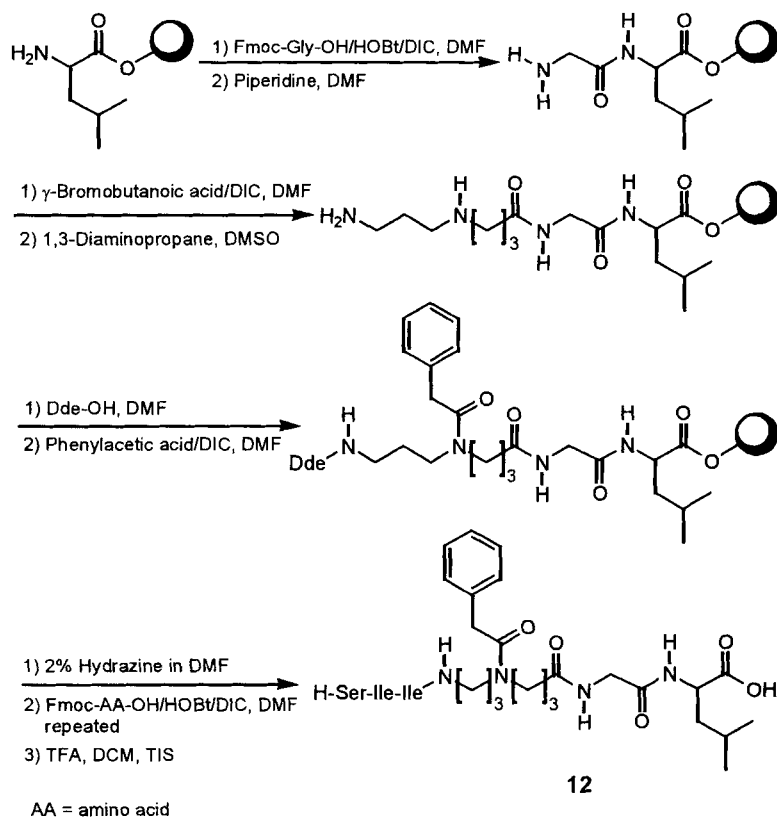
Non-natural peptides involved in the stabilization of the MHC class I molecules can be considered as the starting point for the discovery of new immune modulators such as anti-tumor vaccines and T-cell receptor agonists or antagonists [44–48]. In this context, poly-N-acylated amines (PAAs) as a new class of synthetic oligomers were developed and tested as ligands for the murine class I molecule H-2K<sup>b</sup> [49]. Based on the natural cytotoxic T-cell epitope SIINFEKL, non-peptidic structural elements were introduced in the C-terminal part of the ligand including the anchor positions 5 (Phe) and 8 (Leu) (Fig. 7.7).



**Figure 7.7.** Molecular structures of three potential MHC class I ligands consisting of poly-N-acylated peptidic amines (PAA).

A series of peptide-PAA oligomers was prepared on solid support using two different synthetic strategies. The more flexible procedure, that can be fully automated and that is based on the assembling of simple commercially available building blocks, is illustrated in Scheme 7.6.

The novel PAA oligomers present different backbone modifications (Fig. 7.7), and have been used to determine, first, the optimal spacing and geometry necessary to guarantee a high binding capacity to the MHC. Then, based on the ligand **12** (Scheme 7.6) which displays the highest binding affinity, five randomized libraries were designed. They contain 26 different aromatic, heteroaromatic or pseudoaromatic side chains for the first dominant MHC anchor position 5 and four aliphatic amino acids (Leu, Ile, Val and Met) for the second dominant anchor position 8 (Fig. 7.8). Some interesting ligands have been isolated from the deconvolution of one PAA library. Moreover, the diversity of the aromatic moieties that mimic the side chain of the anchor residue at position 5 permitted the provision of more insights into the structural requirements for fitting into the MHC peptide binding groove. Besides their potent role as lead compounds for investigating immunological defects, PAAs could find applications as new hormone receptor antagonists or enzyme inhibitors.



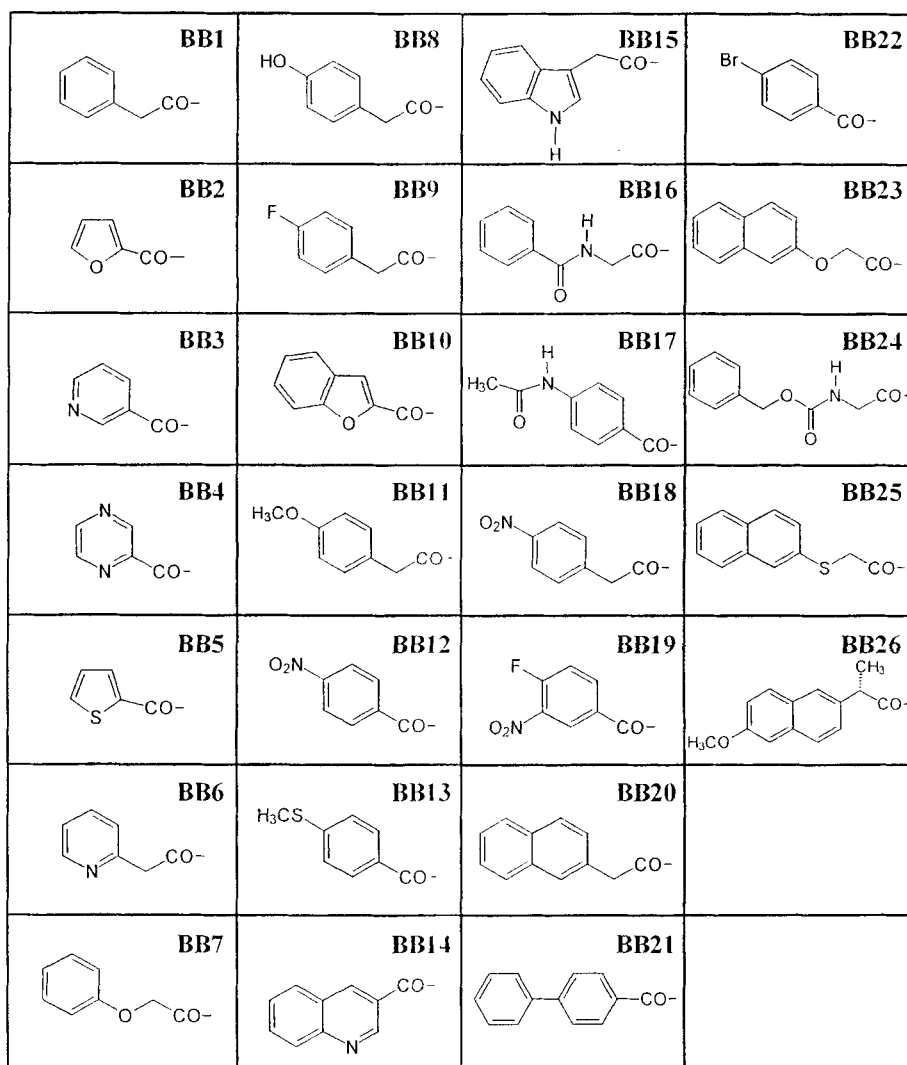
**Scheme 7.6.** Scheme for the automated parallel synthesis of poly-N-acylated amines exemplified for one defined compound, *e.g.*, PAA oligomer **12** exhibiting high binding affinity to MHC class I molecule H-2K<sup>b</sup> [49].

## 7.6 Oligoureas

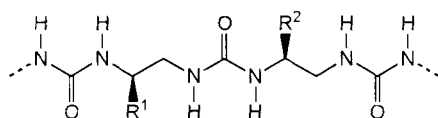
### 7.6.1 Linear Oligoureas

The molecular structure of linear oligoureas is represented in Figure 7.9. These oligomers are easily synthesized on solid phase, starting from activated *p*-nitrophenyl carbamates with amino functions protected as azide, which can be subsequently reduced on the resin using a mixture of SnCl<sub>2</sub>/thiophenol/Et<sub>3</sub>N [50]. Another method on solid-phase synthesis was developed using monophthalimide protected isocyanates as monomers [51, 52].

Oligoureas are characterized by: (i) hydrogen bonding groups inside the backbone; (ii) chiral centers; (iii) significant degree of conformational restriction; and (iv) backbone NH groups available for the introduction of additional side chain functionalities. Such structures should provide new classes of folded polymers with novel conformational features



**Figure 7.8.** Molecular structures of the acylating building blocks used for the synthesis of the PAA libraries (BB = building block) [49].



**Figure 7.9.** Molecular structure of a linear oligourea fragment.

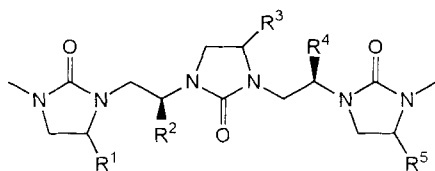
and biological properties. Moreover, the two synthetic methodologies described above are suitable for rapid generation of combinatorial libraries either in solution or on solid phase. A combinatorial library of N,N'-disubstituted linear ureas has been tested for opioid activity at the  $\mu$ ,  $\delta$ ,  $\kappa$ ,  $\sigma$  opioid receptors. After the deconvolution, compounds with an activity for the  $\sigma$  receptor in the nanomolar range have been identified [53].

### 7.6.2 Thioureas

Each urea oxygen atom can potentially form two hydrogen bonds to another properly positioned urea. When the oxygen atom is replaced by a sulfur atom we obtain a thiourea. Due to the increased NH acidity, thioureas form stronger hydrogen bonds than normal ureas and should give rise to complex secondary and tertiary structures (*i.e.*, induction of large two-dimensional pleated sheets) [54]. The solution- and solid-phase synthesis of thioureas is based on the coupling of an isothiocyanate with an amine, the former being easily prepared from a monoprotected diamine with carbon disulfide and carbodiimide [55]. NMR studies of some oligomeric thioureas revealed significant conformational restrictions. These oligomers seem to be very interesting for supramolecular receptor and catalyst design.

### 7.6.3 Oligocycloureas and Cyclothioureas

Cyclic urea moieties are found in many biologically active compounds [56] and expected to have a greater membrane permeability and limited degrees of freedom relative to the corresponding linear oligoureas (Fig. 7.10).



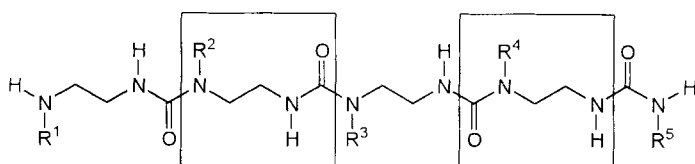
**Figure 7.10.** Molecular structure of a linear oligourea consisting of cyclic urea moieties.

The synthesis in solution of these peptidomimetics starts from readily available N-Boc-protected amino acids. After the coupling with benzylamine and Boc deprotection, the amide bond underwent a selective reduction using  $\text{BH}_3$ . Cyclic urea was then generated by triphosgene-induced cyclization. This monomer was further elongated by acylation of NH ring group with N-Fmoc-amino acid chloride [57]. The solid-phase synthesis of cyclic urea scaffolds has been reported by Nefzi et al. [58]. The strategy is based first on the selective alkylation of the amide bond through which the C-terminal amino acid is linked to the solid support. After the coupling of the second residue and subsequent acylation using acetic anhydride, all the carbonyl groups were reduced with diborane. The two secon-

dary amino groups were finally cyclized to obtain the five-membered ring urea using carbonyldiimidazole or triphosgene. When the cyclization step was performed with thiophosgene, the corresponding cyclic thioureas were obtained. Four combinatorial libraries based on the cyclic urea and thiourea scaffolds have been also synthesized [53].

#### 7.6.4 Ureapeptoids

Combining different classes of peptidomimetics the diversity of the compounds can be increased enormously. One class of these hybrid oligomers is that of ureapeptoids (Fig. 7.11) [59].



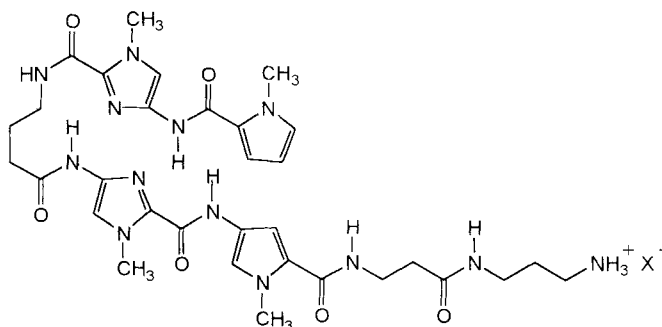
**Figure 7.11.** Structure of an ureapeptoid. The frames highlight two possible synthons for its synthesis.

The building blocks used in the synthesis are N-Boc-protected N-substituted ethylene diamines, which are initially prepared in solution. Until now, syntheses of ureapeptoid dimers and trimers have been carried out in solution, but synthesis on solid phase is presently under implementation.

### 7.7 Heterocyclic Ring-Containing Oligomers

#### 7.7.1 Poly-N-Methyl Pyrroles and Imidazoles

New polyamides based on homo- and/or hetero-pyrrole and imidazole amino acids have been designed for high-affinity recognition of the minor groove of DNA (Fig. 7.12).



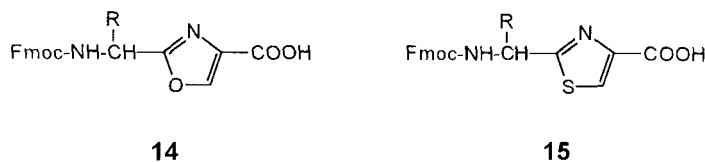
**Figure 7.12.** Molecular structure of a hairpin pyrrole-imidazole tetramer.

A convenient automated solid-phase synthesis, using N-Boc-protected pyrrole and imidazole amino acids, has been developed to increase the number and the complexity of potential pyrrole–imidazole DNA ligands [60]. This effort aims at the discovery of new oligomers which should be able to display a sequence-specific interaction with DNA grooves, and which allow an understanding of the chemical rules for the readout of DNA double-helix. Different hairpin N-methyl-pyrrole-imidazole polyamides, which present also additional aliphatic amino acids within the sequence and at the C-terminal part, have been found to modulate both binding affinity and sequence specificity in the recognition of the DNA minor groove [61–65]. Moreover, these ligands are able to permeate living cells and can inhibit the transcription of specific genes [66].

The rapid solid-phase synthesis of pyrrole–imidazole polyamides with a wide variety of sequences and the generation of combinatorial libraries should increase the possibilities for the recognition of any desired DNA double strand.

### 7.7.2 Thiazole and Oxazole Ring-Containing Peptides

A large variety of thiazole and oxazole heterocycles are produced by many marine organisms and microorganisms. An interesting 43-peptide antibiotic, microcin B17, which inhibits DNA-gyrase, presents inside the sequence eight heterocyclic backbone modifications. For the total synthesis of microcin B17, an efficient procedure for obtaining Boc- and Fmoc-protected amino acids derived from oxazole and thiazole has been developed (Fig. 7.13) [67, 68]. N-Fmoc-2-aminomethylthiazole-4-carboxylic acid was prepared from Boc-glycineamide, converted into the corresponding thioamide using the Lawesson's reagent, and cyclized with 3-bromo-2-oxo-propionic acid. As a final step, the Boc protecting group was removed and substituted with Fmoc. The analogue containing the oxygen atom inside the ring was prepared transforming the starting glycineamide into an imino ether, subsequently cyclized with a serine methyl ester, oxidized, and finally saponified. Again, the Boc group was replaced by Fmoc. Amino- and carboxy-functionalized bis-thiazole, oxazolyl-thiazole and thiazolyl-oxazole building blocks, which can be considered as di- and tripeptide mimetics, have been also synthesized.



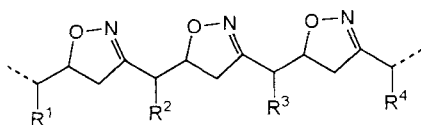
**Figure 7.13.** Molecular structures of N-Fmoc-protected oxazole (**14**) and thiazole (**15**) amino acids used as monomers of oligomer libraries.

The total synthesis of microcin B17 was performed on solid support using Fmoc chemistry. The synthetic microcin revealed the same antibacterial activity as the natural ana-

logue [68]. Besides the preparation of this antibiotic, N-protected thiazole and oxazole amino acids can be considered interesting building blocks for the synthesis of other bioactive compounds, and suitable for solid-phase combinatorial chemistry applications.

### 7.7.3 Polyisoxazolines

The strategy for the generation of polyisoxazolines (Fig. 7.14) is based on subsequent C-C bond-forming steps.



**Figure 7.14.** Molecular structure of a polyisoxazoline fragment.

Each isoxazoline subunit is built up starting from a resin-bound olefin submitted to a 1,3-dipolar cycloaddition of a nitrile oxide, followed by a selenide oxidation/elimination reaction to regenerate the olefin functionality [69]. Using different nitroseleno ethers and nitroalkanes, a library of triisoxazolines has been prepared and characterized by mass spectrometry. This flexible method to create heterocyclic oligomers via C-C bond ring closure is then versatile to combinatorial syntheses.

### 7.7.4 Oligothiophenes

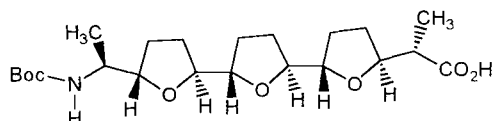
Oligothiophenes are interesting macromolecules due to their optical and electronic properties [70]. A solid-phase synthesis of oligothiophenes from dimer to pentamer has been successfully achieved using an alternating sequence of bromination and Stille reactions [71]. This new procedure avoids several difficulties encountered in the solution synthesis of oligothiophenes, such as loss of functionalization and/or homocoupling, affording products in high yield and purity. Starting from a thiophene functionalized chloromethylated macroporous resin, the  $\alpha$ -position of the heterocyclic ring was brominated with N-bromosuccinimide, followed by coupling to 2-(trimethylstannyl)-4-octylthiophene in the presence of a palladium catalyst [72]. This solid-phase approach is suitable for generation of variable oligothiophenes exploiting different asymmetric thiophene building blocks. Besides the easy access to new materials, the cleavage from the resin yields oligomers with a functionalized handle that may be later anchored to other substrates.

### 7.7.5 Oligotetrahydrofurans

Oligotetrahydrofurans (oligo-THF) have been prepared to form potential polyether helices with ion channel activity [73]. The first generation of these polymers were too



short to span a membrane bilayer. Therefore, to achieve longer oligotetrahydrofurans, Wagner et al. [74] turned to oligo-THF peptides (Fig. 7.15). The preparation of the THF-trimer amino acid building block suitable for oligomerization starts from N-tosyl-alanine and involves several cyclization steps.



**Figure 7.15.** Molecular structure of the N-Boc-protected THF-trimer amino acid building block used in the synthesis of oligo-THF peptides.

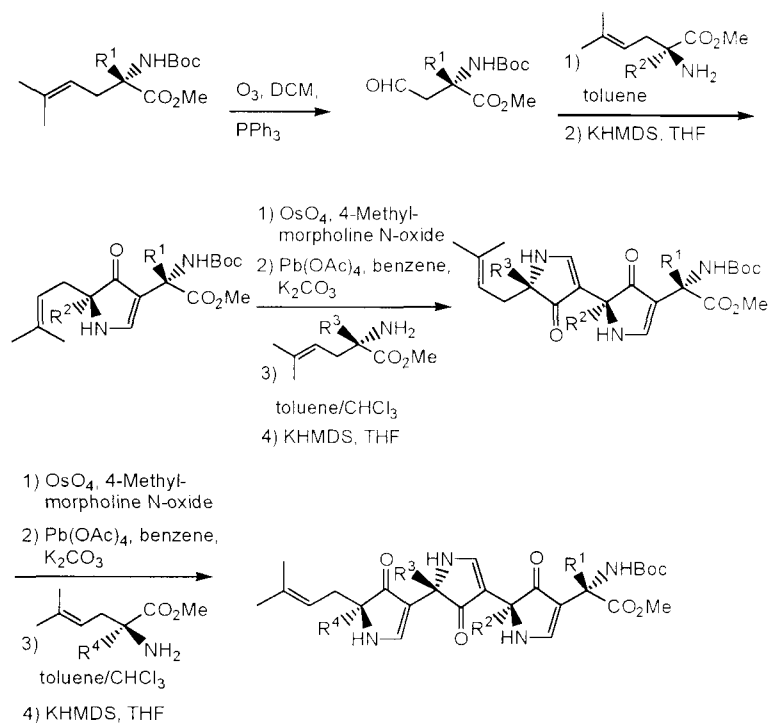
Since the thickness of the lipid bilayer is about 30 Å, the minimum length to form an ion channel is assured by an oligo-THF decapeptide, which can adopt a helical conformation with five turns. The tetrahydrofuran polyamides have been studied in their capacity to interact with membranes [74]. They showed conductance-increasing properties and, therefore, can be considered interesting structures for the design of artificial ion channels. Although oligo-THF peptides have been synthesized in solution, N-protected THF-trimer amino acid, represented in Figure 7.15, is an interesting building block suitable for extension to solid-phase combinatorial syntheses of THF-containing oligomers.

### 7.7.6 Pyrrolinone-Containing Oligomers

Oligomeric peptidomimetics containing two to three 3,5-linked pyrrolinone rings have been designed to mimic  $\beta$ -strand conformations [75]. The pyrrolinone NH protons were shown to maintain the H-bonding donor capabilities of their peptide counterparts, allowing the synthetic oligomer to form parallel and antiparallel  $\beta$ -pleated sheets. Moreover, the side chains appended at position 5 of the heterocyclic ring are oriented in the same direction as in the native peptide. Polypyrrolinones have been synthesized in solution as illustrated in Scheme 7.7 [75]. Each ring closure exploits the condensation reaction between the  $\alpha$ -amino function of an  $\alpha,\alpha$ -tetrasubstituted amino acid ester (see Section 7.8.6) and the side chain aldehyde group of a second N-Boc-protected  $\alpha,\alpha$ -tetrasubstituted amino acid ester, followed by cyclization in the presence of KHMDS [potassium bis(trimethylsilyl)amide].

3,5-Linked pyrrolinone scaffolds, bearing different side chains and C- and N-terminal groups, have been demonstrated to be interesting renin and HIV-1 protease inhibitors with improved transport properties [76]. Recently, a pyrrolinone-peptide hybrid has been proved to be an effective ligand for a class II MHC protein, showing the potential utility of these peptidomimetics also in the treatment of autoimmune diseases [48]. The synthesis of this oligomer has been performed on solid phase. A N-Fmoc-protected bispyrrolinone carboxylic acid was first prepared in solution and then inserted in the central part of the natural epitope to replace four amino acid residues [48]. This example demonstrates

that a pyrrolinone building block, appropriately functionalized, could be suitable for solid-phase combinatorial chemistry applications.

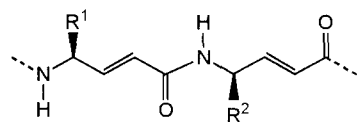


**Scheme 7.7.** Synthesis of a tetrapeptide mimetic pyrrolinone oligomer containing three heterocyclic rings [75].

## 7.8 Other Examples of Synthetic Oligomers

### 7.8.1 Vinylogous Polypeptides

The planar structure of the peptide bond may be extended as represented by the vinylogous polypeptides which contain an (E)-ethenyl unit inserted between the carbonyl carbon and the  $\alpha$ -carbon (Fig. 7.16).



**Figure 7.16.** Molecular structure of a vinylogous peptide fragment.

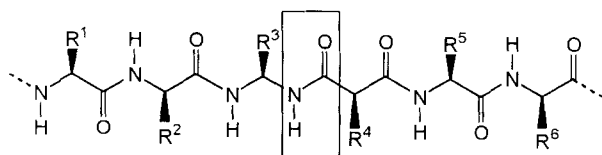
Some vinylogous peptides have been synthesized in solution and shown to possess novel secondary structures. They can adopt a hairpin turn and a novel helical sheet conformation, stabilized by two intramolecular hydrogen bonds of 10- and 12-membered rings, respectively [77, 78].

The monomers are prepared by the homologation of N-protected amino acid aldehydes. Boc-protected amino aldehyde was synthesized in solution via Weinreb amide reduction with  $\text{LiAlH}_4$ . Then, using a Wittig reaction, it was transformed into an olefin bearing an ethyl ester group, which was subsequently saponified.

A new solid-phase approach to olefinic peptidomimetics was recently elaborated by Rotella [79]. A resin-bound amino alcohol was first converted to a N-linked  $\alpha$ -amino aldehyde. Subsequently, using a Wittig reaction, it was transformed into an olefin, functionalized with a carboxylic group, which allows a further elongation with a N-free amino acid ester. This methodology can be readily extended to the preparation of combinatorial libraries suitable for novel enzyme inhibitors discovery. Peptidomimetics containing an (E)-alkene amide bond have been also prepared using solid-phase protocols from alkenylaziridines appropriately protected at the nitrogen [80]. The ring opening, accomplished by alkylcyanocuprates, is the key step for the olefination. Deprotection and elongation allow then combinatorial synthesis of these peptidomimetics.

### 7.8.2 Retro-inverso Pseudopeptides

Another possibility to convert biologically active peptides into more stable molecules is the introduction of the retro-inverso amide bond modification  $\psi[\text{NHCO}]$  (Fig. 7.17) [81, 82].



**Figure 7.17.** Molecular structure of a peptide fragment characterized by a retro-inverso amide bond motif.

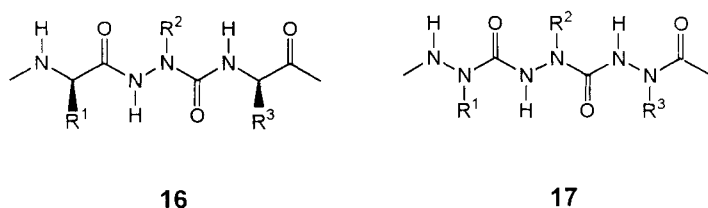
A totally retro-inverso peptide is an isomer of a linear peptide in which the direction of the amino acid sequence is reversed and the chirality at the  $\text{C}^\alpha$  carbon of the residues is inverted. This modification has been widely applied [83]. Antigenicity and immunogenicity have been demonstrated by preparation of retro-inverso isomers of natural antigenic peptides [84]. Retro-inverso forms derived from mellitin and cercropin A maintain antimicrobial activity [85]. A cyclic retro-enantiomeric D amino acid peptide has been designed and shown to be active as synthetic inhibitor of the interaction between immunoglobulin E (IgE) and its high-affinity receptor  $\text{Fc}\epsilon\text{RI}$  [86].

The properties of an active peptide can also be modified simply by inverting the direction of only one amide bond within the backbone. This has been demonstrated by Dürr

et al. on MHC-class I ligands [44]. Then, the retro-inverso motif was introduced in a series of analogues of the antigenic peptide M58-66 derived from the influenza virus matrix protein [45]. Such replacements can be achieved using a 2-substituted malonate derivative and a *gem*-diaminoalkyl residue. Many other biologically relevant partially retro-inverso modified peptides have been prepared and evaluated for their biological activities. They include hormone analogues, enzyme inhibitors and sweeteners. These examples have been recently reviewed by Fletcher and Campbell [82]. The authors gave also detailed information on the synthesis of the oligomers in solution and on solid phase, which is the most promising approach for further combinatorial chemistry.

### 7.8.3 Azatides and Azapeptides

Azatide oligomers are biopolymer mimetics consisting of  $\alpha$ -aza-amino acids linked in a repetitive manner. The molecular structures of azatide and azapeptide fragments are compared in Figure 7.18.



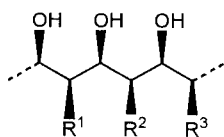
**Figure 7.18.** Molecular structures of azapeptide (**16**) and azatide (**17**) fragments possessing urea- and hydrazide-derived structural elements (compare Fig. 7.11).

Azapeptides are peptides in which one (or more) of the  $\alpha$ -carbons have been replaced by a trivalent nitrogen atom [87, 88]. This transformation results in a loss of asymmetry associated with the  $\alpha$ -carbon and affords a structure that can be considered intermediate in configuration between D- and L-amino acids [89]. Interest in this  $\alpha$ -carbon replacement units derives from its ability to provide resistance to enzymatic cleavage and its capacity to act as a selective inhibitor of cysteine and serine proteases. Complete substitution of all asymmetric trisubstituted  $\alpha$ -carbons is represented by 'pure azapeptides', termed azatides [90].

For their preparation it is necessary, as a first step, to synthesize Boc-protected  $\alpha$ -aza-amino acid monomers. Liquid-phase procedures have been optimized for the synthesis of a MeO-PEG-supported Leu-enkephalin azatide ( $Y^aG^aG^aF^aL^a$ ) analogue [90]. Unfortunately, this peptidomimetic does not compete with the natural peptide in an ELISA for a specific monoclonal antibody. This was not completely unexpected by the authors because azatide adopts an extended conformation within the critical glycine region, while the bioactive peptide conformation is thought to be a sort of  $\beta$ -turn. Starting from N-protected  $\alpha$ -aza amino acid residues, the generation of combinatorial azatide libraries seems to be feasible.

### 7.8.4 Polyketides

Polyketides of bacterial origin include a wide range of bioactive molecules with anti-biotic, antitumor, antifungal and immunosuppressant action (see Chapter 12). Polyketide libraries are considered as a new source of molecular diversity accessible by chemical and biological synthetic methodologies. The chemical synthesis of these polyoxygenated molecules with a hydrocarbon backbone requires sophisticated stereochemical controls of the several chiral centers, protection and reestablishment of the functional groups (Fig. 7.19). Recently, some efforts have been made towards a new approach to prepare polyketide using solid-phase procedures [91–93].

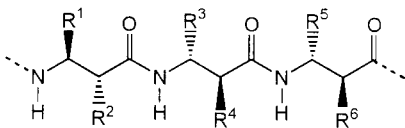


**Figure 7.19.** Molecular structure of an unnatural polyketide fragment.

The assembly of polyketide libraries is based on iterative and asymmetric aldol reactions on the resin using the Evans and Weinreb procedures. Each cycle allows the introduction of a different side chain and the creation of a new stereocenter, which contribute to enhance the potential for diversification [91]. Using a different strategy, based on the iterative application of the boron-mediated aldol reaction of an ethyl ketone, followed by a reduction, a polyketide library was prepared in solution [92]. Interestingly, despite the presence of a flexible backbone, polyketide oligomers seem to adopt predominantly a single extended conformation with a helical twist, reinforced by an intramolecular hydrogen bonding pattern [93].

### 7.8.5 $\beta$ -Polypeptides

$\beta$ -Polypeptides are a new type of oligomers that are currently under extensive scrutiny by the groups of Gellman [3] and Seebach [94].  $\beta$ -Peptides are comprised of  $\beta$ -amino acid monomers (Fig. 7.20).



**Figure 7.20.** Molecular structure of a  $\beta$ -peptide fragment, containing chiral monomers.

The presence of an additional carbon atom between the amino and the carboxy function increases the number of possible constitutional and configurational isomers. Several methods have been developed to prepare N-protected  $\beta$ -amino acid building blocks for the synthesis of  $\beta$ -peptide oligomers [95, 96]. The Arndt–Eistert homologation and the Evans' methodology are two versatile strategies to obtain enantiomerically pure N-protected monomers with the side chains at the C(3) and C(2) carbon atoms, respectively [97, 98].

One of the main interest in the synthesis of systematically varied sets of  $\beta$ -peptides is focused on the discovery of novel well-defined three-dimensional structures [99]. Surprisingly, it has been found that short  $\beta$ -peptides are able to adopt very stable helical structures, characterized by different intramolecular H-bonding patterns [3, 100]. Moreover, the folding propensities of  $\beta$ -peptide oligomers can be modulated by changing the nature of the residues [99, 101]. It has been demonstrated that  $\beta$ -peptides are also resistant to proteases [102], and this may open the door to the possibility of pharmacological applications of  $\beta$ -peptides as potential therapeutic agents.

Recently, Fmoc-N-protected  $\beta$ -amino acid synthons have been prepared and used for the synthesis of  $\beta$ -peptides on solid phase [103]. This methodology facilitates enormously the search for new bioactive compounds, above all through the generation of combinatorial  $\beta$ -peptide libraries. In this context,  $\beta$ -amino acids have been used as building blocks for RGD cyclic peptides and for the synthesis of an inhibitor of human cathepsin L, previously identified by screening and deconvolution of pentapeptide amide collections [104, 105].

With the introduction of a further carbon atom between the amino and carboxy groups, we pass from  $\beta$ - to  $\gamma$ -amino acids. Research on  $\gamma$ -amino acids and on the preparation of  $\gamma$ -peptides is still in its infancy, but will undoubtedly increase during the next few years, as these unnatural peptides have already been shown to have interesting structural features, for example the tendency to adopt new stable helical structures [106, 107].

### 7.8.6 $\alpha,\alpha$ -Tetrasubstituted Amino Acid-Containing Peptides

Non-protein, conformationally restricted  $C^{\alpha,\alpha}$ -tetrasubstituted amino acids, also termed  $\alpha,\alpha$ -dialkylated glycines, became the object of detailed studies as they are able to modify the conformational behavior, the enzymatic stability and the biological properties of naturally occurring peptides. The conformational preferences of these amino acids have been extensively analyzed by the group of Toniolo [108, 109].  $C^{\alpha,\alpha}$ -Tetrasubstituted amino acids show a much more limited conformational freedom compared to protein amino acids, and may be used to freeze a peptide in a single predetermined folded structure. Rigid peptide scaffolds can be exploited for different applications, from complexation and selective transport to catalysis and synthesis of artificial receptors, enzymes and molecular spacers [110]. The synthesis of peptides containing  $\alpha,\alpha$ -dialkylated amino acids has been performed mainly in solution because of difficulties during the coupling steps arising from the extreme steric hindrance of the residues. In general, solution synthesis of such peptides requires long reaction times. Recently, a new method of activation of the carboxylic function of the  $\alpha,\alpha$ -disubstituted glycines has permitted easier and more rapid

solid-phase procedures for the preparation of peptides containing sterically hindered residues [111, 112].

Another means of overcoming the sterically very demanding coupling reactions is to reduce the bulkiness of the residues, inherent to the presence of both the C<sup>α,α</sup>-disubstitution, and the urethane protecting group. This has been demonstrated by the use of α-azido acids in which the azide is the precursor of the amino function [113]. These monomers can be activated as acid chlorides, and their preparation is also compatible with the presence of the side chain protecting group used in Fmoc-based peptide synthesis. This approach has been exploited in the solid-phase synthesis of α-aminoisobutyric acid (Aib)-rich peptides [114].

The synthetic strategies for incorporation of non-protein amino acids within a peptide sequence described above are based on the employment of individually prepared building blocks [115–118]. A more direct and flexible methodology would be to build the α,α-tetrasubstituted residues directly during the peptide elongation using simple solid-phase alkylation chemistry [119]. Peptides containing racemic α,α-dialkylated glycines have been successfully achieved by Scott and co-workers [120]. A resin-bound protein amino acid was activated as Schiff base with an aldehyde and subsequently alkylated at the α-position with an alkyl halide in the presence of a strong base. The free amino function was then regenerated by hydroxylamine transamination.

This mild procedure for the solid-phase synthesis of peptides containing sterically hindered residues can be easily extended to generate combinatorial libraries of unnatural peptides, which may display biological properties superior to those of natural counterparts.

### 7.8.7 Peptide Nucleic Acids

In recent years, research in the field of oligonucleotide mimetics has become intensive, largely due to the encouraging possibility of discovering new therapeutic agents, including antisense or antigene drugs.

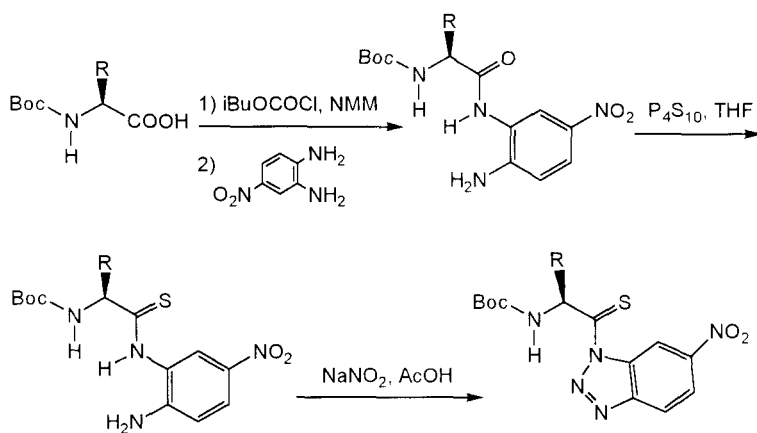
Peptide nucleic acids (PNAs) are one of the most elegant and original class of analogues which are able to interact with DNA and RNA helical strands [121]. In PNA, the phosphate and sugar backbone moieties which are present in the natural oligonucleotide have been replaced by a pseudopeptide scaffold composed of N-(2-aminoethyl)glycine monomers. Surprisingly, this drastic structural modification of the oligonucleotide backbone afforded oligomeric nucleic acids which are able to bind to complementary DNA and RNA with an affinity even higher than the natural counterparts [122]. Synthesis, chemical, biological and conformational properties of PNAs and PNA/DNA chimeras have been elegantly reviewed by Uhlmann et al. [123]. Information on the pharmacology and toxicity of PNA analogues has not yet been published, but these authors believe that the therapeutic use of these compounds is feasible in the near future. Besides their potential therapeutic applications, PNAs are also interesting tools for novel studies on the conformational and biological properties of DNA and RNA themselves. Indeed, many results on the structure of PNA/PNA, PNA/DNA and PNA/RNA chimeras have been already published [124].

## 7.9 Single Position Peptide Bond Modifications

### 7.9.1 Thioamide Pseudopeptides

One of the most subtle peptide bond modifications is the insertion of a thioamide  $\psi$ [CSNH] group, in which the carbonyl oxygen has been replaced by a sulfur atom [125, 126]. Thioamide formations became more easily accessible with the introduction of the thionating Lawesson's reagent.

Recently, a mild, efficient and racemization-free route for the incorporation of a thioamide linkage into a peptide either in solution or on solid phase was provided by Shalaby et al. [127]. First, the N-Boc-amino acid anilide derivative was thionylated with  $P_4S_{10}$  and subsequently activated by transformation to the benzotriazole derivative, affording a stable thioacylating reagent (Scheme 7.8) suitable for combinatorial syntheses.



**Scheme 7.8.** Synthetic route to the stable N-Boc-protected thioamino acid benzotriazole active ester [127].

The introduction of the thioamides in the backbone produces only minor physical changes, but has profound effects on the biological activities and selectivity. The peptide conformation is strongly influenced: (i) sulfur is a weaker hydrogen bond acceptor than oxygen; and (ii) the amide nitrogen adjacent to the thiocarbonyl group is a stronger hydrogen donor. Therefore, thionation scans may be helpful to clarify conformationally induced structure–activity relationships of bioactive molecules [128]. A series of tetrapeptides containing a thiooxo-peptidyl-prolyl bond have been used, for example, to study the mode of action of a cyclophilin enzyme that belongs to peptidyl-prolyl *cis-trans* isomerases [129]. A peptide containing the thioamide moiety was found to be a competitive inhibitor for the enzyme. The *cis-trans* ratio changed when thioamide analogues were compared with the normal amide bond-containing substrates.



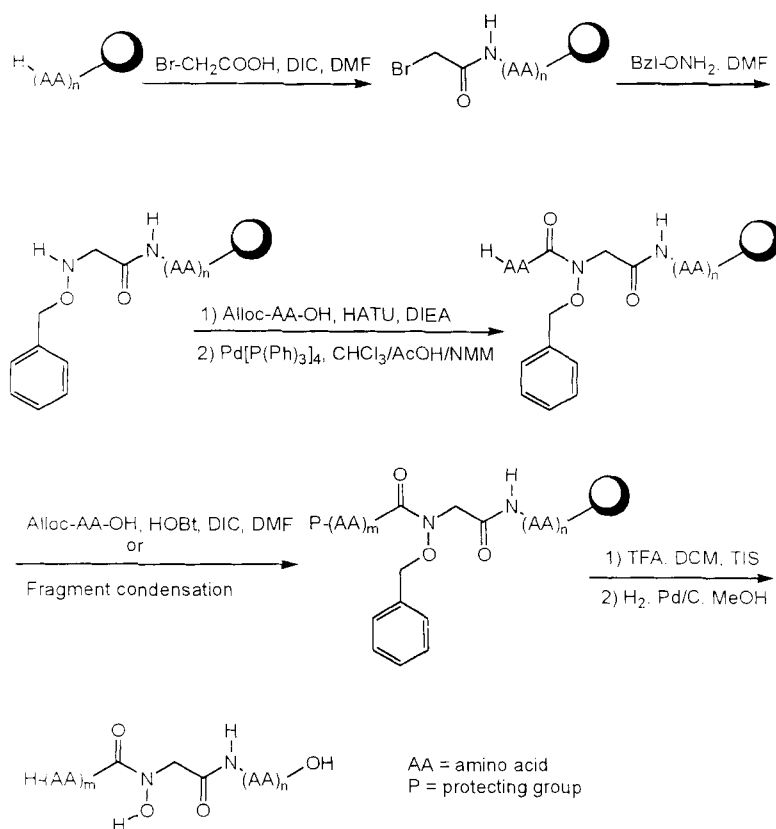
### 7.9.2 Reduced-Amide Bond Peptides

One of the first motifs that have been proposed to modify the natural peptide backbone was the reduction of the amide carbonyl to a methylene group. The introduction of the CH<sub>2</sub>NH peptide bond isostere often produces interesting biological properties due to effects on the backbone conformation and/or increased resistance to peptidases. A common means of preparing the  $\psi$ [CH<sub>2</sub>NH] pseudopeptide on solid phase is the use of the reductive alkylation of an amino function with an aldehyde, employing NaBH<sub>3</sub>CN in acidic dimethylformamide solution [130, 131].

Recently, a series of reduced peptide bond analogues of the C-terminal hexapeptide of histone H3 (Ile-Arg-Gly-Glu-Arg-Ala) have been synthesized in solution and examined for their ability to bind polyclonal and monoclonal antibodies generated against the parent natural peptide and the protein [132, 133]. The results show that the  $\psi$ [CH<sub>2</sub>NH] analogues can mimic the parent peptide. This is the first example in which reduced peptide bond pseudopeptides has been used in immunochemistry. A second example is the synthesis of a series of analogues in which each peptide bond of the antigenic peptide Ser-Tyr-Ile-Pro-Ser-Ala-Glu-Lys-Ile has been systematically replaced by  $\psi$ [CH<sub>2</sub>NH]. These oligomers were tested for their ability to bind to a murine MHC class I molecule [134], and five of them displayed high-affinity binding. This is of great interest for the design of peptidomimetics with potential immunotherapeutic properties (*i.e.*, development of peptidomimetic-based cytotoxic T-lymphocytes vaccines). This peptide bond isostere was also applied for anticancer drug development. Reduced pseudopeptides inhibit the RAS farnesyl-protein transferase [135], while reduced peptide bond analogues of dinorphin A also exhibited a surprisingly strong binding affinity for the opioid  $\mu$ ,  $\delta$  and  $\kappa$  receptors [136]. Linear and cyclic libraries containing one, two or three  $\psi$ [CH<sub>2</sub>NH] amide modifications have been synthesized on solid phase and characterized by Wen and Spatola [137]. These mixtures are based on biologically relevant structures including enkephalin and RGD peptide analogues.

### 7.9.3 N-Hydroxyamide Bond-Containing Peptides

Several natural products, for example siderophores, contain the N-hydroxy amide  $\psi$ [CON(OH)] motif [138]. Within a peptide backbone, this group increases the stability to enzyme degradation and induces characteristic conformational behavior [139]. In addition to the synthesis in solution of N-hydroxy amide-containing peptides (which is not trivial), a new solid-phase approach has recently been developed [140]. To explore the features of the N-hydroxy amide moiety using automated and combinatorial techniques, a method for the preparation of  $\psi$ [CON(OH)] peptide ligands for MHC-I molecules has been elaborated [140]. The strategy for the parallel preparation of these peptidomimetics on a solid support is illustrated in Scheme 7.9. The key step is the nucleophilic substitution reaction of resin-bound bromocarboxylic acids by O-benzylhydroxylamine, which requires several days.

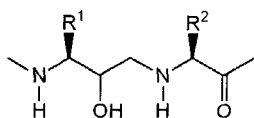


**Scheme 7.9.** Solid-phase synthesis of N-hydroxy amide  $\psi[\text{CON(OH)}]$  peptide bond analogues [140].

A series of  $\psi[\text{CON(OH)}]$  peptides has been synthesized on the basis of the MHC class I H-2K<sup>b</sup> restricted SIINFEKL epitope, by systematically replacing each amino acid position by a N-hydroxy glycine residue. It has been shown that some of the analogues display an MHC binding affinity in the nanomolar range, even higher than the natural peptide. More interestingly, N-hydroxylation of the peptide backbone can convert a ligand that has agonist activity on T-cell receptors into an antagonist. This is the first example in which a T-cell antagonist is generated by a main chain modification at a site of the peptide that is not exposed to the T-cell receptor [141].

### 7.9.4 Hydroxyethylamine Peptide Bond Isosteres

Hydroxyethylamine modification  $\psi[\text{CH(OH)CH}_2\text{NH}]$  has been found to generate potent protease inhibitors (Fig. 7.21) [142].

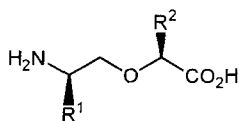


**Figure 7.21.** Molecular structures of a hydroxyethylamine peptide isostere.

This modification seems to mimic the presumed tetrahedral intermediate involved in the aspartyl protease-catalyzed hydrolysis of the peptide bond. A HIV-1 protease inhibitor containing a hydroxyethylene bond isostere has been prepared on solid phase [143]. A Boc-N-protected bromomethyl ketone was coupled to the resin-bound N-free Pro-Ile-Val tripeptide, and the keto function was reduced to an OH group with  $\text{NaBH}_4$ . The peptide analogue was then elongated on the resin. This procedure allows rapid access to a variety of hydroxyethylamino peptide bond isosteres, and in good yield.

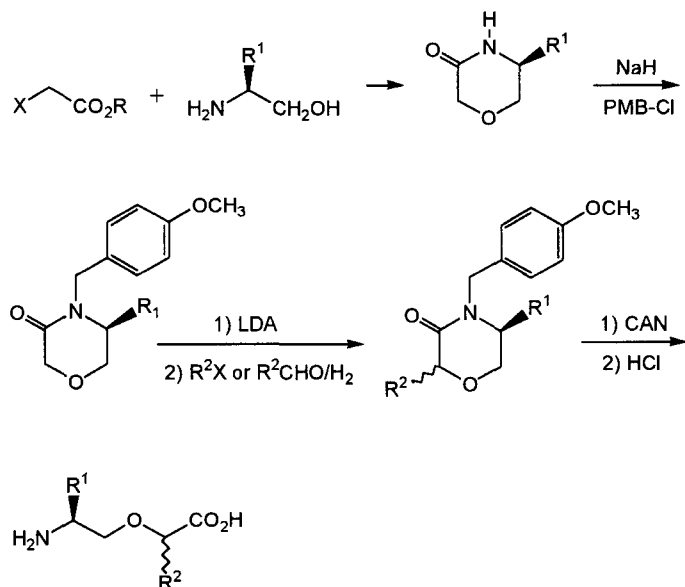
### 7.9.5 Methylene Ether Isosteres

The introduction of the methylene ether isostere  $\psi[\text{CH}_2\text{O}]$  has not been fully exploited (Fig. 7.22) [144, 145]. Several methods have been suggested for the preparation of these peptidomimetics. The methylene ether bond can be obtained by substitution reac-



**Figure 7.22.** Molecular structure of a methylene ether-containing pseudo-peptide.

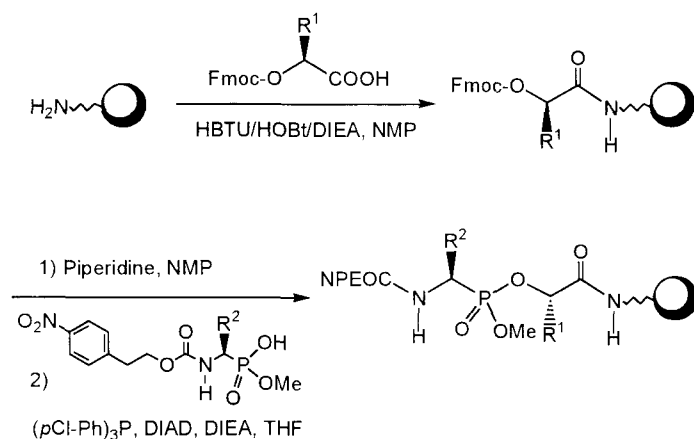
tion of N-protected amino alcohols with  $\alpha$ -bromo carboxylic acid esters or by reaction of  $\alpha$ -hydroxy acid esters with N-protected amino halides [145]. Another procedure for synthesizing a methylene ether-containing dipeptide mimetic is based on the ring opening of a morpholinone derivative in acidic conditions [145]. The method illustrated in Scheme 7.10 presents the advantage that side chain functionalities can be introduced at a later stage in the synthesis. This is also amenable for combinatorial chemistry to generate  $\psi[\text{CH}_2\text{O}]$  dipeptide libraries [145].



**Scheme 7.10.** Synthetic routes to the introduction of methylene ether isostere  $\psi[\text{CH}_2\text{O}]$  [145].

### 7.9.6 Phosphono- and Phosphinopeptides

Peptidylphosphonates have been demonstrated to mimic effectively the transition state involved in the enzymatic hydrolysis of the peptide bond, and are commonly used as inhibitors of peptidases, and in particular of metallo-proteases [146]. A solid-phase method for the introduction of the  $\psi[\text{PO}(\text{OMe})\text{O}]$  motif exploits a modified Mitsunobu reaction (Scheme 7.11) [147].



**Scheme 7.11.** Solid-phase synthesis for the introduction of a peptidylphosphonate bond [147].

A resin-linked  $\alpha$ -hydroxy acid is condensed with a methyl  $\alpha$ -[N-(4-nitrophenethoxy-carbonyl)amino] alkylphosphonic acid, affording a phosphonate-containing peptide analogue with a coupling yield higher than 90%. One advantage of this method is that it enables the generation of combinatorial peptidylphosphonate libraries. Several trimer peptidylphosphonate inhibitors of thermolysin have been prepared [148, 149]. After screening and deconvolution of a phosphonic library, a series of active compounds were identified.

A second possibility of amide bond replacement by a phosphorus atom-containing moiety is represented by phosphinic-based peptides. Pseudopeptides with the  $\psi$ [PO(OH)CH<sub>2</sub>] modification display a strong inhibition activity of zinc proteases such as bacterial collagenase [150]. Solid-phase synthesis of phosphinic peptides became more easily feasible with the discovery of the adamantyl (Ad) orthogonal protecting group for the phosphinyl hydroxy function. Different Fmoc-Xaa $\psi$ [PO(OAd)CH<sub>2</sub>]Xaa-OH building blocks have been prepared [151] and used for the synthesis of phosphinic peptid libraries. A screening of a targeted library enabled the characterization of a potent and selective inhibitor of a mammalian zinc endopeptidase [152].

## 7.10 Summary

In the past few years, intensive efforts have been made to create novel and original synthetic oligomers. Many different modifications have been introduced to replace the labile peptide bond. Among the various positive developments in this field, the discovery of molecules with improved pharmacological properties compared to those of natural peptides can be considered one of the most important. The study of the new conformational features of these synthetic frameworks can supply additional information on the structure–activity relationships.

The identification of interesting synthetic oligomers as new lead structures for therapeutic applications is strictly related to the generation of a large number of compounds which should be evaluated for their biological activities. The remarkable progress made in this direction is clearly due to the development of combinatorial chemistry techniques. Automated solid-phase methods have been extended to the rapid synthesis of peptidomimetic libraries in order to create diverse mixtures which should contain few highly active compounds. The screening and deconvolution of such combinatorial pools then afford novel biologically relevant molecules which can be further developed as therapeutic agents.

In conclusion, research in the field of combinatorial synthetic oligomers for drug discovery is showing both continuous and beneficial progress, and new creative scaffolds will undoubtedly be proposed in the near future.

## 7.11 Acknowledgements

The author wishes to thank Prof. Günther Jung for his enthusiastic support during the writing of this chapter, and for his critical review of the text. Dr. Gilles Guichard and Dr. Jean-Paul Briand are also acknowledged for their helpful suggestions and comments.

## 7.12 References

- [1] A. Giannis, T. Kolter, *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 1244–1267.
- [2] R. M. J. Liskamp, *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 633–636.
- [3] S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173–180.
- [4] J. J. Plattner, D. W. Norbuk, in *Drug Discovery Technologies* (Eds.: C. R. Clark, W. H. Moos), Ellis Horwood Ltd, Chichester, **1990**, pp. 92–126.
- [5] G. Jung, *Combinatorial Peptide and Nonpeptide Libraries*, Wiley-VHC, Weinheim, **1996**.
- [6] R. B. Merrifield, *Science* **1986**, *232*, 341–347.
- [7] R. N. Zuckermann, *Curr. Opin. Struct. Biol.* **1993**, *3*, 580–584.
- [8] R. N. Zuckermann, J. M. Kerr, S. B. H. Kent, W. H. Moos, *J. Am. Chem. Soc.* **1992**, *114*, 10646–10647.
- [9] R. N. Zuckermann, E. J. Martin, D. C. Spellmeyer, G. B. Stauber, K. R. Shoemaker, J. M. Kerr, G. M. Figliozzi, D. A. Goff, M. A. Siani, R. J. Simon, S. C. Banville, E. G. Brown, L. Wang, L. S. Richter, W. H. Moos, *J. Med. Chem.* **1994**, *37*, 2678–2685.
- [10] J. E. Murphy, T. Uno, J. D. Hamer, F. E. Cohen, V. Dwarki, R. N. Zuckermann, *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 1517–1522.
- [11] Y. Pei, W. H. Moos, *Tetrahedron Lett.* **1994**, *35*, 5825–5828.
- [12] P. Armand, K. Kirshenbaum, R. A. Goldsmith, S. Farr-Jones, A. E. Barron, K. T. V. Truong, K. A. Dill, D. F. Mierke, F. E. Cohen, R. N. Zuckermann, E. K. Bradley, *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 4309–4314.
- [13] R. J. Simon, R. S. Kania, R. N. Zuckermann, V. D. Huebner, D. A. Jewell, S. C. Banville, S. Ng, L. Wang, S. Rosenberg, C. K. Marlowe, D. C. Spellmeyer, R. Tan, A. D. Frankel, D. V. Santi, F. E. Cohen, P. A. Bartlett, *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 9367–9371.
- [14] J. A. Krutzler, L. J. F. Hofmeyer, W. Heerma, C. Versluis, R. M. J. Liskamp, *Chem. Eur. J.* **1998**, *4*, 1570–1580.
- [15] R. A. Zambias, D. A. Boulton, P. R. Griffin, *Tetrahedron Lett.* **1994**, *35*, 4283–4286.
- [16] A. Boeijen, R. M. J. Liskamp, *Tetrahedron Lett.* **1998**, *39*, 3589–3592.
- [17] Y. Feng, G. Melacini, M. Goodman, *Biochemistry* **1997**, *36*, 8716–8724.
- [18] T.-A. Tran, R.-H. Mattern, M. Afergan, O. Amitay, O. Ziv, B. A. Morgan, J. E. Taylor, D. Hoyer, M. Goodman, *J. Med. Chem.* **1998**, *41*, 2679–2685.
- [19] S. Østergaard, A. Holm, *Mol. Div.* **1997**, *3*, 17–27.
- [20] V. Nikotaiev, A. Stierendová, V. Krchnák, B. Seligmann, K. S. Lam, S. E. Salmon, M. Lebl, *Pept. Res.* **1993**, *6*, 161–170.
- [21] J. M. Ostresh, G. M. Husar, S. E. Blondelle, B. Dörner, P. A. Weber, R. A. Houghten, *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 11138–11142.
- [22] B. Dörner, G. M. Husar, J. M. Ostresh, R. A. Houghten, *Bioorg. Med. Chem.* **1996**, *4*, 709–715.
- [23] S. C. Miller, T. S. Scalan, *J. Am. Chem. Soc.* **1997**, *119*, 2301–2302.
- [24] J. F. Reichwein, R. M. J. Liskamp, *Tetrahedron Lett.* **1998**, *39*, 1243–1246.
- [25] B. C. Hamper, S. A. Kolodziej, A. M. Scates, R. G. Smith, E. Cortez, *J. Org. Chem.* **1998**, *63*, 708–718.
- [26] C. Y. Cho, E. J. Moran, S. R. Cherry, J. C. Stephans, S. P. A. Fodor, C. L. Adams, A. Sundaram, J. W. Jacobs, P. G. Schultz, *Science* **1993**, *261*, 1303–1305.
- [27] E. J. Moran, T. E. Wilson, C. Y. Cho, S. R. Cherry, P. G. Schultz, *Biopolymers* **1995**, *37*, 213–219.
- [28] S. J. Paikoff, T. E. Wilson, C. Y. Cho, P. G. Schultz, *Tetrahedron Lett.* **1996**, *37*, 5653–5656.
- [29] D. Marcovici-Mizrahi, H. E. Gottlieb, V. Marks, A. Nudelman, *J. Org. Chem.* **1996**, *61*, 8402–8406.
- [30] X. Wang, I. Huq, T. M. Rana, *J. Am. Chem. Soc.* **1997**, *119*, 6444–6445.
- [31] R. Warrass, P. Walden, K.-H. Wiesmüller, G. Jung, *Let. Pept. Sci.* **1998**, *5*, 125–128.
- [32] R. Warrass, K.-H. Wiesmüller, G. Jung, *Tetrahedron Lett.* **1998**, *39*, 2715–2716.
- [33] W. J. Moore, A. Schouten, J. Kroon, R. J. M. Liskamp, *Int. J. Pept. Protein Res.* **1995**, *45*, 501–507.
- [34] C. Gennari, B. Salom, D. Potenza, A. Williams, *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2067–2069.
- [35] F. J. M. Dujols, M. E. Mulliez, *J. Org. Chem.* **1996**, *61*, 5648–5649.
- [36] C. Gennari, B. Salom, D. Potenza, C. Longari, E. Fioravanzo, O. Carugo, N. Sardone, *Chem. Eur. J.* **1996**, *2*, 644–655.
- [37] C. Gennari, H. P. Nestler, B. Salom, W. C. Still, *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1763–1765.
- [38] C. Gennari, H. P. Nestler, B. Salom, W. C. Still, *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1765–1768.

- [39] C. Gennari, C. Longari, S. Ressel, B. Salmon, U. Piarulli, S. Ceccarelli, A. Mielgo, *Eur. J. Org. Chem.* **1998**, 2437–2449.
- [40] M. Gude, U. Piarulli, D. Potenza, B. Salom, C. Gennari, *Tetrahedron Lett.* **1996**, 37, 8589–8592.
- [41] C. Gennari, M. Gude, D. Potenza, U. Piarulli, *Chem. Eur. J.* **1998**, 4, 1924–1931.
- [42] D. W. P. M. Löwik, S. J. E. Mulders, Y. Cheng, Y. Shao, R. M. J. Liskamp, *Tetrahedron Lett.* **1996**, 37, 8253–8256.
- [43] P. W. P. M. Löwik, M. D. Weingarten, M. Brakema, A. J. Broumer, W. C. Still, R. M. J. Liskamp, *Angew. Chem. Int. Ed. Engl.* **1998**, 37, 1846–1850.
- [44] H. Dürr, M. Goodman, G. Jung, *Angew. Chem. Int. Ed. Engl.* **1992**, 31, 785–787.
- [45] G. Guichard, F. Connan, R. Graff, M. Ostankovitch, S. Muller, J.-G. Guillet, J. Choppin, J.-P. Briand, *J. Med. Chem.* **1996**, 39, 2030–2039.
- [46] D. Rognan, L. Scapozza, G. Folkers, A. Daser, *Proc. Natl. Acad. Sci. U. S. A.* **1995**, 92, 753–757.
- [47] G. A. Weiss, E. J. Collins, D. N. Garboczi, D. C. Wiley, S. L. Schreiber, *Chem. Biol.* **1995**, 2, 401–407.
- [48] A. B. Smith III, A. B. Benowitz, M. C. Guzman, P. A. Sprengeler, R. Hirschmann, E. J. Schweiger, D. R. Bolin, Z. Nagy, R. M. Campbell, D. C. Cox, G. L. Olson, *J. Am. Chem. Soc.* **1998**, 120, 12704–12705.
- [49] A. Bianco, C. Brock, C. Zabel, T. Walk, P. Walden, G. Jung, *J. Biol. Chem.* **1998**, 273, 28759–28765.
- [50] J.-M. Kim, Y. Bi, S. J. Paikoff, P. G. Schultz, *Tetrahedron Lett.* **1996**, 37, 5305–5308.
- [51] K. Burgess, D. S. Linthicum, H. Shin, *Angew. Chem. Int. Ed. Engl.* **1995**, 34, 907–909.
- [52] K. Burgess, J. Ibarzo, D. S. Linthicum, D. H. Russel, H. Shin, A. Shitangkoon, R. Totani, A. J. Zhang, *J. Am. Chem. Soc.* **1997**, 119, 1556–1564.
- [53] A. Nefzi, C. Dooley, J. M. Ostresh, R. A. Houghten, *Bioorg. Med. Chem. Lett.* **1998**, 8, 2273–2278.
- [54] J. S. Nowick, D. L. Holmes, G. Mackin, G. Noronha, A. J. Shaka, E. M. Smith, *J. Am. Chem. Soc.* **1996**, 118, 2764–2765.
- [55] J. Smith, J. L. Liras, S. E. Schneider, E. V. Anslyn, *J. Org. Chem.* **1996**, 61, 8811–8818.
- [56] P. Y. S. Lam, P. K. Jadhav, C. J. Eyermann, C. N. Hodge, Y. Ru, L. T. Bacheler, J. L. Meek, M. J. Otto, M. M. Reyner, Y. N. Wong, C.-H. Chang, P. C. Weber, D. A. Jackson, T. R. Sharpe, S. Erickson-Viitanen, *Science* **1994**, 263, 380–384.
- [57] J.-M. Kim, T. E. Wilson, T. C. Norman, P. G. Schultz, *Tetrahedron Lett.* **1996**, 37, 5309–5312.
- [58] A. Nefzi, J. M. Ostresh, J.-P. Meyer, R. A. Houghten, *Tetrahedron Lett.* **1997**, 38, 931–934.
- [59] J. A. W. Kruijtzter, D. J. Lefeber, R. M. J. Liskamp, *Tetrahedron Lett.* **1997**, 38, 5335–5338.
- [60] E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1996**, 118, 6141–6146.
- [61] M. E. Parks, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1996**, 118, 6147–6152.
- [62] S. E. Swalley, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1997**, 119, 6953–6961.
- [63] W. A. Greenberg, E. E. Baird, P. B. Dervan, *Chem. Eur. J.* **1998**, 4, 796–805.
- [64] S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, *Nature* **1998**, 391, 468–471.
- [65] J. M. Turner, S. E. Swalley, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1998**, 120, 6219–6226.
- [66] J. M. Gottesfeld, L. Neely, J. W. Trauger, E. E. Baird, P. B. Dervan, *Nature* **1997**, 387, 202–205.
- [67] G. Videnov, D. Kaiser, C. Kempter, G. Jung, *Angew. Chem. Int. Ed. Engl.* **1996**, 35, 1503–1506.
- [68] G. Videnov, D. Kaiser, M. Brooks, G. Jung, *Angew. Chem. Int. Ed. Engl.* **1996**, 35, 1506–1508.
- [69] M. J. Kurth, L. A. Ahlberg Randall, K. Takenouchi, *J. Org. Chem.* **1996**, 61, 8755–8761.
- [70] J. Roncali, *Chem. Rev.* **1997**, 97, 173–205.
- [71] P. R. L. Malenfant, J. M. J. Fréchet, *J. Chem. Soc., Chem. Commun.* **1998**, 2657–2658.
- [72] P. R. L. Malenfant, L. Groenendaal, J. M. J. Fréchet, *J. Am. Chem. Soc.* **1998**, 120, 10990–10991.
- [73] U. Koert, M. Stein, K. Harms, *Angew. Chem. Int. Ed. Engl.* **1994**, 33, 1180–1182.
- [74] H. Wagner, K. Harms, U. Koert, S. Meder, G. Boheim, *Angew. Chem. Int. Ed. Engl.* **1996**, 35, 2643–2646.
- [75] A. B. Smith III, T. P. Keenan, R. C. Holcomb, P. A. Sprengeler, M. C. Guzman, J. L. Wood, P. J. Carrol, R. Hirschmann, *J. Am. Chem. Soc.* **1992**, 114, 10672–10674.
- [76] A. B. Smith III, R. Hirschmann, A. Pasternak, R. A. Akaishi, M. C. Guzman, D. R. Jones, T. P. Keenan, P. A. Sprengeler, P. L. Darke, E. A. Emini, M. K. Holloway, W. A. Schleif, *J. Med. Chem.* **1994**, 37, 215–218.
- [77] M. Hagihara, N. J. Anthony, T. J. Stout, J. Clardy, S. L. Schreiber, *J. Am. Chem. Soc.* **1992**, 114, 6568–6570.
- [78] M. Hagihara, S. L. Schreiber, *J. Am. Chem. Soc.* **1992**, 114, 6570–6571.

- [79] D. P. Rotella, *J. Am. Chem. Soc.* **1996**, *118*, 12246–12247.
- [80] P. Wipf, T. C. Henning, *J. Org. Chem.* **1997**, *62*, 1586–1587.
- [81] M. Chorev, M. Goodman, *Acc. Chem. Res.* **1993**, *26*, 266–273.
- [82] M. D. Fletcher, M. M. Campbell, *Chem. Rev.* **1998**, *98*, 763–795.
- [83] M. Chorev, M. Goodman, *Trends Biotechnol.* **1995**, *13*, 438–445.
- [84] G. Guichard, N. Benkirane, G. Zeder-Lutz, M. H. V. van Regenmortel, J.-P. Briand, S. Muller, *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 9765–9769.
- [85] R. B. Merrifield, P. Juvvadi, D. Andreu, J. Ubach, A. Boman, H. Boman, *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 3449–3453.
- [86] J. M. McDonnell, A. J. Beavil, G. A. Mackay, B. A. Jameson, R. Korngold, H. J. Gould, B. J. Sutton, *Nat. Struct. Biol.* **1996**, *3*, 419–425.
- [87] J. Gante, *Synthesis* **1989**, 405–413.
- [88] F. André, M. Marraud, T. Tsouloufis, S. J. Tzartos, G. Boussard, *J. Pept. Sci.* **1997**, *3*, 429–441.
- [89] A. Aubry, M. Marraud, *Biopolymers* **1989**, *28*, 109–122.
- [90] H. Han, K. J. Janda, *J. Am. Chem. Soc.* **1996**, *118*, 2539–2544.
- [91] M. Reggelin, V. Brenig, *Tetrahedron Lett.* **1996**, *37*, 6851–6852.
- [92] I. Peterson, J. P. Scott, *Tetrahedron Lett.* **1997**, *38*, 7441–7444.
- [93] I. Peterson, J. P. Scott, *Tetrahedron Lett.* **1997**, *38*, 7445–7448.
- [94] D. Seebach, J. L. Matthews, *J. Chem. Soc., Chem. Commun.* **1997**, 2015–2022.
- [95] E. Juaristi, *Enantioselective Synthesis of  $\beta$ -Amino Acids*, Wiley-VHC, New York, **1997**.
- [96] N. Sewald, K. D. Hiller, M. Körner, M. Findeisen, *J. Org. Chem.* **1998**, *63*, 7263–7274.
- [97] D. Seebach, S. Abele, K. Gademann, G. Guichard, T. Hintermann, B. Jaun, J. L. Matthews, J. V. Schreiber, L. Oberer, U. Hommel, H. Widmer, *Helv. Chim. Acta* **1998**, *81*, 932–982.
- [98] A. Müller, C. Vogt, N. Sewald, *Synthesis* **1998**, 837–841.
- [99] D. H. Appella, L. A. Christianson, D. A. Klein, D. R. Powell, X. Huang, J. J. Barchi Jr, S. H. Gellman, *Nature* **1997**, *387*, 381–384.
- [100] X. Daura, W. F. van Gusteren, D. Rigo, B. Jaun, D. Seebach, *Chem. Eur. J.* **1997**, *3*, 1410–1417.
- [101] D. Seebach, P. E. Ciceri, M. Overhand, B. Jaun, D. Rigo, L. Oberer, U. Hommel, R. Amstutz, H. Widmer, *Helv. Chim. Acta* **1996**, *79*, 2043–2066.
- [102] T. Hintermann, D. Seebach, *Chimia* **1997**, *51*, 244–247.
- [103] G. Guichard, S. Abele, D. Seebach, *Helv. Chim. Acta* **1998**, *81*, 187–206.
- [104] A. Müller, F. Schumann, M. Kokschi, N. Sewald, *Lett. Pept. Sci.* **1997**, *4*, 275–281.
- [105] A. Brinker, G. Raddatz, E. Weber, D. Stoll, J. Voigt, A. Müller, N. Sewald, G. Jung, P. Bohley, K.-H. Wiesmüller, in *Peptides 1998* (Eds.: S. Bajusz, F. Hudecz), Proc. 25<sup>th</sup> Eur. Pept. Symp., Akadémiai Kiadó, Budapest, **1999**, pp 842–843.
- [106] T. Hintermann, K. Gademann, B. Jaun, D. Seebach, *Helv. Chim. Acta* **1998**, *81*, 983–1002.
- [107] S. Hanessian, X. Luo, R. Schaum, S. Michnick, *J. Am. Chem. Soc.* **1998**, *120*, 8569–8570.
- [108] C. Toniolo, E. Benedetti, *Trends Biochem. Sci.* **1991**, *16*, 350–353.
- [109] A. Polese, F. Formaggio, M. Crisma, G. Valle, C. Toniolo, G. M. Bonora, Q. B. Broxterman, J. Kamphuis, *Chem. Eur. J.* **1996**, *2*, 1104–1111.
- [110] C. Toniolo, A. Bianco, F. Formaggio, M. Crisma, G. M. Bonora, E. Benedetti, V. Del Duca, M. Saviano, B. Di Blasio, C. Pedone, A. Aubry, *Bioorg. Med. Chem.* **1995**, *3*, 1211–1221.
- [111] H. Wenschuh, M. Beyermann, H. Haber, J. K. Seydel, E. Krause, M. Bienert, L. A. Carpino, A. El-Faham, F. Albericio, *J. Org. Chem.* **1995**, *60*, 405–410.
- [112] H. Wenschuh, M. Beyermann, R. Winter, M. Bienert, D. Ionescu, L. A. Carpino, *Tetrahedron Lett.* **1996**, *37*, 5483–5486.
- [113] M. Meldal, M. Renil, M. A. Giuliano, A. M. Jansson, E. Meinjohanns, J. Buchardt, A. Schleyer, in *Peptides 1996* (Eds.: R. Ramage, R. Epton), Proc. 24<sup>th</sup> Eur. Pept. Symp., Mayflower Scientific Ltd, Kingswinford, **1998**, pp. 141–152.
- [114] M. Meldal, M. A. Giuliano, A. M. Jansson, *Tetrahedron Lett.* **1997**, *38*, 2531–2534.
- [115] I. Ojima, *Acc. Chem. Res.* **1995**, *28*, 383–389.
- [116] H. E. Schoemaker, W. H. J. Boesten, B. Kaptein, E. C. Roos, Q. B. Broxterman, W. J. J. van den Tweel, J. Kamphuis, *Acta Chem. Scand.* **1996**, *50*, 225–233.
- [117] D. Seebach, A. R. Sting, M. Hoffmann, *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2709–2748.
- [118] A. Avenoza, J. H. Busto, C. Cativiela, J. M. Peregrina, F. Rodríguez, *Tetrahedron* **1998**, *54*, 11659–11674.
- [119] M. J. O'Donnell, C. Zhou, W. L. Scott, *J. Am. Chem. Soc.* **1996**, *118*, 6070–6071.



- [120] W. L. Scott, C. Zhou, Z. Fang, M. J. O'Donnell, *Tetrahedron Lett.* **1997**, 38, 3695–3698.
- [121] P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt, *Science* **1991**, 254, 1497–1500.
- [122] M. Egholm, O. Buchardt, L. Christensen, C. Beherens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen, *Nature* **1993**, 365, 566–568.
- [123] E. Uhlmann, A. Peyman, G. Breiphof, D. W. Will, *Angew. Chem. Int. Ed. Engl.* **1998**, 37, 2796–2823.
- [124] S. Sforza, G. Haaime, R. Marchelli, P. E. Nielsen, *Eur. J. Org. Chem.* **1999**, 197–204.
- [125] D. B. Sherman, A. F. Spatola, *J. Am. Chem. Soc.* **1990**, 112, 433–441.
- [126] T. Hoeg-Jensen, A. F. Spatola, A. Holm, *Int. J. Pept. Protein Res.* **1996**, 47, 190–200.
- [127] M. A. Shalaby, C. W. Grote, H. Rapoport, *J. Org. Chem.* **1996**, 61, 9045–9048.
- [128] H. Kessler, A. Geyer, H. Matter, M. Köck, *Int. J. Pept. Protein Res.* **1992**, 40, 25–40.
- [129] M. Schutkowski, S. Wöllner, G. Fischer, *Biochemistry* **1995**, 34, 13016–13026.
- [130] Y. Sasaki, D. H. Coy, *Peptides* **1987**, 8, 119–121.
- [131] D. H. Coy, S. J. Hocart, Y. Sasaki, *Tetrahedron* **1988**, 44, 835–844.
- [132] G. Guichard, N. Benkirane, R. Graff, S. Muller, J.-P. Briand, *Pept. Res.* **1994**, 7, 308–321.
- [133] N. Benkirane, G. Guichard, J.-P. Briand, S. Muller, *J. Biol. Chem.* **1996**, 271, 33218–33224.
- [134] G. Guichard, S. Calbo, S. Muller, P. Kourilsky, J.-P. Briand, J.-P. Abastado, *J. Biol. Chem.* **1995**, 270, 26057–26059.
- [135] S. L. Graham, S. J. deSolms, E. A. Giuliani, N. E. Kohl, S. D. Mosser, A. I. Oliff, D. P. Pompliano, E. Rands, M. J. Breslin, A. A. Deana, V. M. Garsky, T. H. Scholz, J. B. Gibbs, R. L. Smith, *J. Med. Chem.* **1994**, 37, 725–732.
- [136] J.-P. Meyer, P. Davis, K. B. Lee, F. Porreca, H. I. Yamamura, V. J. Hruby, *J. Med. Chem.* **1995**, 38, 3462–3468.
- [137] J. J. Wen, A. F. Spatola, *J. Pept. Res.* **1997**, 49, 3–14.
- [138] H. Drechsel, G. Jung, *J. Pept. Sci.* **1998**, 4, 147–181.
- [139] H. C. J. Ottenheim, J. D. M. Herscheid, *Chem. Rev.* **1986**, 86, 697–707.
- [140] A. Bianco, C. Zabel, P. Walden, G. Jung, *J. Pept. Sci.* **1998**, 4, 471–478.
- [141] S. Hin, C. Zabel, A. Bianco, G. Jung, P. Walden, *J. Immunol.* **1999**, in press.
- [142] N. A. Roberts, J. A. Martin, D. Kinchington, A. V. Broadhurst, J. C. Craig, I. B. Duncan, S. A. Galpin, B. K. Handa, J. Kay, A. Kröhn, R. W. Lambert, J. H. Merrett, J. S. Mills, K. E. B. Parkes, S. Redshaw, A. J. Ritchie, D. L. Taylor, G. J. Thomas, P. J. Machin, *Science* **1990**, 248, 358–361.
- [143] P. F. Alewood, R. I. Brinkworth, R. J. Dancer, B. Garnham, A. Jones, S. B. H. Kent, *Tetrahedron Lett.* **1992**, 33, 977–980.
- [144] P. Breton, M. Monsigny, R. Mayer, *Int. J. Pept. Protein Res.* **1990**, 35, 346–351.
- [145] B. H. Norman, J. S. Kroin, *J. Org. Chem.* **1996**, 61, 4990–4998.
- [146] A. P. Kaplan, P. A. Bartlett, *Biochemistry* **1991**, 30, 8165–8170.
- [147] D. A. Campbell, J. C. Bermak, *J. Am. Chem. Soc.* **1994**, 116, 6039–6040.
- [148] D. A. Campbell, J. C. Bermak, T. S. Burkoth, D. V. Patel, *J. Am. Chem. Soc.* **1995**, 117, 5381–5382.
- [149] E. M. Gordon, M. A. Gallop, D. V. Patel, *Acc. Chem. Res.* **1996**, 29, 144–154.
- [150] A. Yiotakis, A. Lecoq, S. Vassiliou, I. Raynal, P. Cuniassé, V. Dive, *J. Med. Chem.* **1994**, 37, 2713–2720.
- [151] A. Yiotakis, S. Vassiliou, J. Jiráček, V. Dive, *J. Org. Chem.* **1996**, 61, 6601–6605.
- [152] J. Jiráček, A. Yiotakis, B. Vincent, A. Lecoq, A. Nicolau, F. Checler, V. Dive, *J. Biol. Chem.* **1995**, 270, 21701–21706.

## 8 Glycopeptide and Oligosaccharide Libraries

*Phaedria M. St. Hilaire and Morten Meldal*

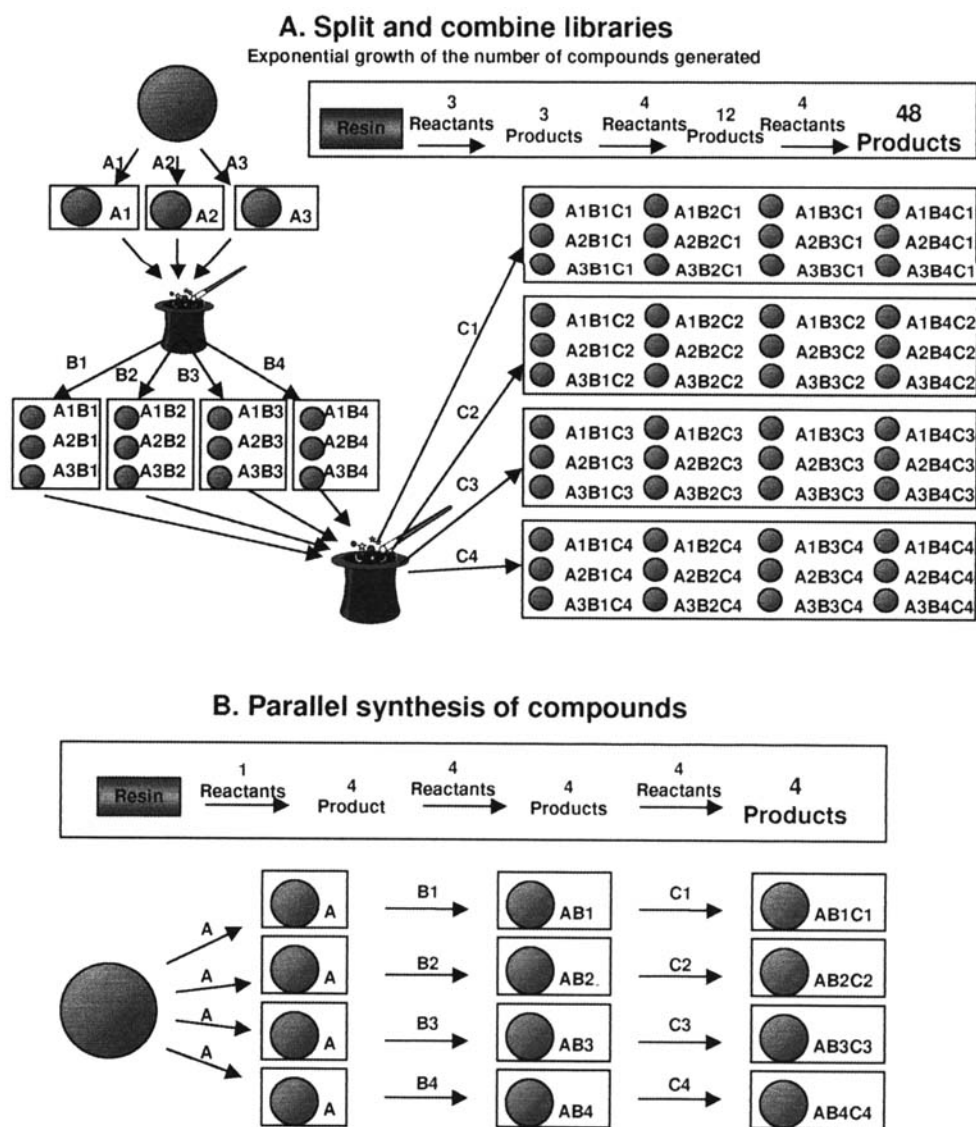
### 8.1 Parallel Arrays versus Libraries of Compounds

Combinatorial methodology has been a driving impetus in many areas of research and has gained considerable ground since its introduction at the turn of last decade. This progress has not been reflected in the field of oligosaccharide libraries probably due to the difficulty in their synthesis and analysis. Therefore glycopeptides as potential mimics of oligosaccharides have been exploited. A few reviews have appeared recently describing carbohydrate-based combinatorial libraries [1–5]. In this chapter, combinatorial libraries have been defined as assemblies of compounds, which are derived through a real combinatorial step leading to exponentially increasing numbers of compounds relative to the number of reagents used (Fig. 8.1A). This may be achieved either by reacting mixtures of reagents to form product mixtures, or by introducing mixing of intermediates on solid phase or in solution, *e.g.*, by the split-and-combine approach. Parallel syntheses are only briefly described; however, these are considered to be extensions of conventional synthesis of analogues and not libraries, as they do not easily give rise to the very large number of compounds required, *i.e.*, to take advantage of combinatorial exponential increase in products (Fig. 8.1B). Furthermore, in descriptions of parallel synthesis no distinction is made as to whether all analogues in an array were synthesized by stringent simultaneous reactions or not.

### 8.2 Proteins Recognizing Carbohydrates

Carbohydrate-binding proteins (CBPs) located in cellular membranes or recirculating in the cytosol or serum are involved in a variety of important biological functions. These include communication and intercellular adhesion, adhesion of bacteria or viruses, activation of the innate immune system, leukocyte rolling, hepatic clearance of aged serum proteins, and sorting of newly synthesized glycoproteins [6, 7]. Based on their mode of binding, mammalian CBPs have been divided into three major groups: the C-type, the S-type and the P-type lectins.

The calcium-dependent C-type lectins include the E-, L- and P-selectins, while the galectins are calcium-independent S-type lectins. The latter have a binding site sized for tight interaction with di- or trisaccharides. The former bind their ligand mainly through co-ordination of two vicinal hydroxyl groups of a single sugar moiety to a bound calcium



**Figure 8.1.** The conceptual difference between a combinatorial library and parallel arrays of compounds.

ion in the carbohydrate recognition domain (CRD), while the surrounding sugars of the oligosaccharide ligand add to the binding specificity through relatively weak additional interactions. Due to the simple nature of this interaction, the specificity of selectin binding is quite broad [8]. However, the high *in vivo* activity observed with the selectin binding could be associated with sulfation of a GlcNAc residue, or with the presentation of multiple ligands, though this has yet to be explained.

The calcium-independent receptors of the P-type are involved in the clearing and sorting of glycoproteins. These include the mannose 6-phosphate receptors (MPR), the hepatic Gal/GalNAc receptors [9, 10] and the Man/GlcNAc receptors isolated from liver extracts [11]. These receptors are truly multivalent and bind to the termini of bi-, tri- or tetra-antennary N-linked oligosaccharides. The binding of the natural ligands are in the nanomolar (nM) or micromolar ( $\mu\text{M}$ ) range, and they are highly specific.

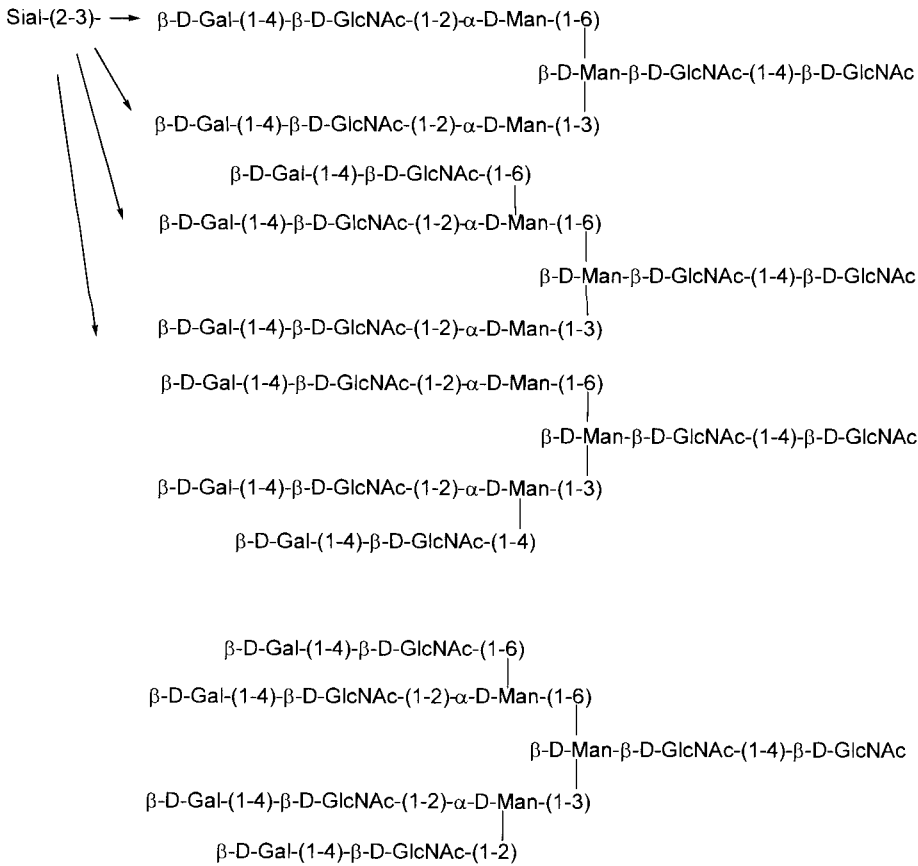
The collectins, a group of CBPs remaining in evolutionary terms from the early development of life, are mostly involved in the innate immune system [12]. They are large surface or serum proteins composed of bundles of structural collagen stalks with trimer heads containing calcium-dependent CRDs, and may be considered C-type by structural similarity. The trimer heads of the mannose-binding protein have been crystallized and the distance between the binding sites determined to be 53 Å [13]. The binding to simple mannose oligosaccharides is relatively weak, and the activation of the complement cascade requires the protein to interact multivalently with large polymannans on foreign cell surfaces, thus creating an activation template for complement processing enzymes [14]. Due to the unspecific nature of this multivalent interaction, and the long distance between the sites of interaction, multivalent synthetic ligands for these proteins are difficult to design and prepare. However, each CRD may be targeted with high-affinity monovalent ligands.

### 8.3 The Carbohydrate Ligands

Mammalian glycoproteins carry between 5% and 90% of glycan structures, and these fall into several major groups. There are the N-glycosylations, including high-mannose, complex and hybrid oligosaccharides (Fig. 8.2) and O-glycosylations including the mucins, the blood group determinants, and the proteoglycans. Glycolipids comprise another group of oligosaccharide signaling molecules. The high-mannose oligosaccharides are based on 1 $\rightarrow$ 2- $\alpha$ -mannosylation of the penta-mannose,  $\alpha$ -D-Man-(1 $\rightarrow$ 3)-{ $\alpha$ -D-Man-(1 $\rightarrow$ 3)-[ $\alpha$ -D-Man-(1 $\rightarrow$ 6)-] $\alpha$ -D-Man-(1 $\rightarrow$ 6)-} $\beta$ -D-Man-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ N), and these oligosaccharides exist as a mixture of conformational populations around the 1 $\rightarrow$ 6 bonds, as depicted in Figure 8.3. Either of these conformers may be the active one in contact with a receptor. Several levels of mannosylation, from five to at least 12 residues, can be observed in a single glycoprotein.

The complex antennary glycans are structures obtained by attachment of  $\beta$ -1-4-linked lactosamine unit to a  $\alpha$ -D-Man-(1 $\rightarrow$ 3)-{ $\alpha$ -D-Man-(1 $\rightarrow$ 6)-} $\beta$ -D-Man-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcNAc-(1 $\rightarrow$ N) core. These form di- to tetra-antennary oligosaccharides by attachment of oligo-LacNAc at the 2 and 6 positions of the 6-branch and 2 and 4 positions of the 3-branch of the mannose core and with the 2-positions of the branching mannoses always glycosylated. The terminal galactoses are often sialylated at the 3 or 6 position.

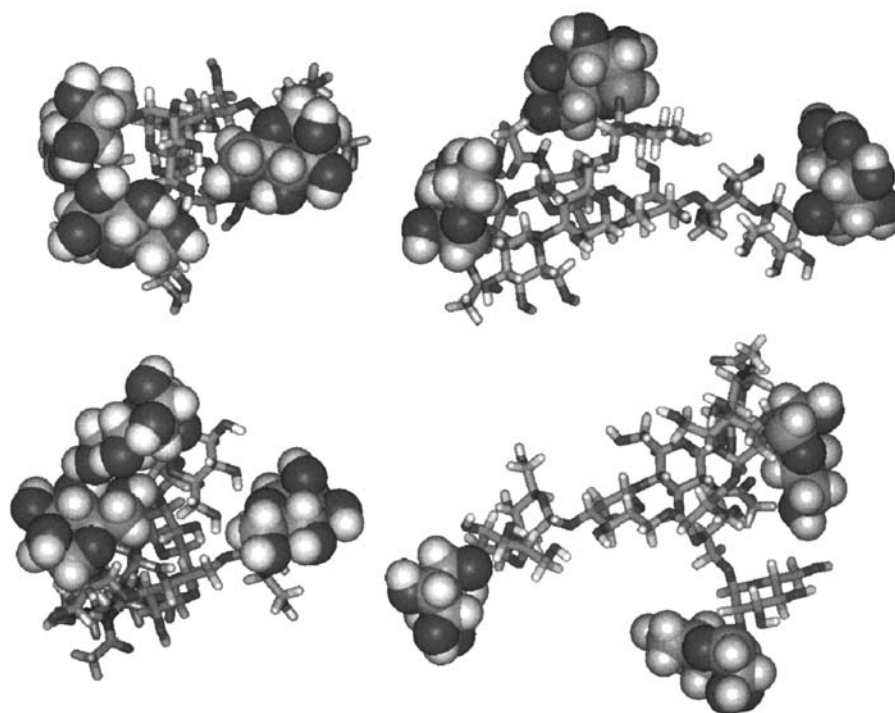
The tetra-antennary oligosaccharide has been shown to present an umbrella-like structure as one of the major conformational populations [15, 16]. The distance between terminal sugar residues interacting with receptors may vary from  $\sim$ 8 to  $\sim$ 30 Å. In the case of the phosphorylated high-mannose ligands for the MPR the distance is  $\sim$ 15 Å [17]. In biolog-



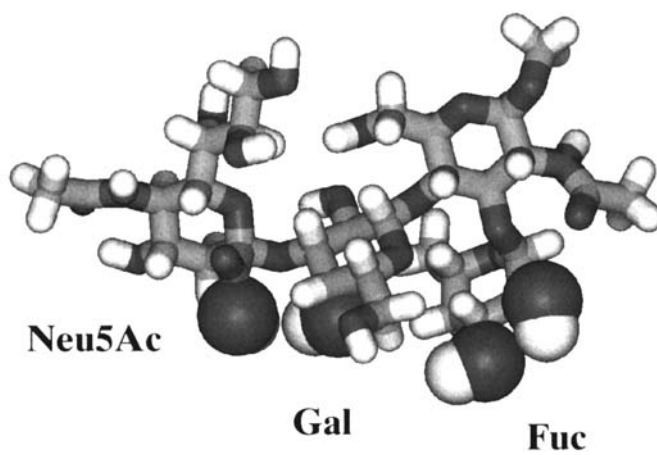
**Figure 8.2.** The most common complex oligosaccharides found on mammalian N-linked glycoproteins. The terminal Gal may or may not be sialylated [100].

ical systems, sialylation is a dynamic process during which sialidases deshield and sialyltransferases shield the underlying glycan structure for receptor recognition [18]. Most receptors recognize Gal/GalNAc or Man/GlcNAc; however, there are also receptors, *e.g.*, the macrophage sialic acid receptor for recognition of sialylated complex glycans [19].

The mucin oligosaccharides do not *per se* have many mammalian receptors. However, bacteria in the gastro-intestinal tract often display receptors that adhere to the mucin-coated epithelia [20]. Recognition of the aberrantly glycosylated forms present in malignant tissue is essential for the rejection of tumors [21]. The blood group determinants are essential for maintenance of the distinction between innate and foreign tissue. The most thoroughly investigated ligands are analogues of the S-Le<sup>x</sup> antigen [22] involved in leukocyte rolling, adhesion and shedding by interaction with the E-, P- or L-selectins, for which the optimal ligands and mode of recognition for biological activity are still sought. The fu-



**Figure 8.3.** Structure models of four different conformer populations of a tri-antennary high-mannose undecasaccharide (M9) as inferred from NOE connectivities of fragments. Terminal mannoses are highlighted as CPK-models.



**Figure 8.4.** Bound conformation of the sialyl Lewis x tetrasaccharide with tightly bound hydroxyl groups shown as spheres.

cose residue, one hydroxyl group from the central galactose unit, and the carboxyl group of the sialic acid, are essential for the interaction (Fig. 8.4).

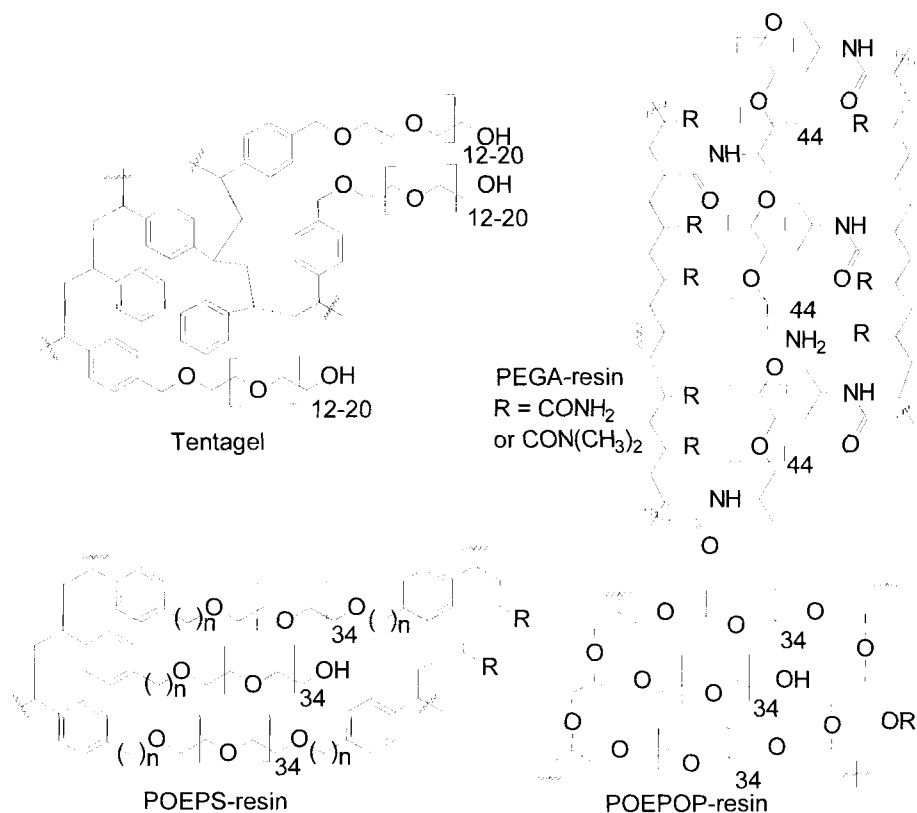
Other types of glycosylation could be important to mimic in the future. Thus, Ser and Thr residues of RNA-polymerases are dynamically glycosylated/deglycosylated or phosphorylated/dephosphorylated during the control of transcription [23, 24]. Proteoglycans act as template ligands for assembly of protein complexes leading to protein processing.

## 8.4 Supports for Solid-Phase Libraries

Parallel arrays, as well as libraries of glycopeptides and oligosaccharides, can be synthesized in solution or on solid phase. The advantages of the using solid-phase methods are well documented, and a detailed description has been recently published [25]. The choice of solid support is crucial to the success of solid-phase synthesis of glycopeptides and oligosaccharides and depends on the types of reactions to be carried out, as well as the screening methods to be employed. The two most commonly used types of solid supports are polystyrene (PS) -based (*e.g.*, Merrifield and Wang resin, TentaGel) or polyethylene glycol (PEG) -based (*e.g.*, PEGA, POEPOP, POEPS and SPOCC) (Fig. 8.5). Polystyrene resins are comprised of mainly 1–2% divinylbenzene cross-linked backbones with short linkers, causing reactions on PS-based resins to be strongly affected by the PS backbones. Polystyrene-based gels are unsuitable for direct on-bead screening due to their poor swelling in polar media and consequent exclusion of bio-molecules from their interior due to the hydrophobic core polymer. This is also true for PS-resins grafted with PEG such as ArgoGel or TentaGel [26], although swelling in water is improved. Furthermore, the polystyrene material absorbs light, thus interfering with fluorescence assays, and the hydrophobic core may result in non-specific protein binding.

PEG-based resins, on the other hand, are cross-linked with long-chain PEG macromonomers and the PEG chains present amino or hydroxyl functional groups. Therefore, the mechanical and chemical properties of the resins are strongly influenced by the PEG chains. Additionally, PEG chains are highly miscible with most solvents, creating a quasi-homogenous reaction medium. PEG-based resins are macroporous, thus allowing bio-molecules access to the entire bead. These resins are suitable for enzymatic reactions and screening in aqueous media.

Since PEG-based polymers confer the advantage of a support that is amenable to both synthesis as well as screening, efforts have been made to generate such polymers. They may be obtained by radical polymerization of long-chain PEG macromonomers to give PEG-cross-linked oligostyrene (POEPS-3) [27] or PEG-polyacrylamide (PEGA) copolymers [28, 29]. Recently, two novel types of gel supports obtained either by anion-catalyzed bulk polymerization of PEG derivatized with epichlorohydrin (POEPOP) [30] or by cation-catalyzed bulk polymerization of PEG derivatized with oxetane (SPOCC) [31] were introduced. The inert character of the polymers allows the application of harsh organic reactions. Because of the excellent swelling in aqueous buffers, all the above resins have been used in bioassays for enzymes, and PEGA resin showed no non-specific binding in protein binding studies [32–35].



**Figure 8.5.** The chemical structure of TentaGel and PEG-based resins.

Polystyrene-based resins have been used successfully for the solid-phase synthesis of glycopeptides and oligosaccharides. Experiments that compare the rates of a reaction on a PS-resin compared to TentaGel have revealed that the rate of reaction completely depends on the nature of the reaction itself [36]. Some reactions perform better on hydrophobic resins, while others are better on hydrophilic resins. Another issue of particular importance for the synthesis of oligosaccharides on solid phase is the influence of the solid support on the stereochemical outcome of the reaction. There are few detailed studies that address this issue; however, preliminary results are contradictory, but indicate that as in solution-phase reactions, the anomeric ratio is strongly influenced by the solvent, C-2 group on the donor and temperature [37–39]. In some cases there seems to be an enhancement of anomer selectivity on the solid phase [38].

At present, there are four examples of successful synthesis of oligosaccharide and glycopeptide libraries in the literature, and these have been generated using either TentaGel or a PEG-based resin (PEGA and POEPOP), thus enabling rapid solid-phase screening of the library in two of the cases.



## 8.5 Analytical Tools for Oligosaccharide and Glycopeptide Libraries

### 8.5.1 Tagging Techniques in Oligosaccharide Libraries

The number of methods for analyzing non-peptide libraries is increasing, and these fall into two categories: direct methods usually based on mass spectrometry and NMR spectroscopy; or indirect methods employing encoding, chemical [40–42], chemoluminescent [43] or others [44–46]. Many of the methods of chemical encoding are restricted by the additional synthetic effort required and the need to design orthogonal reaction conditions required for the two sets of syntheses. However, chemical coding is the only method which has been successfully used for identifying components in an oligosaccharide library [41]. The method utilizes tagging with a carbene insertion reaction to attach polyhalogenated aromatic tags [42]. The tags may be released photo-chemically from single beads and analyzed by GC-MS.

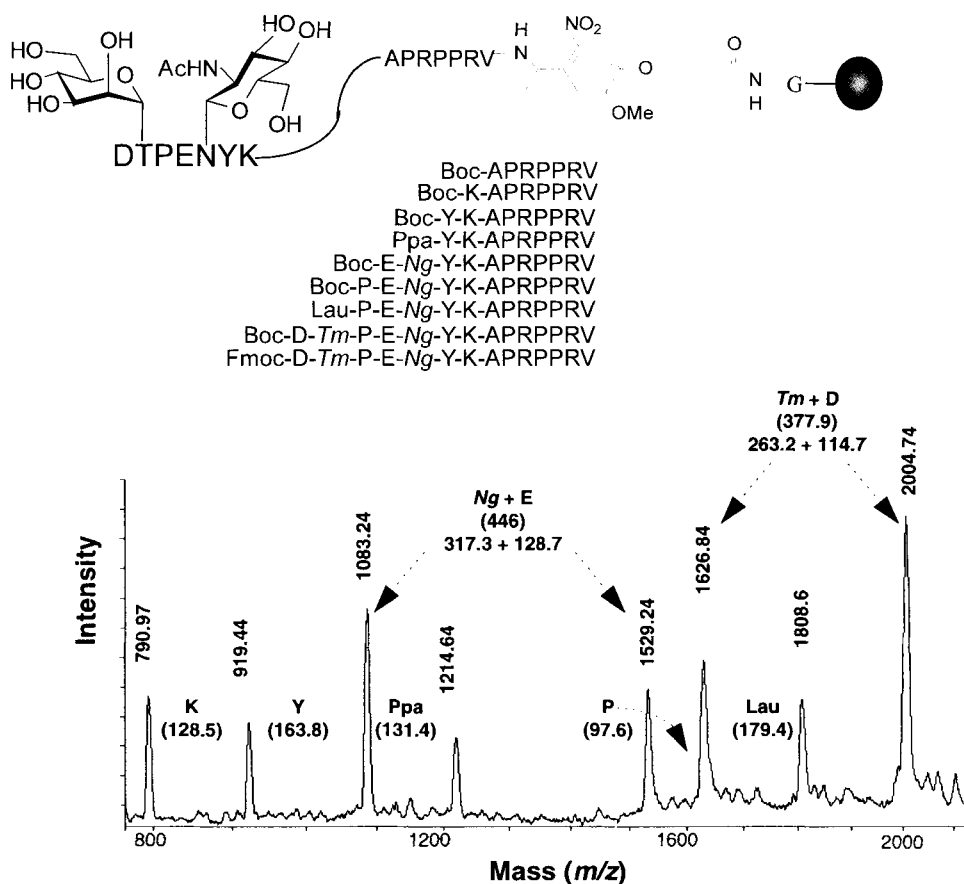
### 8.5.2 Analysis by Mass Spectrometry

The development of oligosaccharide and glycopeptide libraries obtained by the split-mix method is severely hampered by concurrent development of (or lack thereof) general, facile separation and characterization methodologies. Some headway has been made with chemical coding of the libraries, but very few direct methods of analysis exist. One of the promising methods that could be applied to the direct characterization of both types of libraries is mass spectrometry, more specifically Matrix Assisted Laser Desorption/Ionization Time of Flight mass spectrometry (MALDI-TOF-MS) or Electrospray Tandem Mass Spectrometry (ES/MS/MS). The high resolution achieved by Fourier Transform technology has allowed an ES/MS source to be used directly in characterization of peptide libraries [47, 48]. MALDI-TOF-MS and ES/MS have also been used to characterize glycopeptides [49–52] and complex oligosaccharides [53–57].

The analysis of carbohydrates by mass spectrometry presents special challenges in distinguishing between the branch points, isobaric monosaccharides, and the anomeric configuration of the glycosidic bond. Several methods have emerged using MALDI-TOF-MS. In one strategy, the oligosaccharide is treated sequentially with glycosidases, the mass of the resulting fragments recorded, and the identity of the original oligosaccharide determined by a process of deduction [56, 58]. Alternatively, the oligosaccharide fragments are generated during the ionization process in several ways: by post-source decay (PSD), delayed extraction with high-energy, collision-induced decomposition (CID), in-source fragmentation or tandem MALDI (high-energy CID) [53, 54]. In all cases, fragmentation occurs across the glycosidic bond in a manner that is dependent in part on the type (*i.e.*, 1→6 or 1→3) of linkage, the anomeric configuration, and the nature of the monosaccharide [57, 59]. Several types of ions are formed, the most during high-energy collisions. The best method for fragmentation is tandem MALDI since the parent ion can be selected at high resolution, there is high sensitivity, high-energy collisions and a clean background against which the entire range of ions can be seen. Interpretation of PSD spectra can be somewhat difficult and challenging for complex saccharides, but should be much more

simple if the oligosaccharide libraries are restricted to 3–4 monosaccharides. Because free neutral sugars lack a natural charge site, they have been regarded as difficult to ionize under electrospray conditions. However, the addition of sodium ions or derivatization of the glycan have been shown to promote ionization, leading to enhanced sensitivity [53]. Generally, analysis of oligosaccharides has been carried out using MALDI rather than ES/MS/MS.

Similarly, the analysis of glycopeptides has been carried out using primarily MALDI-PSD. Two strategies are employed: the glycan portion is first cleaved enzymatically or by base-catalyzed  $\beta$ -elimination [50, 52] and separately characterized by other means (MS or NMR), while the peptide is sequenced using PSD, or the entire glycopeptide is fragmented by PSD and the spectra analyzed [49, 51]. Again, spectral analysis can be time consuming. The only example to date of a facile, direct method of characterizing glycopeptide libraries generated using the building block approach is one that combines chemical coding (to



**Figure 8.6.** Representative mass spectra showing the ‘ladder’ of a bis-glycosylated glycopeptide obtained from screening of the deprotected glycopeptide library. The phenylpropionic acid (Ppa) and lauric acid (Lau) encode for Asn(GlcNAc) and Thr(Man), respectively.

determine the glycan portion) and a capping step to generate fragments for the mass analysis (ladder synthesis). An example of the mass spectrum of a glycopeptide from such a library is shown in Figure 8.6. The principle of this method and its application in the synthesis of a glycopeptide library are detailed in Section 8.11.

### 8.5.3 Structural Analysis of Compounds Linked to Single Beads by MAS-NMR

Magic Angle Spinning (MAS)-NMR spectroscopy is the analytical technique which holds most promise for direct identification of single compounds bound to one polymer particle. Gel-phase NMR can give respectable  $^{13}\text{C}$ -NMR spectra of large amounts of resin, but for single beads such analysis is not possible and proton spectra do not show useful resolution. The properties of the polymer have a major influence on the quality of the spectra recorded and in general, the more mobile and liquid-like the backbone of the polymer, the better the resolution of the  $^1\text{H}$ -NMR-spectra. This also holds true for MAS-NMR at intermediate spinning rates of 2000–4000 Hz, where compounds on the PEG-based resins are much more resolved than on polystyrene or even TentaGel [60]. In fact, solid-phase spectra displaying a resolution indistinguishable from that of solution spectra may be obtained with these resins [61]. The ability to analyze the structure of compounds by MAS-NMR on the solid support will not only quickly identify active compounds from a solid-phase compound library, it will also provide insight into the course of reactions on solid phase, and allow proper characterization of individual reaction steps used in the library construction [62].

So far, no complete structural elucidation of unknown compounds on solid phase has been published, and furthermore, the analysis presented all required substantial amounts of resin beads, except in one case where the so-called macro-beads of the TentaGel type were used [62]. Most other work has been performed with polystyrene beads, and although generally the 1D- $^1\text{H}$ -MAS-NMR spectra showed little or no resolution of coupled resonances, useful 2D-MAS-NMR spectra could be obtained [63]. Only known reaction products and known compounds on polystyrene or TentaGel have been analyzed so far [64–67]. However, the complete structural elucidation of an unknown heptapeptide on a single POEPOP bead [30] (10 nmol) has recently been achieved [68]. The minimum amount of resin-bound material required for a complete structural elucidation on PEG-based resins is 1–10 nmol of compound at 500 MHz, depending on the complexity of the structure and the resin [68]. Efficient use of MAS-NMR spectroscopy in combinatorial chemistry will require development of the solid support for the purpose of obtaining high resolution and S/N-ratios, as well as compatibility with organic chemistry and screening. The time-consuming nature of NMR-structural analysis also demands a more robust (no false positives) and quantitative biological assay in which the rank of the compound activities may directly be quantified in order to limit the number of structural analyses necessary [69]. MAS-NMR spectroscopy has been used for the monitoring of solid-phase glycosylation reactions of peptide templates with the aim of making 'libraries of libraries' using different glycans in consecutive glycosylations of peptide template libraries [70].

## 8.6 Introduction to Array and Library Synthesis of Oligosaccharides

While tremendous breakthroughs have been made in the generation of libraries comprised of small molecules, peptides or nucleotides, the synthesis and characterization of carbohydrate-based libraries has remained elusive. That is, if a library is defined in its strictest sense as the systematic assembly of a variety of building blocks to generate exponentially increasing numbers of products, in which all the possible combinations are most probably represented. This assembly is by no means a facile process, being complicated by the chemical complexity of the carbohydrate itself as well as issues related to linking together and modifying the building blocks. In the first case, monosaccharides are polyalcohols with five potential reactive sites, and therefore a well-planned orthogonal protecting group strategy must be implemented in order to selectively mask and unmask the relevant hydroxyl groups for glycosylation. Another salient issue deals with the fundamentally difficult nature of the reaction required for generating the library: glycosylation of the hydroxyl groups in a stereospecific manner and in high yields. While over the years, great strides have been made in the development of glycosyl donors and conditions for increased stereoselectivity and yields, each glycosylation reaction is essentially an individual problem, and it is difficult to develop a singular methodology that works perfectly in all cases. In this regard, two types of glycosyl donors, the glycosyl sulfoxides [71] and trichloroacetamidates [72], have proven very effective in coupling reactions with high yield and good stereoselectivity. Furthermore, both require homogeneous reaction conditions and may therefore be used for solid-phase glycosylation reactions. It must be noted however, that it is still difficult, even in solution, to generate important oligosaccharide linkages such as the  $\beta$ -Man,  $\alpha$ -NeuAc and  $\alpha$ -Kdo in greater than 50% yield. In solid-phase libraries another difficulty – which is not trivial – is the transfer of glycosylation reactions that perform well in solution, to the solid phase. These challenges notwithstanding, there have been significant breakthroughs in the generation of oligosaccharide libraries of small to medium size obtained through combinatorial steps.

## 8.7 Parallel Synthesis of Oligosaccharide Arrays

The structural diversity of oligosaccharides may be generated through parallel solution synthesis based on a common core entered at the first synthetic step. If each previous product portion after each successive step of deprotection and glycosylation is divided into a larger number of portions corresponding to number of next reaction sequences, a true library may be obtained. This recent, elegant solution approach [73] employs four orthogonal protecting groups for the monosaccharide hydroxyl groups. It has been used to produce a 45-member array of di- to pentasaccharides synthesized by parallel glycosylation of the common acceptor with seven different glycosyl donors. The key monosaccharide acceptor contained the chloroacetyl, levulinyl, *p*-methoxybenzyl and *tert*-butyldiphenylsilyl protecting groups. The various hydroxyl groups were sequentially unmasked and coupled with methythiophenyl donors under DMTST activation, yielding primarily the  $\alpha$ -anomer in very high yield (90–99%). As the size of the oligosaccharide increased however, there was a decline in selectivity and yield. Compounds synthesized were complete-

ly deprotected and characterized by mass spectrometry and NMR. This strategy as described in solution phase is extremely labor-intensive, and limited by the requirement of an orthogonal protection scheme. The highly orthogonal protection scheme can be avoided by use of a variety of enzymes with the desired glycosylation specificity. In particular, inversion of the reaction of glycohydrolases has been found to be useful and a medium-sized array also including the mucin core 1 and core 2 structures have been synthesized by this procedure [74].

Solid-phase methodology for parallel synthesis of oligosaccharides has also been devised [38, 39, 71, 75–77]. The orthogonal linkers used include those based on the TMSET protection group cleaved with  $\text{BF}_3$  etherate [77], the *p*-alkoxybenzyl group cleaved with cerium ammonium nitrate [76], the hydroxymethylbenzoic acid linker cleaved with dilute base, and photolabile linkers. Acid-labile linkers are sometimes problematic under acidic glycosylation conditions. The glycosylations must be carried out under homogeneous conditions, and those used are the trichloroacetimidate procedure [38, 72], the sulfoxide/ $\text{TiF}_2\text{O}$  procedure [71], pentenylglycoside procedure with iodonium activation (NIS, TES-OTf) [39] and the glycal procedures using either iodonium intermediates or epoxide [75] with Lewis acid activation. As these methods can be performed on solid phase, they would be particularly suited for oligosaccharide library synthesis.

## 8.8 Synthesis of Oligosaccharide Libraries

Oligosaccharide libraries have been generated successfully either in solution or on solid phase using primarily chemical synthesis. Solution-phase synthesis of carbohydrate libraries is carried out employing two main strategies: reacting protected glycosyl donors either with unprotected acceptors [5, 78, 79], or with protected acceptors [80]. The former strategy is designed to eliminate the need for the intense synthetic efforts required to synthesize orthogonally protected saccharide building blocks to ensure regioselectivity, thereby reducing the time and cost involved in the synthesis of oligosaccharide libraries. In the seminal work [5], a library of all six possible  $\alpha$ -linked trisaccharides was synthesized by coupling of the perbenzylated trichloroacetimidate to the unprotected LacNAc acceptor. Key features for the success of this methodology is the incorporation of a UV-active grease 'tail' in the acceptor to aid separation and detection of the products in the mixture, the use of polar solvents (DMF) for the glycosylation, and the termination of the reaction after 30% completion to prevent multiple glycosylation. Interestingly, the regioisomers were obtained in near-equimolar quantities, despite previous evidence suggesting that the different hydroxyl groups have highly differing reactivities toward glycosylation in non polar solvents. Minute (5%) amounts of the  $\beta$ -isomers were produced. However, changing the solvent to dioxane resulted in the formation of larger quantities of the  $\beta$ -anomers [79]. The mixtures were used to screen for the activity of glycosyl transferases. This was successfully carried out, even on substrates that made up 5% of the mixture. The main problems of this strategy are separation and identification of the products; despite the use of the hydrophobic tail, separation of the isomers was quite difficult. Additionally, precise identification of the products was possible by NMR only after comparison of the spectra with the individual product that had all been separately synthesized as single

compounds. This approach eliminates any advantage gained by the use of the library methodology, except for the initial screening.

While the glycosylation using unprotected acceptors can be truly random, caution must be exercised in order to prevent multiple glycosylation, which would make even more difficult the deconvolution and separation of the library components. Therefore, others have opted to use the approach wherein both the donor and acceptor are protected, thus enabling one to drive the reaction to completion without fear of multiple glycosylations. For example, a small library of 20 (including  $\alpha/\beta$  anomers) trisaccharides was generated using the vinyl latent-active strategy [80]. The 3-methyl allyl glycoside with benzyl protecting groups and one acetyl group was used as the common building block, and was converted into the glycosyl donor by isomerization to the vinyl glycoside or to the acceptor by removal of the acetyl group. TMSOTf-promoted glycosylation reactions proceeded in greater than 70% yield, producing a 1:1 mixture of  $\alpha/\beta$  anomers. This procedure is in fact a combination of parallel synthesis and the split-mix methodology. Because a common monosaccharide building block with only two different protecting groups was used, very limited diversity was introduced by the split-mix step. At the end of the synthesis of the trisaccharide library, the individual members were not separated but analyzed simply on the basis of the trisaccharide mass and monosaccharide composition. Mass spectrometry was also used to analyze completion of the glycosylation reactions.

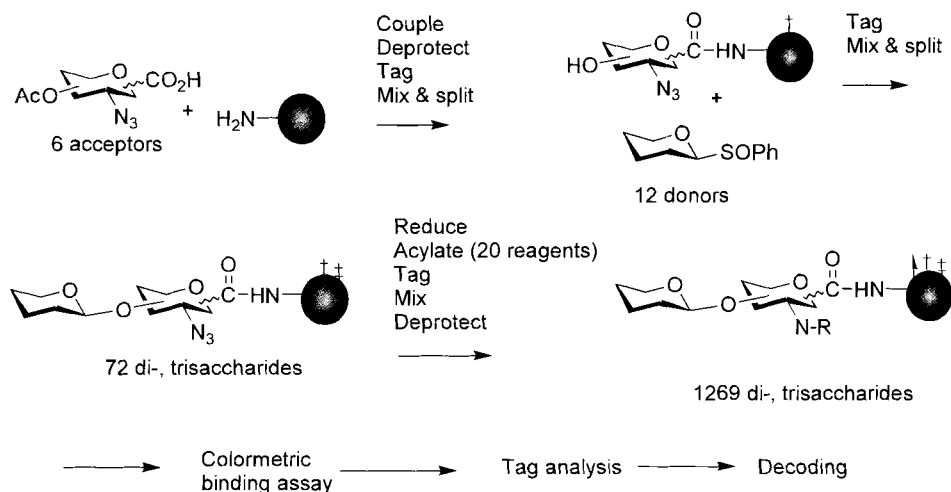
In another example of utilizing only the split-mix methodology, a small, four-member library of 2,6 dideoxy trisaccharides was generated also using a common TMS-protected glycal building block [81]. Thus, the glycal was first converted to its 6-trifluoroacetamido-hexyl 2-iodo glycoside by IDCP activation. The protecting groups in the 3 and 4 positions were removed and then simultaneously glycosylated stereoselectively at low temperature with an equimolar amount of the glycal donor to yield a mixture of the two  $\alpha$ -linked regio-disaccharides in equal amounts. Although reactions were allowed to proceed for as long as possible, surprisingly, there was no mention of the formation of trisaccharides resulting from glycosylation of both the 3 and 4 positions, or of unreacted starting material. The remaining hydroxyl groups were then acetylated and the TMS groups removed. The new hydroxyl groups were glycosylated as previously to yield a mixture of four 2-iodo trisaccharides. The iodo group was reduced by treatment with tributyl tin hydride. Again, in this case, the products were not separated but analyzed as a mixture using mass spectrometry and NMR.

Although most oligosaccharide libraries have been synthesized in solution, there are two examples of solid-phase synthesis of oligosaccharide libraries. Oligosaccharide synthesis on solid phase is not a straightforward extrapolation from solution phase, and a great deal of effort is required to develop conditions for stereoselective glycosylations in high yield. In the first example, a small library of 12 different trisaccharides (including  $\alpha/\beta$  anomers) was synthesized on TentaGel resin using the thioethyl glycoside as the donor by the split-mix method [82]. This library was very limited in its diversity since only one hydroxyl group was used as the acceptor. The starting thioethyl glycoside was attached to the solid support via a succinic ester linker at C-6, thus allowing it to be used both as the donor and acceptor in glycosylation reactions. Upon completion of library synthesis, the mixture was cleaved by treatment with base, and the mixture analyzed by determination of monosaccharide composition. While solid-phase synthesis of libraries by the split-mix

methodology confers the advantage of compartmentalization of the products (since each bead may be obtained with an individually distinct compound – one-bead–one-compound), accurate identification of the compound attached is a problem which escalates with library size. In order to be sure of the nature of the active ligand after the screening process, it is crucial that each glycosylation step proceeds stereospecifically.

In a more extensive effort, the split-mix method was used for the synthesis of ~1300 di- and trisaccharides as a ‘one-bead–one-compound’ library (Fig. 8.7) [41]. Six suitably protected azide-containing monosaccharides were attached to TentaGel and glycosylated with 12 different mono- and disaccharide sulfoxide donors to yield 72 different compounds. Additional diversity was achieved by reduction of the azide and acylation of the oligosaccharide amine with 20 different carboxylic acids. The identity of compounds attached to the beads was established using a halo-aromatic tagging technique employing carbene insertion reactions [42]. While this technique provided a way of decoding which carbohydrates were coupled to each other on each bead, the precise nature of the glycosidic linkage could not be conclusively determined. It was assumed, based on previous experience, that the glycosylation reactions under those conditions used were completely stereoselective. The library was screened for active ligands for *Burhinia purpurea* lectin using a solid-phase enzyme-linked assay. The screen yielded several non natural oligosaccharide ligands for the lectin, but surprisingly, the known natural ligand that should be present in the library was not detected. The natural ligand and the mimics found all showed comparable binding affinity.

There are a few examples where mixtures of complex oligosaccharide have been isolated from natural sources such as bovine fetuin, porcine fibrinogen or conalbumin, and the mixture converted to the glycosyl amine and then coupled to an amino acid before separation of the individual components by HPLC. In one example described, 15 N-linked



**Figure 8.7.** A true solid-phase oligosaccharide library [41], showing exponential increase in number of compounds with each reaction step. The library was screened for active compounds using a solid-phase, enzyme-linked assay and the components were identified using chemical coding.

complex oligosaccharides containing 6–11 monosaccharide units were isolated from hen ovalbumin and characterized by a combination of high-field NMR and FAB MS after HPLC separation [83].

Synthesis of oligosaccharide libraries both in solution and on solid phase is by no means trivial. The main impediment to progress in carbohydrate libraries has been the yields and selectivities of the glycosylation reactions and the technology for separation and analysis. In all the examples described above, apart from one instance, the library of compounds was never separated, nor were the individual products precisely identified and characterized. A recent method of using orthogonal protecting groups in parallel synthesis in solution [73] has great potential for library synthesis on solid phase, provided that the analytical methods (see Section 8.8.5) can be developed. It should also be noted that only in one case of oligosaccharide libraries described [41] has the exponential increase of compound numbers by consecutive reactions been exploited to any significant extent.

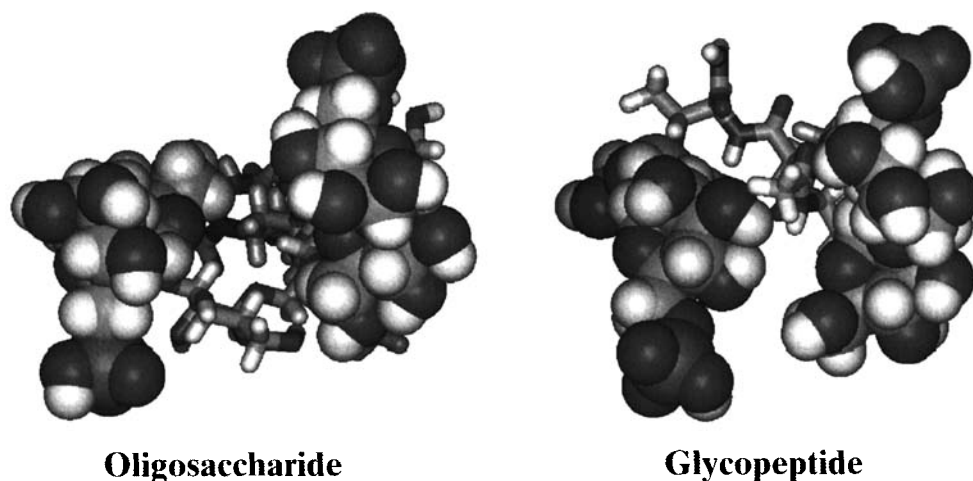
## 8.9 Glycopeptide Templates as Oligosaccharide Mimetics

During the quest for a carbohydrate-based therapeutic, the synthesis of large numbers of diverse oligosaccharides is an extremely expensive, synthetically challenging, and time-consuming process. Whereas the use of oligosaccharide libraries would save time during synthetic operations and screening, the expense, synthetic challenge and analysis of active compounds from the library would still be problematic. Furthermore, it has been shown that modifications of the natural oligosaccharide ligand rarely confer any increase in binding efficacy.

When glycopeptides were first introduced as mimetics of oligosaccharides [84] it was envisaged that the saccharide moiety would provide the specificity of the binding by directing the ligand to the oligosaccharide binding site, while the peptide would function as a scaffold for optimal orientation of the glycan. Peptide ligands are known generally to bind with high affinity to their receptors, and it is expected that the glycopeptide could furthermore interact favorably with the receptor through the peptide scaffold, leading to increased binding affinity [17]. The source of the enhanced affinity of peptides may lie in the ability to perform an induced fit by the relatively limited flexibility around peptide bonds, leading to fast on-rates for the binding. In the peptide architecture, this flexibility is obtained with the minimum penalty in loss of entropy. Conversely, the carbohydrates generally display conformational assemblies located closely around one global energy minimum conformation.

The principle was consolidated through binding studies with an array of phosphorylated glycopeptides and the divalent mannose 6-phosphate receptor [17]. Ligands binding 20-fold stronger than the natural phosphorylated pentamannose ligand were found. The structural similarity between the oligosaccharide and the most active glycopeptide ligand was confirmed by molecular dynamics (Fig. 8.8). It was demonstrated that a minimum of two disaccharides on a scaffold were required for the specific interaction with the receptor. Attempts to increase the affinity through cyclization of the peptide were unsuccessful [85].





**Figure 8.8.** MD-models of phosphorylated high-mannose  $M_7$  and a glycopeptide mimic, as seen from the point of view of the mannose 6-phosphate receptor. Disaccharide phosphates are emphasized as spheres. The structure of the mimic was obtained after 40 ps at 400 K and 10 ps at 300 K using the Insight/Discover program package.

Similar results were later obtained with glycopeptides designed for binding to the selectins. A glycopeptide mimic of the sialyl-Le<sup>x</sup> tetrasaccharide containing fucose on a peptide scaffold had a greater than ten-fold increased binding affinity for E-selectin [86]. There was no significant further increase in affinity when the ligands were immobilized in a polymeric multivalent arrangement in liposomes [87]. In another example, a high-affinity di-valent ligand adhesin containing Gal- $\alpha$ -1-4-Gal- $\alpha$ -linked via peptide bonds to an aromatic nucleus was prepared for *Streptococcus suis*. The interaction was found to be truly divalent since the preparation of putative tetravalent ligands showing no significant increase in binding [20]. Glycopeptide mimics of galactose-containing oligosaccharides afforded 1.7  $\mu$ M inhibitors of galactosidase [88].

Since glycopeptides are excellent mimics of the complex oligosaccharides involved in myriad important biological events, the affinity of glycopeptides to carbohydrate binding proteins may be utilized in a library format to identify high-affinity ligands. The ease with which glycopeptides are synthesized using preactivated amino acids and glycosylated amino acid building blocks can, by careful assembly of a library, ensure the generation of a single compound in each bead. For the preparation of glycosyl amino acid building blocks, the glycosylation of Fmoc-amino acid-OPfp esters have proved a general and versatile method useful for the preparation of complex compounds for direct incorporation into the glycopeptide libraries [89]. In addition, the ease of characterization of active compounds make libraries of glycopeptide templates a very attractive alternative to oligosaccharide libraries.

## 8.10 Parallel Synthesis of Glycopeptide Arrays

When considerable *a priori* structural knowledge of a protein–carbohydrate interaction is available, it is possible to synthesize a range of active glycopeptide analogues by biased design using a parallel synthesis approach. The requisite knowledge is the nature of the dominant sugars involved in the interaction, and preferably their spatial arrangement in the receptor interaction. With such prominent information available from X-ray crystal data or from transfer NOE-NMR-experiments, a range of high-affinity ligands of the glycopeptide type have been developed by parallel synthesis for several receptors.

As described above, 20 bidentate glycopeptide ligands for the MPR were synthesized by parallel synthesis and high-affinity ligands with approximately 20-fold increased affinity for the receptor were obtained [17]. The array of analogues contained glycosylated linear tri- to pentapeptides and cyclic hexa- to octapeptides. The glycans were phosphorylated 1→2 and 1→6-linked mannodisaccharides and phosphorylated mannose. Glycopeptides were synthesized by solid-phase multiple column peptide synthesis (MCPS) using Fmoc-amino acid-OPfp esters and a PEGA resin. The glycosylated building blocks were protected as acetates or benzoates and the phosphate with trichloroethyl groups. Deprotected glycopeptides were obtained in high yield and purity. Assaying the compounds clearly showed the necessity to have sufficient, but not excessive, flexibility in the scaffold and two 6-P- $\alpha$ -Man-(1→2)- $\alpha$ -Man disaccharides were essential for the interaction.

An array of 120 galactose-containing compounds (30 mixtures of four diastereomers) was prepared in parallel by base-catalyzed Michael addition of  $\beta$ -D-(C<sub>12</sub>H<sub>25</sub>CO)<sub>4</sub>Gal-SAc to 4 different unsaturated ketones and  $\alpha$ -chloro ketones, followed by reductive amination with six amino acids. The 30 products were purified by solid-phase extraction and tested as inhibitors of galactosidase. The galactose provided the specificity for the enzyme, while the affinity was obtained through interaction with the aglyconic scaffold. In this way, 1.7  $\mu$ M inhibitors of the enzyme activity were obtained [88].

An array of glycopeptide ligands binding to MHC class II E<sup>k</sup> and interacting with the T-cell receptor (Fig. 8.9(a)) and in which the glycan part was varied, were synthesized by MCPS as described for the MPR above [90, 91]. The glycosylated building blocks were synthesized by glycosylation of Fmoc-amino acid-OPfp esters, and these were used directly for coupling reactions. The glycans included the mucin Tn ( $\alpha$ -D-GalNAc) antigen, core 1, core 2, core 3 and core 4 structures, as well as a large array of mucin-unrelated saccharides. Using the panel of glycopeptides shown in Figure 8.9(a), it was found that mainly the T and Tn antigens were recognized and were considerably more immunogenic than other glycans (Fig. 8.9(b)) [21]. The results indicated a repertoire of T-cell receptors specifically directed towards tumor-associated antigens and elongation with GlcNAc (*e.g.*, core 2 and 4) gave non immunogenic compounds.

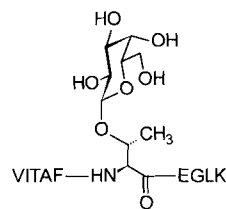
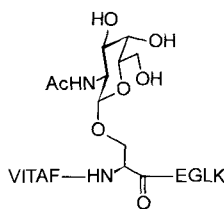
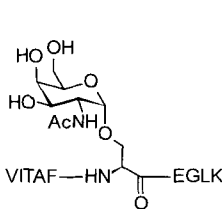
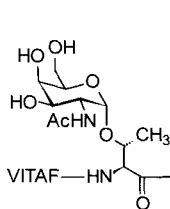
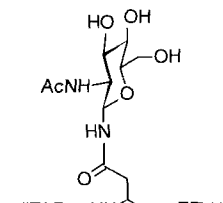
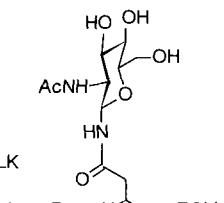
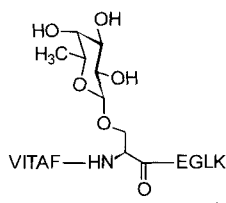
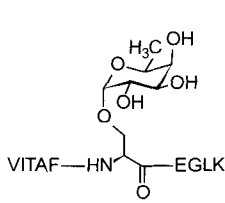
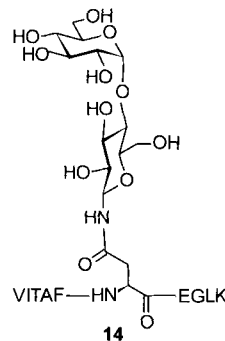
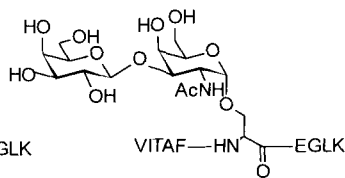
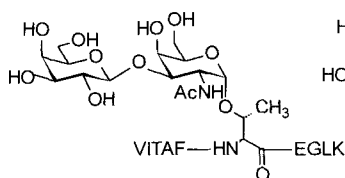
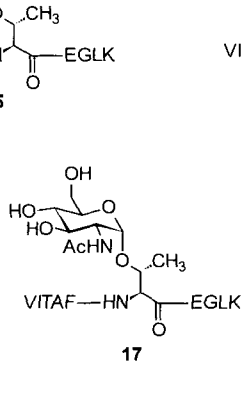
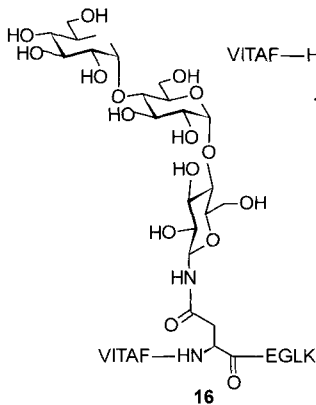
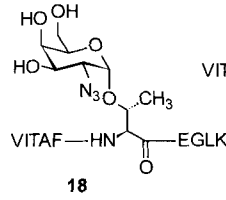
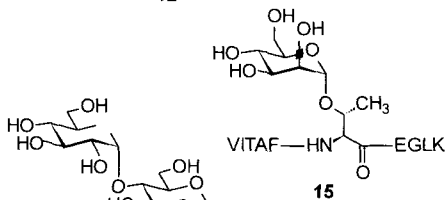
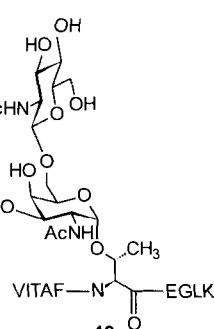
Parallel synthesis of 62 different fucosylated tripeptides resulted in two ligands with submicromolar affinity for the P-selectin; however, the desired activity for the E-selectin was not observed [92]. For the E-selectin selectivity it was necessary to incorporate a hydroxyl group which mimics the 4-hydroxyl of the central Gal in S-Le<sup>x</sup>, in addition to a Fuc-residue and a carboxylate to obtain ligands with >10-fold increased activity over

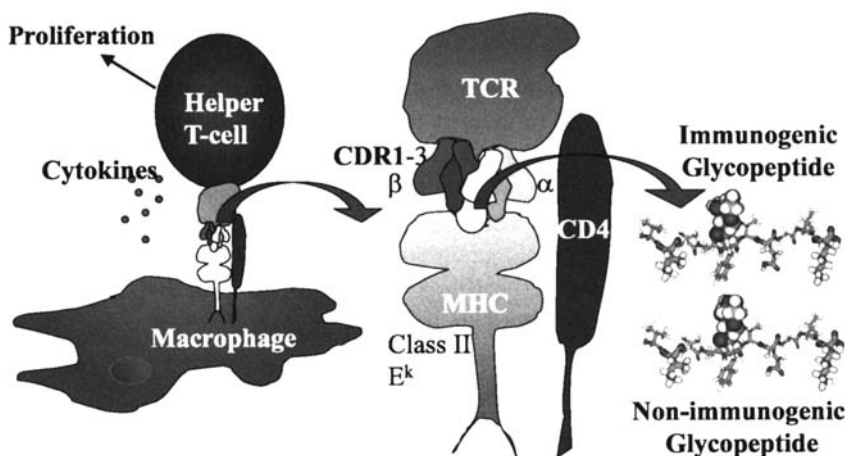
**Hb(67-76)**

VITAFNEGLK

VITAFTEGLK

VITAFSEGLK

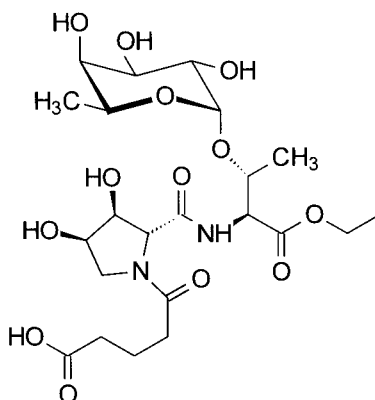
**1****2****3****4****5****6****7****8****9****10****11****12****13****14****18****16****17****19**



**Figure 8.9** (a) Mammalian histocompatibility complex (MHC) binding array of glycopeptides used in the study of glycan specificity of the T-cell receptor (TCR) recognition. Only the Tn and T antigens were strong immunogens. (b) Glycopeptide/MHC/TCR interaction. The minor modification of substituting  $\alpha$ -GalNAc with  $\alpha$ -GlcNAc abolished immunogenicity.

that of the S-Le<sup>x</sup> tetrasaccharide [86]. The best ligands were obtained from Thr( $\alpha$ -Fuc)-OEt, which was first N-acylated with a hydroxyl amino acid and then elongated with a diacid to furnish the acid mimic of the sialic acid carboxylate (Fig. 8.10).

In an elegant and different approach, an array of C-linked, glycopeptide-like mimics of SLe<sup>x</sup> were synthesized in parallel by an Ugi four-component reaction. Reaction of different anomeric two- or three-carbon extended C-glycosyl aldehydes or acids with resin-bound amine, isonitriles and with other acids or aldehydes, respectively, yielded an array of C-linked analogues. The method is easy to perform in good yield on solid phase. How-



**Figure 8.10.** A ligand with high affinity for E-selectin. The receptor interacts with the 2- and 3-hydroxyls of Fuc, probably the hydroxyls of the dihydroxyproline ring and with the carboxylate.

ever, this strategy can only be used for a mixed library, and not for a 'one-bead-one-compound' library. It is derived through a multicomponent reaction where all components are introduced in a single step, and deconvolution strategies such as positional scanning or iterative synthesis are required to identify the components of the library. On the other hand, deconvolution is feasible with a fast screen because the products are generated rapidly.

## 8.11 Preparation and Analysis of Solid-Phase Glycopeptide Template Libraries

While parallel synthesis of an array of glycopeptides is readily achieved by implementation of the building block approach [89], glycopeptide library synthesis in a combinatorial way via the split-mix method has yet to become routine. The difficulty lies in the structural analysis of the vast number of compounds each generated in picomolar quantities on a single bead. Whereas peptides on beads can be conveniently analyzed by solid-phase ladder sequencing [93] or Edman degradation, neither of these methods is suitable for glycopeptides because of the instability of the glycosidic bond under the acidic and basic conditions employed.

Presently, there is only one example of a combinatorially generated glycopeptide library in the literature [94]. Generation of a 300 000-member library and analysis of its components was made possible in part by the development of a mass spectrometry-based technique for identifying the sequence of the individual glycopeptides in the library. In this method, the synthetic history of the glycopeptide is captured on the beads by capping a small percentage of the growing oligomer chain in each synthetic step [95]. Thus, a series of related fragments rather than a single compound is generated on the bead (see Fig. 8.6). The difficulty arises when this technique is applied for a glycopeptide library in which the amino nucleophiles present significant reactivity differences. Both post- and pre-capping techniques are doomed to fail. Even the use of a simple in-situ capping agent in mixture with the activated amino acids was problematic, and resulted in anywhere from 0% to 100% capping. The capping method that was eventually developed is based on the use of structurally closely related capping agents and building blocks in mixtures of 90% Fmoc-amino acid and 10% of the same Boc-amino acid (activated by TBTU) for all the couplings of natural amino acids [96].

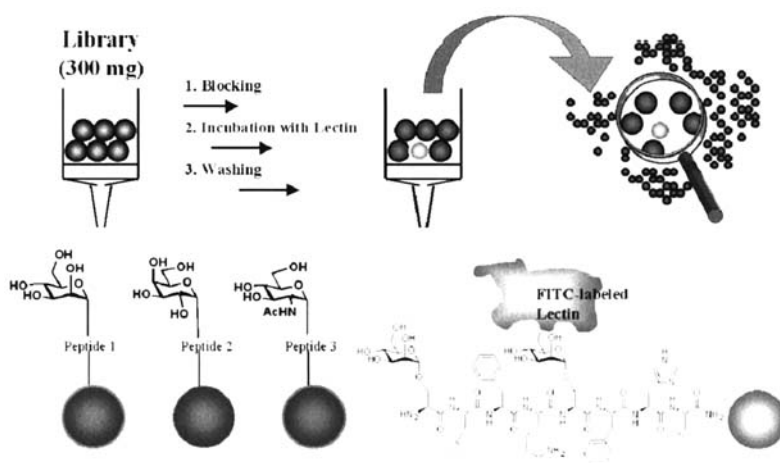
The glycosylated Fmoc-amino acid-OPfp esters were encoded by capping with selected carboxylic acid-OPfp esters with masses different from those of the regular amino acids. The reactivity of the carboxylic acids ester labels was first determined by reaction of a peptide resin with a mixture of a representative glycosylated building block and each of the carboxylic acid OPfp esters. As expected, the reactivities varied ranging from three times less to several times more reactive than the glycosylated amino acid-OPfp ester. The ratio of capping agent to building block used for coupling in the libraries was then adjusted accordingly.

The library was linked to the solid support via the photo-labile linker (see Fig. 8.6) which was cleaved at the 337 nm of the nitrogen laser used in MALDI-TOF mass spectrometers (optimal wavelength for cleavage of linker is 365 nm). The linker facilitated the

immediate analysis of compounds released from the resin beads after application of these to the laser target in the presence  $\alpha$ -cyano-4-hydroxy-cinnamic acid (CHC) matrix. Alternatively, single beads on the target were first irradiated for 20 minutes with a Hg lamp prior to mass analysis.

The synthesis of the libraries was carried out in a MCPS Teflon library generator with 20 columns and a mixing chamber above the columns [97]. During mixing, the columns and half the volume of the chamber are filled with solvent, a circular lid sealed with an O-ring is fitted, and the resin thoroughly agitated for 15 minutes. Washing solvents and 20% piperidine solution for deprotection are added from dispenser bottles connected to 20-line dispenser heads mounted on an aluminum frame. At the end of the assembly the library is deprotected during a 2-hour treatment with a mixture of TFA and scavengers. The *O*-acetates of the sugar are removed with hydrazine hydrate in methanol for 2 hours. The purity of the library was assessed by MALDI-TOF mass spectrometry by collection and analysis of a few beads. Most of the beads collected afforded spectra as ladders which could easily be deciphered using mass difference assignment software from Bioanalysis.

In a departure from the building block approach, glycopeptide libraries can also be obtained by glycosylation of a pre-attached glycan or the hydroxyl group of an amino acid side chain of a peptide library. In preliminary studies [38], good yields of glycopeptides containing di- and trisaccharides were obtained using 5–8 equivalents of the perbenzoylated trichloroacetimidate donor to glycosylate a known glycopeptide. Glycosylation was attempted on four resins: Polyhipe, TentaGel, PEGA 1900 and Macrosorb-SPR250, but was successful only on PEGA and Polyhipe. Interestingly, product anomeric ratio appears to depend on the resin. Attempts [38, 98] to effect direct glycosylation of the amino acid side chain hydroxyl were partially successful, albeit in low yield and initially only on Polyhipe resin. The failure of polar resins such as PEGA may be due to the amide backbone, which interferes with the carbocation intermediate. However, since polar resins are re-

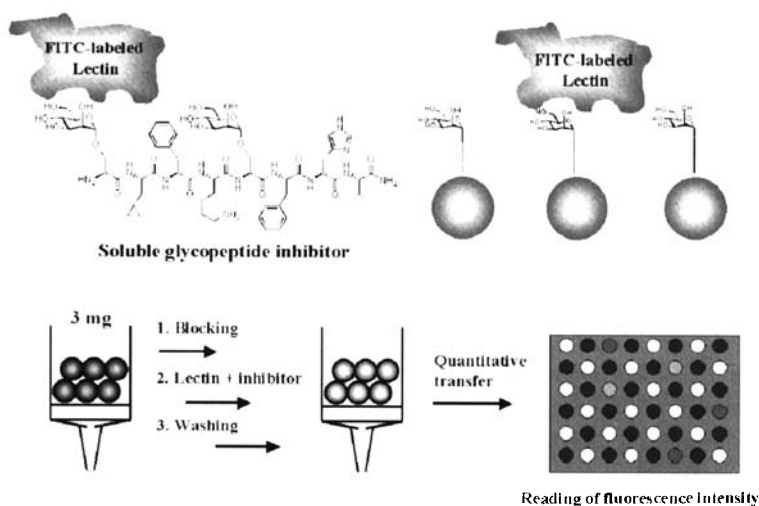


**Figure 8.11.** Screening of a glycopeptide library using a fluorescent-labeled lectin and ligands bound to PEGA-beads.

quired for solid-phase bioassays, new types of polar resin containing only ether bonds were developed for the solid-phase glycosylation of peptides. Quantitative glycosylation of a known peptide [99] was achieved on a novel PEG-based resin (POEPOP) containing only ether bonds [30] using 5–8 equivalents of the peracetylated or benzoyleated trichloroacetimidate. In a ‘one-bead–one-compound’ approach, two resin-bound peptides bearing protected and unprotected hydroxyl groups were first glycosylated with galactose and then with fucose after deprotection of the second hydroxyl group, affording a small library of four glycopeptides. The glycopeptides were cleaved off, separated and characterized by mass spectrometry. While solid-phase glycosylation is undoubtedly a feasible alternative for the generations of truly random glycopeptide library with diversity in both the peptide and glycan portions, analysis of these libraries will present quite a challenge. One possibility is the use of fragmentation of the compounds by mass spectroscopy, or tandem mass spectroscopy; another is the use of MAS-NMR, as discussed in Section 8.5.3.

## 8.12 Screening of a Solid-Phase Glycopeptide Library

An important requirement for the successful application of the combinatorial library approach to the drug discovery process, is the ability to test the library in high throughput screening (HTS) procedures. Rapid screening can be achieved if it is carried out on the resin-bound ligand. With the development of new polar supports for solid-phase synthesis, it is possible to screen libraries with the ligand still attached to the solid support. The library generated as described in the preceding section was incubated with the fluores-



**Figure 8.12.** Inhibition of FITC-labeled lectin binding to resin-bound mannose by soluble glycopeptides obtained from library screen. Percent inhibition was quantified by recording lectin fluorescence. Only every second well of the microtiter plate is used, and non fluorescent beads indicate good inhibitors.

cent-labeled lectin from *Lathyrus odoratus* (Fig. 8.11). The most fluorescent beads were collected, crushed on a target, mixed with CHC-matrix and subjected to MALDI-TOF mass spectrometry.

One-third of the beads contained no glycan, and were pure peptides. Two-thirds of the beads contained either mannose, N-acetylglucosamine, or both. The active sequences were synthesized, and binding of the lectin to resin-bound glycopeptides assessed in a solid-phase fluorescent assay. The most active compounds were glycopeptides containing only mannose (T( $\alpha$ -D-Man)ALKPTHV, LHGGFT( $\alpha$ -D-Man)HV, T( $\alpha$ -D-Man)-EHKGSKV, GT( $\alpha$ -D-Man)-FPGLAV, and T( $\alpha$ -D Man)-LFKGFHV), displaying up to a 25-fold increase in fluorescence than lectin binding to resin-bound mannose. Binding of the lectin to resin-bound mannose was inhibited by the active glycopeptides synthesized (Fig. 8.12), suggesting that the glycopeptides and the natural carbohydrate ligand bind to the same or to related binding sites of the lectin and that glycopeptides can mimic the activity of carbohydrate ligands.

## 8.13 Conclusions

During recent years, novel methodologies have been presented for the synthesis of oligosaccharide and glycopeptide libraries, as well as for the identification of active oligosaccharide or glycopeptide ligands isolated from solid-phase libraries for oligosaccharide-binding proteins. The main limitations which have been encountered, are related to the lack of proper methods for miniaturization and analysis of compounds on solid phase. However, these methods are currently being developed. The use of larger solution-phase libraries requires deconvolution synthesis, which is not practical with oligosaccharide synthesis due to its complexity and lengthy nature. Oligosaccharide synthesis on solid phase is difficult and complicated. On the other hand, glycopeptide templates – which may easily be formed on solid phase and are less difficult to prepare in high purity – hold a lot of promise. Currently, the most versatile method is based on glycopeptide templates as mimics and utilizes in-situ capping in each synthesis step with a mixture of Boc and Fmoc amino acids, while glycosylated amino acids are separately encoded by capping with carboxylic acids. The analysis of structures is performed by direct photolytic release from PEG-based supports in the MALDI-TOF mass spectrometer. Application of this technique yields high-affinity ligands for carbohydrate-binding proteins.

## 8.14 Acknowledgements

Some of these studies have been supported by the Mitzutani foundation and the Danish National Research Foundation.



## 8.15 References

- [1] M. J. Sofia. Carbohydrate-based combinatorial libraries. *Mol. Diversity* **1998**, *3*, 75–94.
- [2] M. J. Sofia. Generation of oligosaccharide and glycoconjugate libraries for drug discovery. *DTT* **1996**, *1*, 27–34.
- [3] D. Kahne. Combinatorial approaches to carbohydrates. *Curr. Opin. Chem. Biol.* **1997**, *1*, 130–135.
- [4] P. Arya, R. N. Ben. Combinatorial chemistry for the synthesis of carbohydrate libraries. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1280–1282.
- [5] Y. Ding, O. Kanté, J. Labbe, M. M. Palcic, B. Ernst, O. Hindsgaul “Synthesis and biological activity of oligosaccharide libraries” in *Glycoimmunology*. (Eds: Alavi A, Axford JS), Plenum Press, New York, **1995**, pp. 261–269.
- [6] A. Varki. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **1993**, *3*, 97–130.
- [7] R. A. Dwek. Glycobiology: Toward understanding the function of sugars. *Chem. Rev.* **1996**, *96*, 683–720.
- [8] G. Kretzschmar, A. Toepfer, C. Hüls, M. Krause. Pitfalls in the synthesis and biological evaluation of sialyl- Lewis<sup>x</sup> mimetics as potential selectin antagonists. *Tetrahedron* **1997**, *53*, 2485–2494.
- [9] P. Y. Tong, W. Gregory, S. Kornfeld. Ligand interactions of the cation-independent mannose 6-phosphate receptor. *J. Biol. Chem.* **1989**, *264*, 7962–7969.
- [10] Y. C. Lee “Binding modes of mammalian hepatic Gal/GalNAc receptors” in *Carbohydrate recognition in cellular function*. (Eds: Bock G, Harnett S), John Wiley & Sons, Chichester, England, **1989**, pp. 80–95.
- [11] R. T. Lee, Y. C. Lee. Difference in binding-site architecture of the serum-type and liver-type mannose-binding proteins. *Glycoconj. J.* **1997**, *14*, 357–363.
- [12] K. Miyamura, K. M. B. Reid, U. Holmskov. The collectins – mammalian containing collagen-like regions. *TIGG* **1994**, *6*, 286–309.
- [13] W. I. Weis, K. Drickamer. Trimeric structure of a C-type mannose-binding protein. *Structure* **1994**, *2*, 1227–1240.
- [14] H. Franzyk, M. Meldal, H. Paulsen, S. Thiel, J. C. Jensenius, K. Bock. Glycopeptide mimics of mammalian Man<sub>6</sub>GlcNAc<sub>2</sub>. Ligand binding to mannan-binding proteins (MBPs). *Bioorg. Med. Chem.* **1996**, *4*, 1881–1899.
- [15] K. Bock. The preferred conformation of oligosaccharides in solution inferred from high resolution NMR data and hard sphere exo-anomeric calculations. *Pure Appl. Chem.* **1983**, *55*, 605–622.
- [16] K. Bock, J. Arnarp, J. Lönngren. The preferred conformation of oligosaccharides derived from the complex-type carbohydrate portion of glycoproteins. *Eur. J. Biochem.* **1982**, *129*, 171–178.
- [17] M. K. Christensen, M. Meldal, K. Bock, H. Cordes, S. Mouritsen, H. Elsner. Synthesis of glycosylated peptide templates containing 6'-O-phosphorylated mannose disaccharides and their binding to the mannose-6-phosphate receptor. *J. Chem. Soc., Perkin Trans. 1* **1994**, 1299–1310.
- [18] R. Schauer. Origin and biological role of the great chemical diversity of natural sialic Acids. *Trends in Glycoscience and Glycotechnology* **1997**, *9*, 315–330.
- [19] M. Vinson, S. Mucklow, A. P. May, E. Y. Jones, S. Kelm, P. R. Crocker. Sialic acid recognition by sialoadhesin and related lectins. *Trends in Glycoscience and Glycotechnology* **1997**, *9*, 283–297.
- [20] H. C. Hansen, S. Haataja, J. Finne, G. Magnusson. Di-, tri, and tetravalent dendritic galabiosides that inhibit hemagglutination by *streptococcus suis* at nanomolar concentration. *J. Am. Chem. Soc.* **1997**, *119*, 6974–6979.
- [21] L. Galli-Stampino, E. Meinjohanns, K. Frische, M. Meldal, T. Jensen, O. Werdelin, S. Mouritsen. The nature of the glycan moiety plays a decisive role in determining glycopeptide immunogenicity. *J. Cancer. Res.* **1997**, *57*, 3214–3222.
- [22] L. Poppe, G. S. Brown, J. S. Philo, P. V. Nikrad, B. H. Shah. Conformation of sLe<sup>x</sup> tetrasaccharide, free in solution and bound to E-, P-, and L-selectin. *J. Am. Chem. Soc.* **1997**, *119*, 1727–1736.
- [23] G. W. Hart. Dynamic O-linked glycosylation of nuclear and cytoskeletal proteins. *Annu. Rev. Biochem.* **1997**, *66*, 315–335.
- [24] G. W. Hart, K. D. Greis, D. L. Y. Dong, M. A. Blomberg, T.-Y. Chou, M.-S. Jiang, E. P. Roquemore, D. M. Snow, L. K. Kreppel, R. N. Cole, F. I. Comer, C. S. Arnold, B. K. Hayes “O-linked N-acetylglucosamine: the “Yin-Yang” of Scr/Thr phosphorylation?” in *Glycoimmunology*. (Eds: Alavi A, Axford JS), Plenum Press, New York, **1995**, pp. 115–123.

- [25] M. Meldal "Properties of the solid support" in *Methods in Enzymology: Solid-Phase Peptide Synthesis*. (Ed: Fields G), Academic Press, London, **1997**, pp. 83–104.
- [26] J. Vagner, G. Barany, K. S. Lam, V. Krchnák, N. F. Sepetov, J. A. Ostrem, P. Strop, M. Lebl. Enzyme-mediated spatial segregation on individual polymeric support beads: application to generation and screening of encoded combinatorial libraries. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8194–8199.
- [27] J. Buchardt, M. Meldal. A chemically inert hydrophilic resin for solid phase organic synthesis. *Tetrahedron Lett.* **1998**, *39*, 8695–8698.
- [28] M. Meldal. PEGA: A flow stable polyethylene glycol dimethyl acrylamide copolymer for solid phase synthesis. *Tetrahedron Lett.* **1992**, *33*, 3077–3080.
- [29] F.-I. Auzanneau, M. Meldal, K. Bock. Synthesis, characterization and biocompatibility of PEGA resins. *J. Peptide Sci.* **1995**, *1*, 31–44.
- [30] M. Renil, M. Meldal. POEPOP and POEPS: Inert polyethylene glycol crosslinked polymeric supports for solid phase synthesis. *Tetrahedron Lett.* **1996**, *37*, 6185–6188.
- [31] J. Rademann, M. Gröthli, M. Meldal, K. Bock. SPOCC: Resin for solid phase organic chemistry and enzyme reactions. *J. Am. Chem. Soc.* **1999**, *121*, 5459–5466.
- [32] M. Meldal, I. Svendsen, K. Breddam, F.-I. Auzanneau. Portion-mixing peptide libraries of quenched fluorogenic substrates for complete subsite mapping of endoprotease specificity. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3314–3318.
- [33] M. Meldal, I. Svendsen. Direct visualization of enzyme inhibitors using a portion mixing inhibitor library containing a quenched fluorogenic peptide substrate. 1: Inhibitors for subtilisin Carlsberg. *J. Chem. Soc., Perkin Trans. 1* **1995**, 1591–1596.
- [34] M. Meldal, F.-I. Auzanneau, O. Hindsgaul, M. M. Paicic. A PEGA resin for use in solid phase chemical/enzymatic synthesis of glycopeptides. *J. Chem. Soc., Chem. Commun.* **1994**, 1849–1850.
- [35] J. C. Spetzler, V. Westphal, J. R. Winther, M. Meldal. Preparation of fluorescence quenched libraries containing inter-chain disulfide bonds for studies of protein disulfide isomerases. *J. Peptide Sci.* **1998**, *4*, 128–137.
- [36] B. Yan. Monitoring the progress and the yield of solid-phase organic reactions directly on resin supports. *Acc. Chem. Res.* **1998**, *31*, 621–630.
- [37] S. Manabe, Y. Ito, T. Ogawa. Solvent effect in glycosylation reaction on polymer support. *Synlett* **1998**, 628–630.
- [38] H. Paulsen, A. Schleyer, N. Mathieux, M. Meldal, K. Bock. A new solid-phase oligosaccharide synthesis on glycopeptides bound to a solid phase. *J. Chem. Soc., Perkin Trans. 1* **1997**, 281–293.
- [39] R. Rodebaugh, S. Joshi, B. Fraser-Reid, H. M. Geysen. Polymer-supported oligosaccharides via *n*-pentenyl glycosides: Methodology for a carbohydrate library. *J. Org. Chem.* **1997**, *62*, 5660–5661.
- [40] F. Balkenhohl, C. v. d. Bussche-Hünnefeld, A. Lansky, C. Zechel. Combinatorial synthesis of small organic molecules. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2288–2337.
- [41] R. Liang, L. Yan, J. Loebach, M. Ge, Y. Uozumi, K. Sekanina, N. Horan, J. Gildersleeve, C. Thompson, A. Smith, K. Biswas, W. C. Still, D. Kahne. Parallel synthesis and screening of a solid phase carbohydrate library. *Science* **1996**, *274*, 1520–1522.
- [42] H. P. Nestler, P. A. Bartlett, W. C. Still. A general method for molecular tagging of encoded combinatorial chemistry libraries. *J. Org. Chem.* **1994**, *59*, 4723–4724.
- [43] B. J. Egner, S. Rana, H. Smith, N. Bouloc, J. G. Frey, W. S. Brocklesby, M. Bradley. Tagging in combinatorial chemistry: the use of coloured and fluorescent beads. *Chem. Commun.* **1997**, 735–736.
- [44] K. C. Nicolaou, X.-Y. Xiao, Z. Parandoosh, A. Senyei, M. P. Nova. Radiofrequency encoded combinatorial chemistry. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2289–2291.
- [45] X.-Y. Xiao, C. Zhao, H. Potash, M. P. Nova. Combinatorial chemistry with laser optical encoding. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 780–782.
- [46] Y.-H. Chu, Y. M. Dunayevskiy, D. P. Kirby, P. Vouros, B. L. Karger. Affinity capillary electrophoresis – Mass spectrometry for screening combinatorial libraries. *J. Am. Chem. Soc.* **1996**, *118*, 7827–7835.
- [47] B. E. Winger, J. E. Campana. Characterization of combinatorial peptide libraries by electrospray ionization fourier transform mass spectrometry. *Rapid Commun. Mass Spectr.* **1996**, *10*, 1811–1813.
- [48] J. W. Metzger, K.-H. Weismüller, S. Kienle, J. Brünjes, G. Jung "Mass spectrometric analysis of peptide libraries" in *Combinatorial Peptide and Nonpeptide Libraries*. (Ed: Jung G), VCH, Weinheim, **1999**, pp. 247–285.

- [49] S. Goletz, M. Leuck, P. Franke, U. Karsten. Structure analysis of acetylated and non-acetylated O-linked MUC1-glycopeptides by post-source decay matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectr.* **1997**, *11*, 1387–1398.
- [50] G. J. Rademaker, J. Haverkamp, J. Thomas-Oates. Determination of glycosylation sites in O-Linked Glycopeptides: A sensitive mass spectrometric protocol. *Org. Mass Spectrometry* **1993**, *28*, 1536–1541.
- [51] M. J. Kieliszewski, M. O'Neill, J. Leykam, R. Orlando. Tandem mass spectrometry and structural elucidation of glycopeptides from a hydroxyproline-rich plant cell wall glycoprotein indicate that contiguous hydroxyproline residues are the major sites of hydroxyproline O-arabinylation. *Jour. Biol. Chem.* **1995**, *270*, 2541–2549.
- [52] J. R. Chapman. Analysis of glycoproteins and glycopeptides using fast-atom bombardment. *Methods in Molecular Biology* **1997**, *61*, 161–253.
- [53] D. J. Harvey, T. J. P. Naven, B. Küster. Identification of oligosaccharides by matrix-assisted laser desorption ionization and electrospray MS. *Biochemical. Mass. Spectrom.* **1996**, *24*, 905–912.
- [54] T. Yamagaki, Y. Mitsuishi, H. Nakanishi. Post-source decay fragment analysis of a highly branched nonasaccharide of xyloglucan using MALDI-TOF mass spectrometry. *Chem. Lett.* **1998**, 57–58.
- [55] D. I. Papac, A. Wong, A. J. S. Jones. Analysis of acidic oligosaccharides and glycopeptides by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal. Chem.* **1996**, *68*, 3215–3223.
- [56] Y. Mechref, M. V. Novotny. Mass spectrometric mapping and sequencing of N-linked oligosaccharides derived from submicrogram amounts of glycoproteins. *Anal. Chem.* **1998**, *70*, 455–463.
- [57] T. Yamagaki, Y. Ishizuka, S.-I. Kawahara, H. Nakanishi. Analysis of glycosidic linkages in saccharide compounds by post-source decay fragment methods in matrix-assisted laser desorption/ionization time-of-flight mass spectroscopy. *Rap. Comm. Mass. Spec.* **1999**, *11*, 527–531.
- [58] Y. Zhao, S. B. H. Kent, B. T. Chait. Rapid, sensitive structure analysis of oligosaccharides. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 1629–1633.
- [59] D. Garozzo, V. Nasello, E. Spina, L. Sturiale. Discrimination of isomeric oligosaccharides and sequencing of unknowns by post source decay matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectr.* **1997**, *11*, 1561–1566.
- [60] M. Meldal, P. M. St. Hilaire, M. Willert, J. Rademann, M. Grötli, J. Buchardt, C. H. Gotfredsen, M. A. Juliano, L. Juliano. "Investigation of enzyme activity and inhibition in the interior of novel solid supports" in *Proc. 5<sup>th</sup> Chinese Pept. Symp., Chendu 1998*. Kluwer Academic Publishers, Dordrecht, **1999**, in press.
- [61] M. Grötli, C. H. Gotfredsen, J. Rademann, J. Buchardt, J. Ø. Duus, M. Meldal, K. Bock. Chemical and physical properties of PEG based resins. **1999**, submitted.
- [62] M. Pursch, G. Schlotterbeck, L.-H. Tseng, K. Albert, W. Rapp. Monitoring the reaction progress in combinatorial chemistry: H MAS NMR investigations on single macro beads in the suspended state. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2867–2869.
- [63] M. Baldus, R. J. Lulucci, B. H. Meier. Probing through-bond connectivities and through-space distances in solids by magic-angle-spinning nuclear magnetic resonance. *J. Am. Chem. Soc.* **1997**, *119*, 1121–1124.
- [64] I. E. Pop, C. F. Dhalluin, B. P. Deprez, P. C. Melnyk, G. M. Lippens, A. L. Tartar. Monitoring of a three-step solid phase synthesis involving a heck reaction using magic angle spinning NMR spectroscopy. *Tetrahedron* **1996**, *52*, 12209–12222.
- [65] R. S. Garigipati, B. Adams, J. L. Adams, S. K. Sarkar. Use of spin echo magic angle spinning <sup>1</sup>H NMR in reaction monitoring in combinatorial organic synthesis. *J. Org. Chem.* **1996**, *61*, 2911–2914.
- [66] P. H. Seeberger, X. Beebe, G. D. Sukanick, S. Pochapsky, S. J. Danishefsky. Monitoring the progress of solid-phase oligosaccharide synthesis by high-resolution magic angle spinning NMR: observations of enhanced selectivity for β-glycoside formation from α-1,2-anhydrosugar donors in solid-phase couplings. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 491–493.
- [67] C. Dhalluin, C. Boutillon, A. Tartar, G. Lippens. Magic angle spinning nuclear magnetic resonance in solid-phase peptide synthesis. *J. Am. Chem. Soc.* **1997**, *119*, 10494–10500.
- [68] C. H. Gotfredsen, M. Grötli, J. Ø. Duus, M. Meldal. Requirements for complete single bead structural elucidation of solid phase libraries by nano-probe MAS-NMR spectroscopy. **1999**, submitted.
- [69] M. Meldal. "Combinatorial solid phase assay for enzyme activity and inhibition" in *Combinatorial Peptide Libraries*. (Ed: Shmuel C), Humana Press, Totowa, New Jersey, **1998**, pp. 51–82.

- [70] J. Ø. Duus, P. M. St. Hilaire, M. Meldal, K. Bock. Carbohydrate chemistry synthetic and structural challenges towards the end of the 20th century. *Pure & Appl. Chem.* **1999**, in press.
- [71] L. Yan, C. M. Taylor, J. Goodnow, D. Kahne. Glycosylation in the Merrifield resin using anomeric sulfoxides. *J. Am. Chem. Soc.* **1994**, *116*, 6953–6954.
- [72] G. Grundler, R. R. Schmidt. Anwendung des Trichloracetimidat-Verfahrens auf 2-Azidoglucose- und 2-Azidogalactose-Derivate. *Liebigs Ann. Chem.* **1984**, 1826–1847.
- [73] C.-H. Wong, X.-S. Ye, Z. Zhang. Assembly of oligosaccharide libraries with a designed building block and an efficient orthogonal protection-deprotection strategy. *J. Am. Chem. Soc.* **1998**, *120*, 7137–7138.
- [74] T. Murata, T. Usui. Preparation of oligosaccharide units library and its utilization. *Biosci. Biotech. Biochem.* **1997**, *61*, 1059–1066.
- [75] C. Zheng, P. H. Seeberger, S. J. Danishefsky. Amidoglycosylation of polymer-bound glycals: a complete solid-phase synthesis of the oligosaccharide domain of the Lewis<sup>b</sup> blood group determinant. *Angew. Chem. Int. Ed.* **1998**, *37*, 786–789.
- [76] K. Fukase, K. Egusa, Y. Nakai, S. usumoto. Novel oxidatively removable protecting groups and linkers for solid-phase synthesis of oligosaccharides. *Mol. Diversity* **1996**, *2*, 182–188.
- [77] D. Weigelt, G. Magnusson. A linker for solid phase carbohydrate synthesis, permitting the introduction of variable anomeric functionality in the release step. *Tetrahedron Lett.* **1998**, *39*, 2839–2842.
- [78] O. Kaine, F. Barresi, Y. Ding, J. Labbe, A. Otter, L. S. Forsberg, B. Ernst, O. Hindsgaul. A strategy of “random glycosylation” for the production of oligosaccharide libraries. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2720–2721.
- [79] Y. Ding, J. Labbe, O. Kanie, O. Hindsgaul. Towards oligosaccharide libraries: a study of the random galactosylation of unprotected *N*-acetylglucosamine. *Bioorg. Med. Chem.* **1996**, *4*, 683–692.
- [80] G.-J. Boons, B. M. Heskamp, F. Hout. Vinyl glycosides in oligosaccharide synthesis: A strategy for the preparation of trisaccharide libraries based on latent-active glycosylation. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2845–2847.
- [81] M. Izumi, Y. Ichikawa. Combinatorial synthesis of oligosaccharide library of 2,6-dideoxysugars. *Tetrahedron Lett.* **1998**, *39*, 2079–2082.
- [82] T. Zhu, G.-J. Boons. A two-directional approach for the solid-phase synthesis of trisaccharide libraries. *Angew. Chem. Int. Ed.* **1998**, *37*, 1898–1900.
- [83] M. L. C. D. Silva, H. J. Stubbs, T. Tamura, K. G. Rice. <sup>1</sup>H NMR characterization of a hen ovalbumin tyrosinamide *N*-linked oligosaccharide library. *Arch. Biochem. Biophys.* **1995**, *318*, 465–475.
- [84] M. Meldal, I. Christiansen-Brams, M. K. Christensen, S. Mouritsen, K. Bock “Synthesis and biological application of glycosylated peptide templates” in *Complex carbohydrates in drug research. Structural and functional aspects*. (Eds: Bock K, Clausen H), Copenhagen, Munksgaard, **1994**, pp. 153–165.
- [85] H. Franzyk, M. K. Christensen, M. Jørgensen, M. Meldal, H. Cordes, S. Mouritsen, K. Bock. Constrained glycopeptide ligands for MPRs. Limitations of unprotected phosphorylated building blocks. *Bioorg. Med. Chem.* **1997**, *5*, 21–40.
- [86] C.-C. Lin, M. Shimazaki, M.-P. Heck, R. Wang, T. Kimura, H. Ritzen, S. Takayama, S.-H. Wu, G. Weitz-Schmidt, C.-H. Wong. Synthesis of sialyl Lewis X mimetics and related structures using the glycosyl phosphite methodology and evaluation of E-selectin inhibition. *J. Am. Chem. Soc.* **1996**, *118*, 6826–6840.
- [87] C.-C. Lin, T. Kimura, S.-H. Wu, G. Weitz-Schmidt, C.-H. Wong. Liposome-like fucopeptides as sialyl Lewis X mimetics. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2755–2760.
- [88] U. J. Nilsson, E. J. Fournier, O. Hindsgaul. Solid-phase extraction on C18 silica as a purification strategy in the solution synthesis of a 1-thio-beta-D-galactopyranoside library. *Bioorg. Med. Chem.* **1998**, *6*, 1563–1575.
- [89] M. Meldal, K. Bock. A general approach to the synthesis of *O*- and *N*-linked glycopeptides. *Glycoconjugate J.* **1994**, *11*, 59–63.
- [90] M. Meldal, E. Meinjohanns, K. Frische, T. Jensen, P. Hansen, O. Werdelin, L. Galli-Stampino, S. Mouritsen, K. Bock “Libraries of synthetic glycopeptides in the characterization of the T-cell response to tumor associated mucin antigens” in *Peptides, Biology and Chemistry*. (Eds: Xu X-J, Ye Y-H, Tam JP), Kluwer Academic Publishers, Dordrecht, **1998**, pp. 59–62.
- [91] K. Frische, M. Meldal, O. Werdelin, S. Mouritsen, T. Jensen, L. Galli-Stampino, K. Bock. Multiple column synthesis of a library of T-cell stimulating T<sub>H</sub>-antigenic glycopeptide analogs for the molecular characterization of glycan T-cell specificity. *J. Peptide Sci.* **1996**, *2*, 212–222.

- [92] S. Peters, T. L. Lowary, O. Hindsgaul, M. Meldal, K. Bock. Solid-phase synthesis of a fucosylated glycopeptide of human factor IX with a fucose- $\alpha$ -(1-*O*)-serine linkage. *J. Chem. Soc., Perkin Trans. I* **1995**, 3017–3022.
- [93] B. T. Chait, R. Wang, R. C. Beavis, S. B. H. Kent. Protein ladder sequencing. *Science* **1993**, 262, 89–92.
- [94] P. M. St. Hilaire, T. L. Lowary, M. Meldal, K. Bock. Oligosaccharide mimetics obtained by novel, rapid screening of carboxylic acid encoded glycopeptide libraries. *J. Am. Chem. Soc.* **1998**, 120, 13312–13320.
- [95] R. S. Youngquist, G. R. Fuentes, M. P. Lacey, T. Keough. Generation and screening of combinatorial peptide libraries designed for rapid sequencing by mass spectrometry. *J. Am. Chem. Soc.* **1995**, 117, 3900–3906.
- [96] P. M. St. Hilaire, T. Lowary, M. Meldal, K. Bock “Synthesis and analysis of glycopeptide libraries and their application in solid phase assays of carbohydrate binding-proteins” in *Peptides 1996, Proc. Eur. Pept. Symp.* (Eds: Epton R, Ramage R, Davies J), Mayflower Worldwide, Kingwinford, **1997**, pp. 817–818.
- [97] M. Meldal. Multiple column synthesis of quenched solid-phase bound fluorogenic substrates for characterization of endoprotease specificity. *Methods: A Companion to Methods in Enzymology* **1994**, 6, 417–424.
- [98] O. Seitz, H. Kunz. Hycron, an allylic anchor for high-efficiency solid-phase synthesis of protected peptides and glycopeptides. *J. Org. Chem.* **1997**, 62, 813–826.
- [99] A. Schleyer, M. Meldal, M. Renil, H. Paulsen, K. Bock. Direct solid phase glycosylations of peptide templates on a novel PE9-based resin. *Angew. Chem. Int. Ed. Engl.* **1997**, 36, 1976–1978.
- [100] J. F. G. Vliegenhart, J. Montreuil “Primary structure of glycoprotein glycans” in *Glycoproteins*. (Eds: Montreuil J, Vliegenhart JFG, Schachter H), Elsevier Science B. V., Amsterdam, **1998**, pp. 13–28.

# 9 RNA and DNA Aptamers

*Michael Famulok*

## 9.1 Introduction

Since the first examples of the use of combinatorial nucleic acid libraries for the in-vitro selection of specific ligand-binding RNAs were reported in 1990, considerable progress has been achieved in this field. Nucleic acid aptamers for more than 100 different targets have been described, showing that aptamers can now routinely be obtained for almost every desired target. Aptamer technology is currently in a state in which it demonstrates its power and impressive potential as a tool in molecular biology, diagnostics, molecular medicine, and bioorganic chemistry. In this review, an update of some novel and promising developments of in-vitro selection and aptamer technology is given.

An in-vitro selection experiment comprises various sequential steps, of which the first is the generation of a nucleic acid library of completely random sequences. The design of such libraries involves the synthesis of a short defined sequence, followed by a random region of variable length, and another defined sequence at the 5'-end. This pool of single-stranded DNA (ssDNA) is amplified in the polymerase chain reaction (PCR) to generate several copies of each DNA in its double-stranded form. To be able to transcribe RNA from these DNA templates, the 5'-primer contains a promotor sequence suitable for in-vitro transcription by RNA polymerase, usually the promotor for the RNA polymerase from phage T7 (T7-polymerase). By in-vitro transcription, a library of RNA molecules is generated which can then be used for the in-vitro selection. In this way, libraries of remarkable complexities can be generated which exceed the diversity of antibodies produced by the immune system by several orders of magnitude ( $10^{14}$ – $10^{15}$  different oligonucleotides).

The challenge is to select active molecules from oligonucleotide libraries of such complexity. As it is known that RNA molecules are able to fold into complex three-dimensional structures, the selection is performed on the assumption that this high pool complexity guarantees the presence of oligonucleotide-structures that are complementary to virtually any shape. Therefore, combinatorial oligonucleotide selections are often performed to isolate specifically binding nucleic acids, the so-called 'aptamers'. It is also possible to use the in-vitro selection technology for the isolation of catalytic nucleic acids, but in this review we will focus on binding rather than catalysis.

## 9.2 Aptamers for Small Molecules

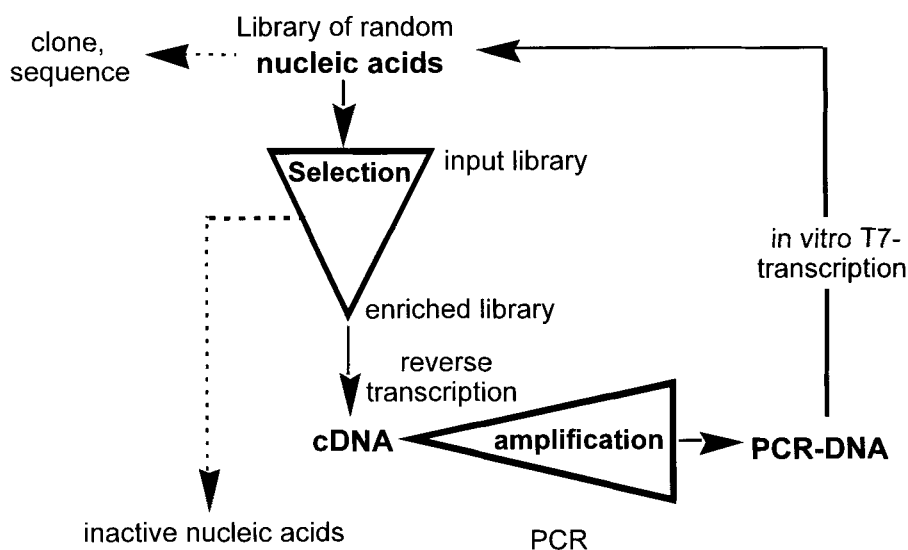
A considerable number of aptamer sequences which bind to molecules of low molecular weight with high specificity and affinity have been described (see Table 9.1).

**Table 1.** Aptamers selected to bind small organic molecules.

Ligand	NMR structure reference	PDB entry for structural data	Binding pocket	Reference
<b>Miscellaneous small organic ligands</b>				
Organic dyes			RNA	[1]
Organic dyes			DNA	[2]
Theophylline	[3]	1 EHT	RNA	[4]
Dopamine			RNA	[5]
Cellobiose			DNA	[6]
<b>Amino acids</b>				
D-tryptophan			RNA	[7]
L-arginine			RNA	[8]
L-arginine	[9]	1 koc	RNA	[10, 11]
L-citrulline	[9]	1 kod	RNA	[10, 11]
L-argininamide	[12]		DNA	[13]
L-arginine			RNA	[14]
L-arginine			mirror-image RNA	[15]
L-valine			RNA	[16]
L-Isoleucine			RNA	[17]
<b>Nucleotides and derivatives, biological cofactors</b>				
AMP	[18–20]	1 AMO, 1 RAW	RNA	[21–23]
AMP	[24]	1 AW4	DNA	[25]
AMP			mirror-image RNA	[26]
Guanosine			RNA	[27]
FMN	[28]	1 FMN	RNA	[22]
FMN			RNA	[29]
NAD			RNA	[29]
Vitamin B <sub>12</sub>			RNA	[30]
5'-Cap			RNA	[31]
Xanthene			RNA	[32]
<b>Antibiotics</b>				
Kanamycin				
Tobramycin	[33, 34]	1 TOB	RNA	[35]
Neomycin B			RNA	[36, 37]
Viomycin			RNA	[38]
Chloramphenicol			RNA	[39]
Streptomycin			RNA	[40]
<b>Others</b>				
Rev peptide vasopressin	[41]	1 ULL	RNA mirror-image RNA	[42, 43] [44]

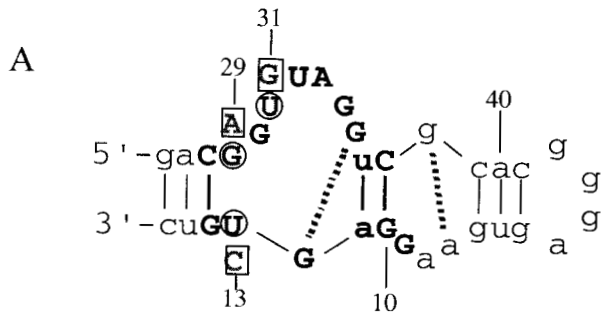
Aptamers for small molecules can be viewed as receptors composed of a structured nucleic acid rather than a protein. Sequences between 35 and 100 nucleotides in length are often sufficient for generating a specific binding pocket into which the target molecule can bind. In almost any case, small-molecule-binding aptamers have been selected by affinity chromatography. Ligands were immobilized on a solid support, incubated with the RNA library, non binding molecules were removed by extensive buffer-washes, the few remaining molecules bound to the immobilized ligand were eluted, the RNA reverse-transcribed, and the cDNA amplified by PCR. In-vitro transcription of the PCR-DNA resulted in an enriched RNA pool which was used as the input for the next selection cycle. Iterative cycles of selection and amplification yielded an RNA library in which specific ligand-binding sequences dominated the population. Cloning and sequencing revealed monoclonal aptamers for which their individual physico-chemical characteristics could be elucidated (Fig. 9.1).

Detailed structural investigations of many aptamer/ligand complexes by multi-dimensional NMR spectroscopy have established that the structure of aptameric binding sites for a given ligand usually is poorly pre-organized. RNA/ligand complex formation is always accompanied by extensive conformational changes and conformational stabilization of the binding RNA-receptor only in the presence of the cognate ligand [45–48]. In other words, while the RNA aptamer exists in a relatively undefined and floppy confor-

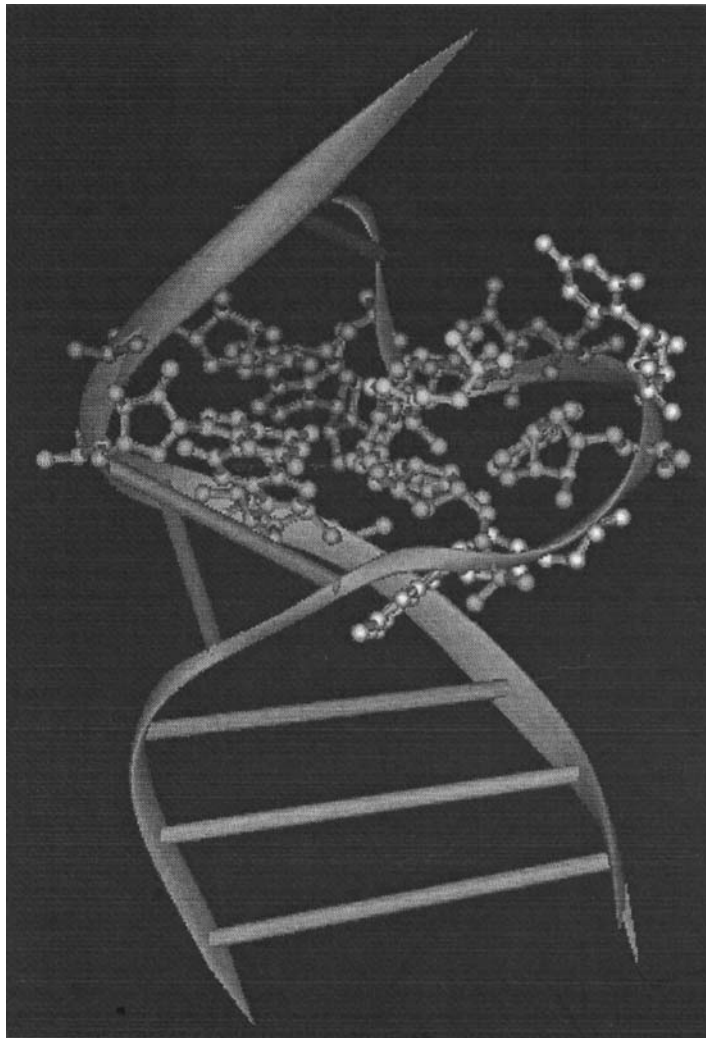


**Figure 9.1.** An in-vitro selection experiment comprises various sequential steps, of which the first is the generation of a nucleic acid library of completely random sequences. This library is subjected to an appropriate selection strategy which allows the separation of functional molecules from non-functional ones. The small proportion of nucleic acids with the desired activity is then amplified enzymatically and re-subjected to the selection procedure. This is necessary as the complexity of the library, which can contain up to  $10^{16}$  different oligonucleotide sequences, makes it impossible to enrich for the active sequences in one single selection and amplification cycle. Therefore, a number of cycles are performed sequentially until the functional sequences are the majority species in the library mix, and these can be characterized by cloning and sequencing.





B



mation in the absence of its ligand, defined and stable conformations are achieved by induced fit as soon as the ligand is added to the aptamer [9, 11, 18, 19, 28, 49].

According to current understanding, binding by an induced-fit mechanism seems to be a unique property of small-molecule-binding RNA aptamers. This property of aptamers might have inspired the idea of fusing aptamer sequences with known catalytic RNAs to introduce the principle of allosteric regulation into ribozyme catalysis. None of the ribozymes described so far was known to operate as allosteric enzymes *in vitro* or *in vivo*.

The hammerhead ribozyme (HHR) was transformed into an allosteric ribozyme by attaching the ATP- [50] or FMN-binding aptamer sequences [22, 28, 51] to its 5'-end [52]. These allosteric HHRs are capable of phosphodiester cleavage only in the absence of the cognate aptamer ligand. In the case of the conjoined aptamer/ribozyme construct, the presence of ATP or FMN results in a ligand-induced conformational change. In the case of the ATP-regulated HHR-construct, ligand binding causes a steric clash between aptamer and ribozyme domains [53] which prevents the HHR from adopting its active structure. The inhibition ratio obtained by this strategy was up to 180-fold. Tang and Breaker also used the principle of allosteric control of ribozyme catalysis to try to select variants of the HHR that would be more active in phosphodiester cleavage than the wild-type HHR by including the inhibitory ligand ATP during the *in-vitro* transcription [54]. In this selection, however, only the natural HHR sequence was isolated, suggesting that this sequence represents a motif which has been optimized by nature for the purpose of phosphodiester cleavage.

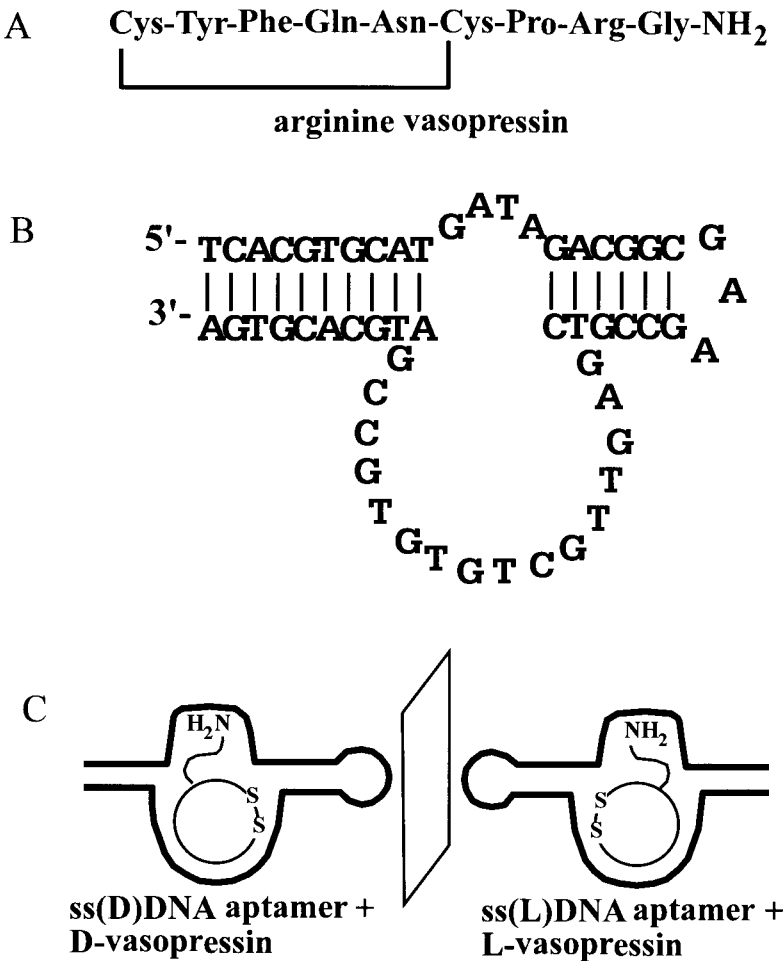
The molecular discrimination achieved by aptamer/small-molecule complexes can be as good as, or even better than with antibodies. The theophylline aptamer discriminates between the related molecules theophylline and caffeine, which differ by only one methyl group, at least ten-fold better than an antibody isolated for the same purpose [4]. An aptamer selected for the specific binding of L-arginine shows a 12000-fold reduced affinity to the D-arginine enantiomer [14]. An aptamer for neomycin B [36, 37] discriminates this antibiotic from a closely related one, paromomycin, in which an amino group is substituted by a hydroxy group. These remarkable specificities suggest that aptamers might soon find applications as very useful and powerful diagnostics. Small-molecule-binding aptamers may also be very useful tools in molecular biology. For example, it was shown by Haller and Sarnow that a 7-methyl guanosine-binding aptamer is able to inhibit the translation of capped mRNAs, but not of uncapped mRNAs, in cell-free lysates of HeLa or yeast cells [31].

Comparisons of various ligand-binding aptamer structures with proteins which bind related molecules showed that nucleic acids and proteins use strikingly similar strategies for the formation of well-defined binding pockets. Structures of nucleic acid/ligand complexes that have been published so far are summarized in Table 9.1.

◀ **Figure 9.2.** (A) Secondary structure proposed previously for the citrulline- and arginine-specific aptamers, based on co-variations of selected sequences, on the chemical footprinting pattern obtained in the presence of the cognate amino acid, as well as in damage selection experiments. The bases which were conserved among different isolates are shown in upper case; variant bases are in lower case. The three nucleotides critical for arginine specificity (13, 29 and 31) are indicated by circles (for citrulline) and boxes (for arginine). (B) Tertiary structure of the L-arginine aptamer complex resolved by NMR spectroscopy. Yellow: L-arginine; red: the three mutations. (Illustration adapted from [9].)

### 9.3 Functional Aptamers for Proteins and their Application in Biotechnology, Molecular Medicine and Diagnostics

The specificity of molecular recognition combined with the ease by which protein-binding aptamers can be isolated, engineered, evolved, and modified chemically makes these molecules very attractive as tools in molecular medicine, biotechnology and diagnostics. A very large number of aptamers which bind to all kinds of protein or peptide targets have been isolated. Because many of these studies have been extensively discussed



**Figure 9.3.** Mirror-image in-vitro selection approach. An aptamer which binds a stereoisomer of the cyclic peptide hormone arginine vasopressin (A) composed of all D-amino acids is isolated by in-vitro selection. The aptamer is synthesized in both enantiomeric forms for vasopressin-binding studies in-vitro (B). Inversion of all chiral centers in both partners yields a mirror-image complex, in which vasopressin is bound by a ligand which is resistant against degradation by cellular components (C).

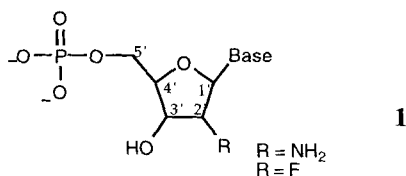
in several excellent and very detailed reviews [55–60], I will concentrate only on some examples or points.

One such point is the fact that RNA is very prone to hydrolysis by ubiquitous ribonucleases. This has led to the widely held bias that nucleic acids are potentially poor therapeutics, diagnostics, or assay-components. However, in contradiction to this opinion, there is now a battery of new results showing that functional nucleic acids not only can be made strikingly small, but also resistant against degradation in cellular environments [15, 26, 44, 59, 61]. One way to circumvent the vulnerability of RNA to nuclease degradation is indirect: in a first step, an aptamer that binds the enantiomer of the target is selected; then, in a second step, the enantiomer of the aptamer is synthesized (from L-phosphoramidites) as a nuclease-insensitive ligand of the natural target. This mirror-image approach, sometimes also designated as the ‘Spiegelmer’ approach, has been applied to L-arginine [15], D-adenosine [26], and the peptide hormone L-vasopressin [44].

Vasopressin is a peptide hormone involved in the regulation of water balance in the body [62]. This peptide plays a key role in various states of disease, including diabetes insipidus as well as hyponatremia and polydipsia in schizophrenic patients [63]. The goal of Williams et al. [44] was to isolate L-ssDNA ligands to vasopressin. In their approach, a library composed of natural D-DNA was applied to first select single-stranded DNA aptamers which specifically recognize D-vasopressin. The enantiomer of the winning D-ssDNA aptamer, designated as L-ssDNA aptamer, was then synthesized by classical phosphoramidite chemistry and its ability to bind L-vasopressin was demonstrated (Fig. 9.3).

Importantly, these mirror-image aptamers exhibited complete resistance against nuclease degradation, similar to the small-molecule-binding mirror-image aptamers previously selected by another group [15, 26]. When tested in a bioassay, the anti-vasopressin L-ssDNA-aptamer inhibited cAMP release mediated by vasopressin, but cAMP release induced by oxytocin (a related peptide hormone) was not affected.

An alternative to the mirror image-approach is the direct selection of an aptamer from libraries of chemically modified RNAs. Many modifications in the ribose moiety of nucleic acids have been shown to dramatically increase their nuclease resistance. Modifications have to be chosen so as to be compatible with nucleic acid replicating enzymes such as reverse transcriptase, or DNA- and RNA-polymerases. The modifications most commonly used are those in which the 2'-OH group of pyrimidines is substituted by a 2'-fluoro-, or a 2'-amino group (**1**) [64, 65].



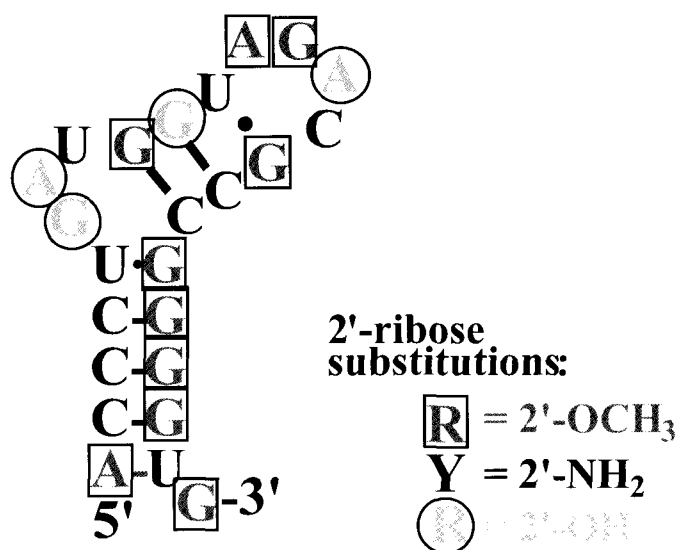
In-vitro selections with chemically modified libraries have been applied in various selection experiments and in many cases, the resulting modified aptamers exhibited a biological function. For example, an RNA library which was modified by substitution of the

2'-OH for a 2'-NH<sub>2</sub>-group at the riboses of pyrimidine residues to improve resistance against degradation by nucleases was used to select aptamers which bind to autoantibodies of patients affected by the muscular disease, myasthenia gravis. The aptamers blocked the binding of autoimmune antibodies to the human acetylcholine receptors from human cells [66]; as a potential therapeutic reagent for patients suffering from myasthenia gravis, these aptamers could be used to depress the undesirable antibody-stimulated immune response. In another case, 2'-fluoro-modified nuclease-resistant aptamers were selected which bind the human keratinocyte growth factor (hKGF) with a dissociation constant of up to 0.3 pM, this being the tightest-binding aptamer reported to date. The aptamers can competitively inhibit hKGF binding to its receptor and inhibit mitogenic activity with  $K_i$  values of 92 pM.

Another interesting target for several aptamer selections is the vascular permeability factor/vascular endothelial growth factor (VPF/VEGF). This protein plays a crucial role in the pathological induction of new blood vessel growth (angiogenesis) in a variety of epithelial hyperproliferative disorders. Aptamers from an unmodified RNA library of 30 randomized positions [67] as well as aptamers from a library with 2'-amino pyrimidine modifications [68] and 2'-fluoropyrimidine modifications [69] were used in these selection experiments. The aptamers selected from the modified libraries were minimized to sequences between 23 and 29 nucleotides and the 2'-OH groups of defined purine residues were subsequently modified by 2'-methoxy groups. This was achieved in a damage selection experiment in which variants of a defined aptamer were transcribed with a mixture of 2'-OH-purines and 2'-OCH<sub>3</sub>-purines, and subsequently screened for enhanced binding to VPF/VEGF [68, 69]. 2'-Methoxy-substituted purine residues were identified by protection from alkaline hydrolysis. In one 24-mer sequence from the 2'-amino-modified library, nine of the 13 purine positions could be substituted by the 2'-methoxy group; the other four could not be changed without significant loss of binding affinity. The final modified aptamer bound to the target protein with a  $K_d$  of 140 pM and specifically blocked the binding of <sup>125</sup>I-labeled VPF/VEGF to cell-surface receptors expressed on human umbilical vein endothelial cells [68]. Equilibrium dissociation constants for the truncated, 2'-O-methyl-modified aptamers from the 2'-fluoro-modified pool ranged between 49 and 130 pM for the interaction of VEGF (Fig. 9.4) [69].

One of the aptamers significantly reduced intradermal VEGF-induced vascular permeability *in vivo*. Thus, these nuclease-resistant molecules may be useful for the development of novel pharmaceutical lead compounds for epithelial hyperproliferative diseases.

Extracellular growth factors and other extracellular transmembrane proteins are particularly interesting targets for aptameric inhibitors. The reason is obvious: an aptamer, whether chemically modified or not, is a highly polar molecule due to the negatively charged phosphate backbone; thus, the molecule clearly faces a delivery problem when targeted to an intracellular protein. However, a number of studies have addressed the question of using aptamers expressed in the context of a living cell to induce a biological response. Induction of a biological response by an intracellularly expressed aptamer depends on a number of factors such as stabilization against rapid degradation, efficient transcription, direction into the correct cellular compartment in which the target molecule exhibits its activity, and its ability to recognize the target molecule when expressed in



**Figure 9.4.** Secondary structure of the 24-mer VPF/VEGF modified aptamer. Pyrimidines (Y) all contained 2'-fluoro-2'-deoxy modifications, purines (R) with 2'-methoxy-2'-deoxy modifications are boxed; purines (R) with unmodified ribose sugar residues are circled. The dissociation constant of the aptamer/protein complex was 190 pM.

the context of flanking RNA sequences of the RNA expression cassette used. I will discuss some recent examples of intracellular aptamers below.

## 9.4 Functional Aptamers In Vivo

Some RNA aptamers that were isolated *in vitro* have also been expressed *in vivo* to study their biological function within the context of a living pro- or eucaryotic cell. Among them is an aptamer which binds to the reverse transcriptase (Rev) protein of the human immunodeficiency virus type 1 (HIV-1) [42, 43, 70]. This anti-HIV-1-Rev aptamer was cloned into an expression cassette based on the U6 snRNA promoter, in which aptamer transcripts are protected against nuclease degradation to some extent. Transient expression in the nucleus of cultured cells led to  $10^7$ – $10^9$  full-length aptamer transcripts per cell. When anti-HIV-1-Rev aptamer-expressing cells were co-transfected with the HIV-1 provirus, viral replication was efficiently inhibited in these cells, as shown by an assay in which the production of HIV-1 reverse transcriptase was measured [71].

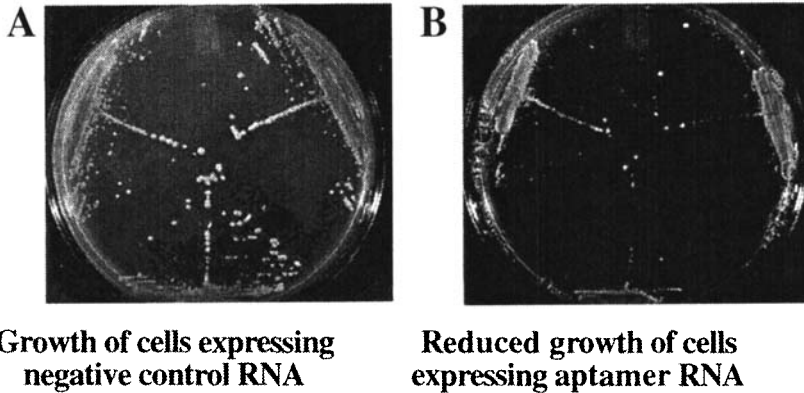
Another study also investigated the effect of various anti-Rev aptamers *in vivo*, but from a somewhat different angle. Symensma *et al.* substituted the wild-type Rev-binding element for the Rev-binding aptamer, and tested whether the 'synthetic' Rev-binding RNA would still be able to substitute the biological function of the wild-type, *i.e.*, to facilitate Rev-dependent mRNA transport [72]. Various classes of anti-HIV-1-Rev-binding aptamers were assayed: One class closely resembled the wild-type sequence, whereas the

other class diverged significantly in primary sequence, but not in its overall tertiary structure, as shown by molecular modeling [70]. The two groups of aptamers were both able to replace the function of the wild-type Rev-responsive element (RRE). When the aptamers were expressed in the presence of sub-saturating concentrations of Rev they facilitated mRNA transport with improved efficiencies compared to the wild-type RRE-sequence. This result was attributed to the fact that the aptamer constructs bound the Rev-protein between 3- and 10-fold better than the wild-type RRE *in-vitro*. This result suggested that the biological response of wild-type RNA sequences can be readily replaced by alternative sequences.

This conclusion, however, is far from general. In a study by Klug et al. [73], another natural RNA-sequence with a specialized biological function was replaced by *in-vitro*-selected aptamer sequences and the question was asked whether the substitution of a wild-type protein-binding RNA element for an *in-vitro*-selected unnatural one would result in differences in the biological effect of the RNA. Klug et al. isolated aptamers which bind the special elongation factor SelB of *E. coli*. The SelB protein promotes incorporation of the rare amino acid selenocysteine into formate dehydrogenase H (*fdhF*) at a UGA codon by simultaneous binding to selenocysteyl-tRNA<sup>Sec</sup>, GTP, and a downstream mRNA hairpin structure (the SECIS element) [74]. To investigate if binding of SelB to the SECIS element could be dissected from its overall biological function, the promotion of selenocysteine incorporation into proteins, the primary RNA sequence of the SECIS element was heavily mutagenized and the resulting RNA library used for the *in-vitro* selection of SelB-binding aptamers. The isolated SelB-binding aptamer variants closely resembled the wild-type SECIS element in its SelB-binding characteristics. When assayed *in vivo* for their ability to promote UGA readthrough in a lacZ fusion construct in *E. coli*, however, it was found that this was the case only for one very closely related sequence. The other three aptamers tested were drastically reduced or unable to replace the natural SECIS-mRNA element *in vivo*, despite their similar secondary structure and binding affinities compared to the wild-type motif. Based on these results it was assumed that, besides providing an RNA binding site for SelB, the biological role of the SECIS element also includes the promotion of a conformational change in the SelB protein which is necessary for selenocysteine incorporation *in vivo*.

As stated by Conrad et al. [75], the two studies discussed above suggest that in some instances wild-type sequences are functionally optimal, whereas in other cases wild-type sequences can be readily replaced by alternative sequences. Some RNA-binding sequences, such as the Rev-binding element, obviously are able to function in a structure-independent manner, whereas other RNA-binding sequences, such as the SelB-binding mRNA-SECIS element, function in a structure-dependent manner.

Thomas et al. also investigated the effect of RNA aptamers selected *in vitro* for recognition of a natural nucleic acid-binding protein, the RNA polymerase II from *Saccharomyces cerevisiae* (Pol II), when expressed in an intracellular context [76]. Pol II-binding RNA aptamers were selected from a completely randomized library, and first investigated *in vitro* for their ability to inhibit RNA polymerization. One selected RNA ligand specifically inhibited Pol II from *S. cerevisiae*, whereas Pol I from the same organism and Pol III and Pol II from *Schizosaccharomyces pombe* or wheat germ cells were not affected. *In vivo* expression of the selected RNA in yeast cells under the control of a Pol III pro-



**Figure 9.5.** Growth of yeast strain YF1971 cells expressing Pol II-specific aptamer RNA (B) compared to non-binding RNAs from the unselected pool (A). When these cells were grown on a leucine-rich medium, the endogenous amount of Pol II is reduced. Aptamer-expressing cells clearly showed reduced growth under conditions of reduced Pol II expression.

motor showed that yeast cells that expressed the anti-Pol II aptamer grew normally. However, when the RNA aptamer was expressed in a yeast strain with an artificially reduced level of Pol II, presumably at sub-saturating concentrations of Pol II, an aptamer-specific cell growth defect was observed (Fig. 9.5).

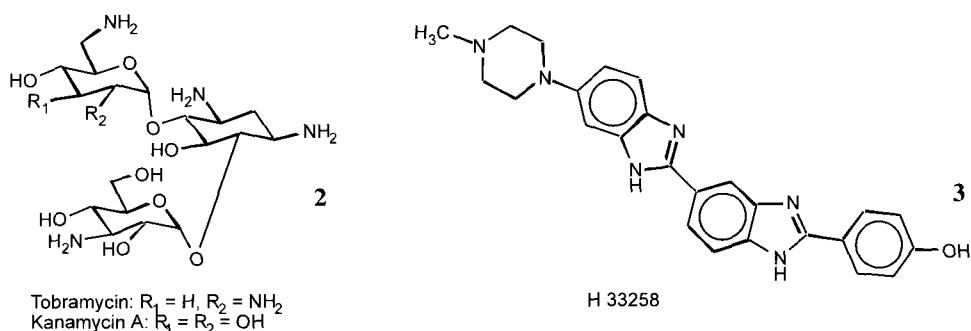
So far, *in-vivo* expression of aptamers was restricted to non natural RNA-aptamers which recognized proteins with a natural affinity to nucleic acids. Furthermore, eucaryotic expression *in vivo* was restricted to the nucleus only. Many macromolecular interactions, however, occur in the cytoplasm and are implicated in unassessable biological responses, the precise function of which in many cases has remained largely elusive. Thus, to understand or control the intracellular functions of specific proteins, substantial efforts are directed at the development of general strategies for the generation and application of specific inhibitors of macromolecular interactions. Therefore, a major biological and medical challenge is to identify such selective tools, and one such tool might be the aptamer-technology.

Blind et al. have recently undertaken an approach in this direction [77]. They selected RNA aptamers directed against the cytoplasmic domain of the  $\beta 2$  integrin CD18, a heterodimeric integral transmembrane protein found exclusively on leukocytes. This integrin is involved in various regulatory steps in immunology, and mediates adhesion of leukocytes in immune and inflammatory responses by binding to cellular ligands, the so-called intercellular adhesion molecules (ICAM-1, -2, and -3). Blind et al. constructed a cytoplasmic aptamer-expression system based on viral vectors which allowed for the inhibition of the  $\beta 2$  integrin-mediated cell adhesion to ICAM-1 and investigations of its participation in signal transduction and cell adhesion *in vivo*. This study established that cytoplasmic aptamers are capable not only of targeting proteins without known affinity to nucleic acids, but also of protein-receptors which are anchored in the plasma membrane compartment. This opens an even wider application potential for aptamer technology, which add to the highly significant aforementioned studies on aptamer expression in nuclei. Following the



most recently established developments in the field of intracellular antibodies ('Intrabodies') [78, 79], it can be predicted that cytoplasmically expressed aptamer-reagents ('Intramers') will play a major role as molecular biological tools for investigating cellular processes, and as novel reagents for cancer therapy and the control of infectious diseases. The significance of this new method lies in the readily accessible range of highly specific reagents designed to influence cell processes directly.

Werstruck and Green recently showed that aptamers binding to small organic molecules can, potentially, be used to control the expression of genes. They showed that aptamers selected to bind the aminoglycosides tobramycin and kanamycin A (**2**) or to certain dye molecules (Hoechst H33258; **3**) *in vitro* were also capable of binding these ligands *in vivo*.



The complementary DNA of these aptamers was inserted in tandem into the untranslated region of an expression plasmid with a  $\beta$ -galactosidase reporter gene. Chinese hamster ovary (CHO) cells were transfected with the expression plasmid and grown in the presence of H33258. Quantification of  $\beta$ -galactosidase activities 24 hours after transfection revealed that translation was prevented at concentrations of 5–10 mM of H33258 [80].

## 9.5 Conclusions

The few examples depicted in this review demonstrate that aptamers possess properties comparable to those of protein monoclonal antibodies, and thus should be potent alternatives to long-established, antibody-based biopharmaceutical or diagnostic products. They can be equipped in a modular manner with additional functions, and therefore tailored specifically for many potential applications in biotechnology, molecular medicine and molecular biology. With affinities in the nanomolar to picomolar range, aptamers can bind their target molecules with a specificity similar to the antigen-binding fragment of monoclonal antibodies. As RNA molecules, aptamers can be simply synthesized by the cell's own transcription machinery. Due to their high affinities, they represent excellent candidates for highly specific inhibitors of signal transduction, cell growth, transcription,

or viral replication. While nucleic acids in cells and in the blood degrade rapidly, chemically modified aptamers or aptamers expressed in a certain structural context show dramatically increased stability profiles, and are often resistant to nuclease digestion. With the development of appropriate gene transfer systems, the use of intracellular aptamers, analogous to intrabodies, can also be considered in gene therapy procedures to combat infectious diseases, and in cancer therapy. An advantage of the nucleic acid-based system would be its greater safety. Being an RNA molecule, the provocation of cytotoxic immune reactions induced by aptamers should not occur, as the normal protein presentation pathway via T cells does not apply.

## 9.6 Acknowledgements

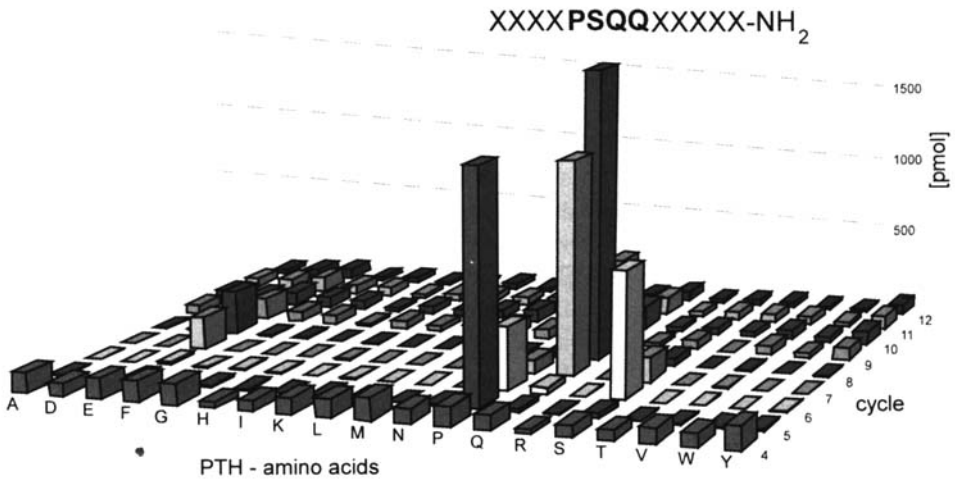
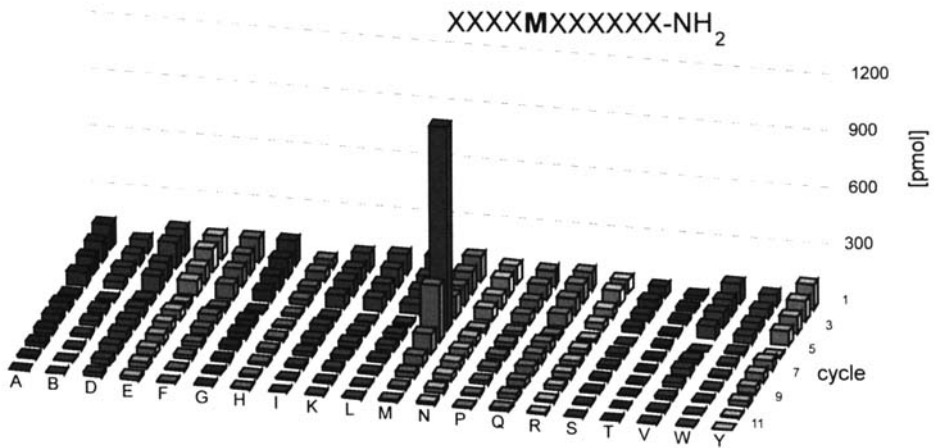
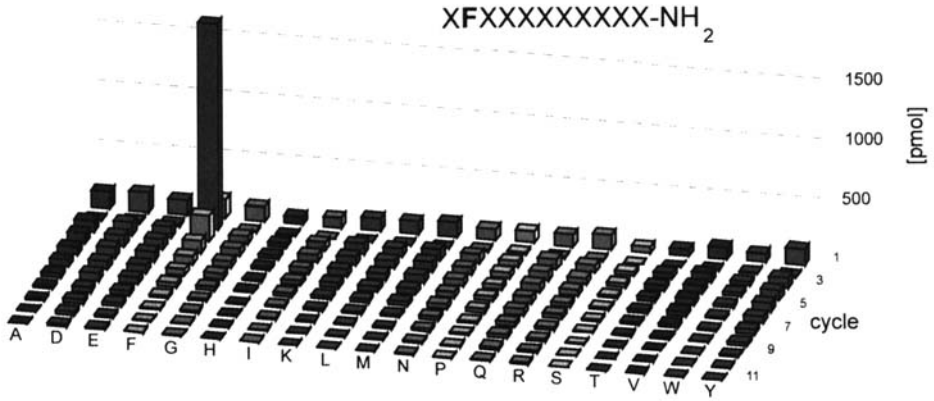
I thank all the members of my research group for many helpful and stimulating discussions and Günter Mayer for critical reading of the manuscript.

## 9.7 References

- [1] Ellington, A. D. & Szostak, J. W. (1990) *Nature* **346**, 818–822.
- [2] Ellington, A. D. & Szostak, J. W. (1992) *Nature* **355**, 850–852.
- [3] Zimmerman, G. R., Jenison, R. D., Wick, C. L., Simorre, J.-P. & Hardy, A. (1997) *Nature Struct. Biol.* **4**, 644–649.
- [4] Jenison, R. D., Gill, S. C., Pardi, A. & Polisky, B. (1994) *Science* **263**, 1425–1429.
- [5] Mannironi, C., Di Nardo, A., Fruscoloni, P. & Tocchini-Valentini, G. P. (1997) *Biochemistry* **36**, 9726–9734.
- [6] Yang, Q., Goldstein, I. J., Mei, H. Y. & Engelke, D. R. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 5462–5467.
- [7] Famulok, M. & Szostak, J. W. (1992) *J. Am. Chem. Soc.* **114**, 3990–3991.
- [8] Connell, G. J., Illangsekare, M. & Yarus, M. (1993) *Biochemistry* **32**, 5497–5502.
- [9] Yang, Y., Kochoyan, M., Burgstaller, P., Westhof, E. & Famulok, M. (1996) *Science* **272**, 1343–1347.
- [10] Famulok, M. (1994) *J. Am. Chem. Soc.* **116**, 1698–1706.
- [11] Burgstaller, P., Kochoyan, M. & Famulok, M. (1995) *Nucleic Acids Res.* **23**, 4769–4776.
- [12] Lin, C. H. & Patel, D. J. (1996) *Nat. Struct. Biol.* **3**, 1046–1050.
- [13] Harada, K. & Frankel, A. D. (1995) *EMBO J.* **14**, 5798–5811.
- [14] Geiger, A., Burgstaller, P., Von der Eltz, H., Roeder, A. & Famulok, M. (1996) *Nucleic Acids Res.* **24**, 1029–1036.
- [15] Nolte, A., Klußmann, S., Bald, R., Erdmann, V. A. & Fürste, J. P. (1996) *Nat. Biotechnol.* **14**, 1116–1119.
- [16] Majerfeld, I. & Yarus, M. (1994) *Nat. Struct. Biol.* **1**, 287–292.
- [17] Majerfeld, I. & Yarus, M. (1998) *RNA* **4**, 471–478.
- [18] Dieckmann, T., Suzuki, E., Nakamura, G. K. & Feigon, J. (1996) *RNA* **2**, 628–640.
- [19] Jiang, F., Kumar, R. A., Jones, R. A. & Patel, D. J. (1996) *Nature* **382**, 183–186.
- [20] Nonin, S., Jiang, F. & Patel, D. J. (1997) *J. Mol. Biol.* **268**, 359–374.
- [21] Sassanfar, M. & Szostak, J. W. (1993) *Nature* **364**, 550–553.
- [22] Burgstaller, P. & Famulok, M. (1994) *Angew. Chem. Int. Ed.* **33**, 1084–1087.
- [23] Burke, D. H. & Gold, L. (1997) *Nucleic Acids Res.* **25**, 2020–2024.
- [24] Lin, C. H. & Patel, D. J. (1997) *Chem. Biol.* **4**, 817–832.
- [25] Huizenga, D. E. & Szostak, J. W. (1995) *Biochemistry* **34**, 656–665.
- [26] Klußmann, S., Nolte, A., Bald, R., Erdmann, V. A. & Fürste, J. P. (1996) *Nat. Biotechnol.* **14**, 1112–1115.

- [27] Connell, G. J. & Yarus, M. (1994) *Science* **264**, 1137–1141.
- [28] Fan, P., Suri, A. K., Fiala, R., Live, D. & Patel, D. J. (1996) *J. Mol. Biol.* **258**, 480–500.
- [29] Lauhon, C. T. & Szostak, J. W. (1995) *J. Am. Chem. Soc.* **117**, 1246–1257.
- [30] Lorsch, J. R. & Szostak, J. W. (1994) *Biochemistry* **33**, 973–982.
- [31] Haller, A. A. & Sarnow, P. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 8521–8526.
- [32] Kiga, D., Futamura, Y., Sakamoto, K. & Yokoyama, S. (1998) *Nucleic Acids Res.* **26**, 1755–1760.
- [33] Jiang, L., Suri, A. K., Fiala, R. & Patel, D. J. (1997) *Chem. Biol.* **4**, 35–50.
- [34] Jiang, L. & Patel, D. J. (1998) *Nat. Struct. Biol.* **5**, 769–774.
- [35] Wang, Y., Killian, J., Hamasaki, K. & Rando, R. R. (1996) *Biochemistry* **35**, 12338–12346.
- [36] Wallis, M. G., Von Ahsen, U., Schroeder, R. & Famulok, M. (1995) *Chem. Biol.* **2**, 543–552.
- [37] Famulok, M. & Hüttenhofer, A. (1996) *Biochemistry* **35**, 4265–4270.
- [38] Wallis, M. G., Streicher, B., Wank, H., von Ahsen, U., Clodi, E., Wallace, S. T., Famulok, M. & Schroeder, R. (1997) *Chem. Biol.* **4**, 357–366.
- [39] Burke, D. H., Hoffman, D. C., Brown, A., Hansen, M., Pardi, A. & Gold, L. (1997) *Chem. Biol.* **4**, 833–843.
- [40] Wallace, S. T. & Schroeder, R. (1998) *RNA* **4**, 112–123.
- [41] Ye, X., Gorin, A., Ellington, A. D. & Patel, D. (1996) *Nat. Struct. Biol.* **3**, 1026–1033.
- [42] Giver, L., Bartel, D. P., Zapp, M. L., Green, M. R. & Ellington, A. D. (1993) *Nucleic Acids Res.* **21**, 5509–5516.
- [43] Giver, L., Bartel, D. P., Zapp, M. L., Green, M. R. & Ellington, A. D. (1993) *Gene* **137**, 19–24.
- [44] Williams, K. P., Liu, X. H., Schumacher, T. N., Lin, H. Y., Ausiello, D. A., Kim, P. S. & Bartel, D. P. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 11285–11290.
- [45] Cech, T. R. & Szewczak, A. A. (1996) *RNA* **2**, 625–627.
- [46] Patel, D. J. (1997) *Curr. Opin. Chem. Biol.* **1**, 32–46.
- [47] Patel, D. J., Suri, A. K., Jiang, F., Liang, L., Fan, P., Kumar, R. A. & Nonin, S. (1997) *J. Mol. Biol.* **272**, 645–664.
- [48] Heus, H. A. (1997) *Nat. Struct. Biol.* **4**, 597–600.
- [49] Feigon, J., Dieckmann, T. & Smith, F. W. (1996) *Chem. Biol.* **3**, 611–617.
- [50] Tang, J. & Breaker, R. R. (1997) *Chem. Biol.* **4**, 453–459.
- [51] Burgstaller, P. & Famulok, M. (1996) *Bioorg. Med. Chem. Lett.* **6**, 1157–1162.
- [52] Araki, M., Okuno, Y., Hara, Y. & Sugiura, Y. (1998) *Nucleic Acids Res.* **26**, 3379–3384.
- [53] Tang, J. & Breaker, R. R. (1998) *Nucleic Acids Res.* **26**, 4214–4221.
- [54] Tang, J. & Breaker, R. R. (1997) *RNA* **3**, 914–925.
- [55] Famulok, M. & Jenne, A. (1998) *Curr. Opin. Chem. Biol.* **2**, 320–327.
- [56] Famulok, M. (1998) in *Combinatorial nucleic acid libraries: the new world of aptamers and ribozymes*, ed. Habenschütz, U. (Springer Verlag, Heidelberg), pp. in press.
- [57] Gold, L., Polisky, B., Uhlenbeck, O. & Yarus, M. (1995) *Annu. Rev. Biochem.* **64**, 763–797.
- [58] Gold, L. (1995) *Harvey Lect.* **91**, 47–57.
- [59] Osborne, S. E. & Ellington, A. D. (1997) *Chem. Rev.* **97**, 349–370.
- [60] Famulok, M. & Mayer, G. (1999) *Curr. Topics Microbiol. Immunol.*, in press.
- [61] Eaton, B. E. (1997) *Curr. Opin. Chem. Biol.* **1**, 10–16.
- [62] Nielsen, S., Chou, C. L., Marples, D., Christensen, E. I., Kishore, B. K. & Knepper, M. A. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 1013–1017.
- [63] Goldman, M. B., Robertson, G. L., Luchins, D. J., Hedecker, D. & Pandey, G. N. (1997) *Arch. Gen. Psychiatry* **54**, 443–449.
- [64] Aurup, H., Tuschl, T., Benseler, F., Ludwig, J. & Eckstein, F. (1994) *Nucleic Acids Res.* **22**, 20–24.
- [65] Pieken, W. A., Olsen, D. B., Benseler, F., Aurup, H. & Eckstein, F. (1991) *Science* **253**, 314–317.
- [66] Lee, S.-W. & Sullenger, B. A. (1997) *Nature Biotechnology* **15**, 41–45.
- [67] Jellinek, D., Green, L. S., Bell, C. & Janjic, N. (1994) *Biochemistry* **33**, 10450–10456.
- [68] Green, L. S., Jellinek, D., Bell, C., Beebe, L. A., Feistner, B. D., Gill, S. C., Jucker, F. M. & Janjic, N. (1995) *Chem. Biol.* **2**, 683–695.
- [69] Ruckman, J., Green, L. S., Beeson, J., Waugh, S., Gillette, W. L., Henninger, D. D., Claesson, W. L. & Janjic, N. (1998) *J. Biol. Chem.* **273**, 20556–20567.
- [70] Leclerc, F., Cedergren, R. & Ellington, A. D. (1994) *Nat. Struct. Biol.* **1**, 293–300.
- [71] Good, P. D., Krikos, A. J., Li, S. X., Bertrand, E., Lee, N. S., Giver, L., Ellington, A., Zaia, J. A., Rossi, J. J. & Engelke, D. R. (1997) *Gene Ther.* **4**, 45–54.
- [72] Symensma, T. L., Giver, L., Zapp, M., Takle, G. B. & Ellington, A. D. (1996) *J. Virol.* **70**, 179–187.

- [73] Klug, S. J., Hüttenhofer, A., Kromayer, M. & Famulok, M. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 6676–6681.
- [74] Zinoni, F., Heider, J. & Böck, A. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4660–4664.
- [75] Conrad, R. C., Symensma, T. L. & Ellington, A. D. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 7126–7128.
- [76] Thomas, M., Chédin, S., Carles, C., Riva, M., Famulok, M. & Sentenac, A. (1997) *J. Biol. Chem.* **272**, 27980–27986.
- [77] Blind, M., Kolanus, W. & Famulok, M. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 3606–3610.
- [78] Marasco, W. A. (1995) *Immunotechnology* **1**, 1–19.
- [79] Marasco, W. A. (1997) *Gene Ther.* **4**, 11–15.
- [80] Werstuck, G. & Green, M. R. (1998) *Science* **282**, 296–298.



# 10 Combinatorial Approaches to Molecular Receptors for Chemosensors

*Dietmar Leipert, Jürgen Mack, Rolf Tünnemann and Günther Jung*

## 10.1 Introduction

Conventional analytical methods barely fulfil present requirements for fast process control and environmental analysis. Quantification of the components often requires their separation by precipitation, centrifugation, TLC, GC, HPLC, CE and other methods. The analytical procedures are rather time-consuming, especially when chromatographic steps are included. However, as many analytical problems demand a rapid and continuous acquisition of analytical data, the ongoing research for appropriate sensors is of growing importance. Aspects of recognition and sensor applications have been described in a selection of recent books and reviews [1–10].

The research for chemosensors began as a branch of analytical chemistry and is now an approved and independent field of activities at the interface between research and application. A chemosensor can be considered as a small unit for the acquisition of analytical data. It has been optimized for one distinct application includes a sensitive layer, whose physico-chemical properties are affected by the interaction with the substance to be detected. These effects are translated into electronic signals by microelectronic devices and can be processed by data acquisition systems [11]. In most cases, mass-sensitive or optical transducers are used, and some of them are listed in Table 10.1.

**Table 10.1.** Principles of transducers for the detection of interactions between analyte and receptor.

Transducer principle	Explanation	Reference
Impedance spectroscopy	Change of impedance (frequency and potential dependent) of a thin layer	[15]
Ellipsometry	Change of polarization-dependence of reflectivity	[16]
White light reflectometry	Change of interference pattern of a thin layer	[17, 18]
Microbalance	Change of frequency induced by mass changes	[19, 20]
Surface plasmon spectroscopy	Change of angle of resonance by change of reflective index	[21, 22]
Resonant mirror	Influence of a guided optical wave on phase relation by evanescent field	[23]
Grating coupler	Shift of coupling angle of an optical grating	[24, 25]

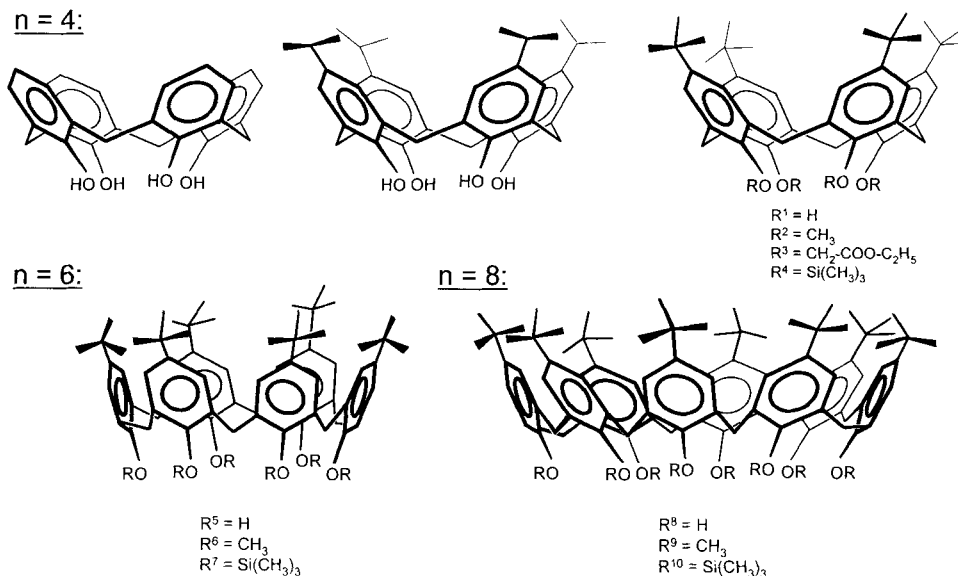
The analyte can hardly be identified if the sensor substrate induces only weak and un-specific interactions. A new approach to overcome these problems is the development, using supramolecular chemistry, of sensor materials which allow the production of selective chemosensors. A large number of supramolecular receptors can be prepared by means of combinatorial chemistry [12], but the selective receptors described so far, such as cyclodextrins and calixarenes, offer only a relatively small number of distinct variations.

Thus, synthetic combinatorial receptor libraries of high diversity are a prerequisite for further progress in the field of chemosensors. As yet, there are no specific receptors available for most of the analytes, but this can be overcome by the use of sensor arrays. A mathematical analysis by pattern recognition can lead to a successful analysis of compound mixtures which are exposed to an array of sensors with different selectivities ('bio-electronic nose') [13, 14].

## 10.2 Supramolecular Recognition Sites

Receptors prepared by organic synthesis which are used as sensor coating with defined molecular recognition properties, may increase the sensitivity of a sensor in the gaseous and in the liquid phases. Such sensors can be considered as quick and simple assay systems to detect supramolecular interactions between the synthesized receptors and analytes at the solid/liquid or solid/gaseous interfaces (Fig. 10.1) [26].

Those classes of substances which form cavities have been found to be the most specific hosts. These cavities can be entered only by guests with appropriate morphology, and



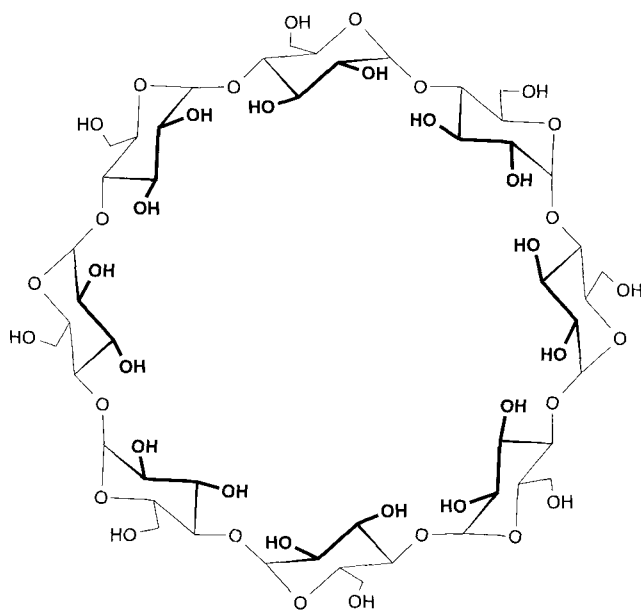
**Figure 10.1.** Different structures of calixarenes used as recognition sites in the gaseous phase.

are formed by cyclodextrins, cyclophanes, calixarenes, resorcinarenes and crown ethers. Their ability to include host molecules has been demonstrated [27, 28].

Calix[n]arenes are cyclic condensation products of *para*-substituted phenol derivatives and formaldehyde [29]. They are highly interesting for the development of sensitive coatings due to their conformational flexibility and the ease by which they may be modified chemically. Chemical modification can be done either in the *meta* position, or by reactions at the hydroxy group. In this way, bulky substituents [30], chelating substituents [31], aromatic residues [32], crown ethers [33, 34], peptides [35, 36], etc. can be introduced. A first approach to combinatorial synthesis of calix[4]arene receptors has been published by Reinhoudt and co-workers [37, 38], who prepared calixarenes with different substituents. In solution, these calixarenes lead to formation of hetero-oligomers with barbiturates, and these hetero-oligomers were detected by MALDI-TOF mass spectrometry and  $^1\text{H-NMR}$  spectroscopy.

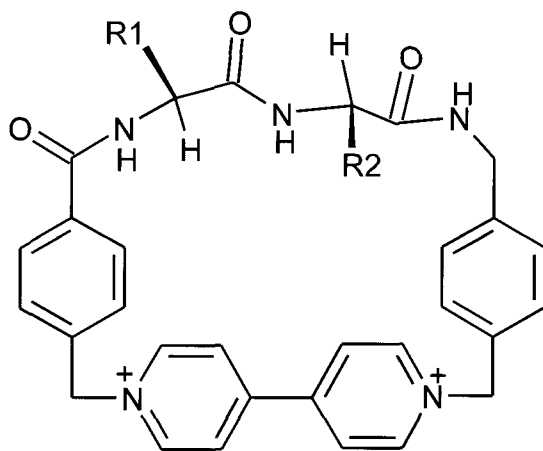
Cyclodextrins are ring-shaped condensation products consisting of six, seven or eight ( $\alpha$ -,  $\beta$ -,  $\gamma$ -) glucopyranose units (Fig. 10.2) [39]. The dimensions of the cavity and the flexibility of the structure depend on the ring size. Cyclodextrins are used in industry for separation of hydrocarbons [40], as scavengers for highly diluted molecules [41] and (appropriately modified) as chiral separation phase in chromatography [42–44]. These applications are due to the simple preparation and negligible toxicity of these compounds, and their ability to form complexes [45]. Many modifications of the cyclodextrin molecules have been described [44], though no combinatorial approaches were reported.

Cyclophanes are cyclic macromolecules with aromatic units connected via short aliphatic chains. The shape and the properties of their cavities can be varied by modification



**Figure 10.2.** Structure of a  $\gamma$ -cyclodextrin.



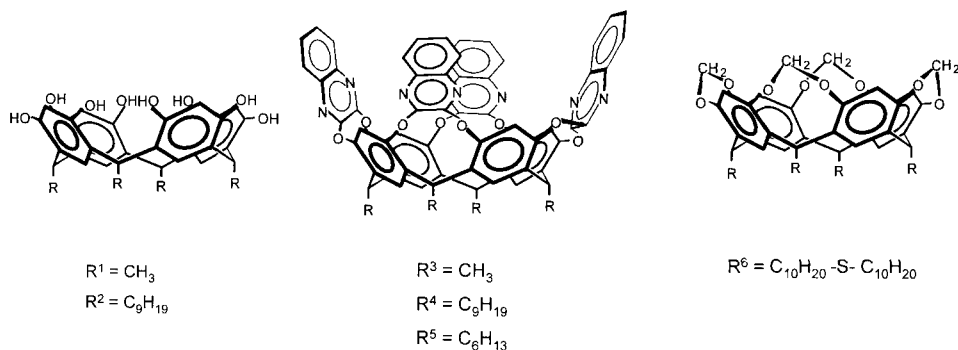


**Figure 10.3.** Cyclophane receptor consisting of a 4,4'-bipyridinium- and a dipeptide fragment [46].

of the aromatic units and of the connecting chains [47]. Garcia et al. [48] performed the parallel synthesis of cyclophanes consisting of a bipyridinium dication and 20 L-dipeptides (Fig. 10.3). The chiral recognition of amino acid derivatives could be verified by NMR titration.

Pedersen [49] discovered the crown ethers and their complexation behavior. The first crown ethers were by-products and have been identified as a new class of substances because of their extraordinary complexation behavior [50]. The synthesis typically starts from brenzcatechin, and one or two aromatic rings and five to ten ether units are included in the macrocyclic.

Resorcinarenes are synthesized by condensation of resorcinol and aldehydes. The structures of calix- and resorcinarenes are similar, the compounds differing only with respect to the number and position of substituents [51]. Since formaldehyde can be replaced by other aldehydes, the bridging methylene groups can be modified in the condensation reaction (Fig. 10.4) [52]. This approach may be suitable for the parallel synthesis of more diverse recognition sites.



**Figure 10.4.** Different structures of resorcinarenes used as recognition sites in the gaseous phase.

Phthalocyanines have been developed as new sensor materials [53]. Like the crown ethers, phthalocyanines have been discovered as by-products of a synthesis, namely the reaction of *o*-cyanobenzamide with phthalodinitrile [54]. They consist of four indole units, and most of them have square-planar structure; a metal ion can be co-ordinated in the center of the square. According to this structure, molecular interaction with an analyte is not an effect of a cavity.

### 10.3 Macrocyclic Peptides

To date, the capability of macromolecular peptides as molecular recognition sites has been poorly characterized. Due to the analytical techniques available, first investigations on host-guest complexes were done in the homogeneous phase [55]. One of the most popular cyclic compounds is the depsipeptide, valinomycin, which can complex potassium ions selectively and carry them through lipid membranes [56]. Another interesting receptor is the peptide antibiotic, vancomycin [57]. This consists of a complicated tricyclic peptidic structure and can be used as chiral selector [58] in capillary electrophoresis and HPLC.

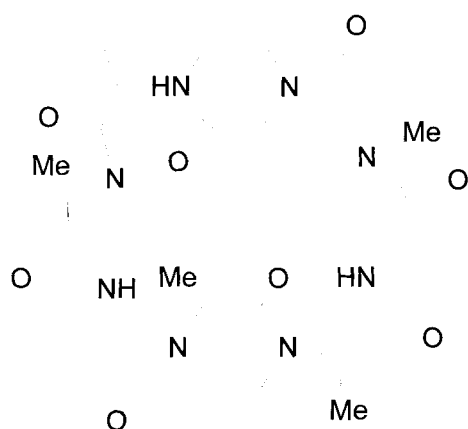
In 1974, Deber and Blout [59, 60] reported for the first time that they could discriminate between ions of D- and L-amino acids with the help of macrocyclic peptides. In <sup>13</sup>C-NMR spectra cyclopeptides c(L-Pro-L-Gly)<sub>3</sub> or c(L-Pro-L-Gly)<sub>4</sub> showed different chemical shifts for the carbon atoms of D- and L-amino acids salts.

Kojima and co-workers [61, 62] prepared a number of macrocyclic pseudopeptides and investigated chiral recognition of amine derivatives. Non natural amino acids like N-methylglycine or non natural dipeptides like (2S,3'S)-2-(2'-oxo-3'-methyl-1'-piperazinyl)propionic acid and (2S,3'S)-4-methyl-2-(2'-oxo-3'-isobutyl-1'-piperazinyl)pentanoic acid have been tested as building blocks for these macrocyclic pseudopeptides (Fig. 10.5). According to <sup>13</sup>C-NMR in CDCl<sub>3</sub>, ammonium or amine groups of guest molecules are localized in the cavity of these receptors [63–65]. Additionally, a hydrophobic interaction between host and guest has been proposed [66, 67]. These macrocyclic peptides could also be used as carriers of chiral amino acids in liquid membrane systems, while L-amino acids have been transported more effectively than D-amino acids [68, 69].

Still and Yoon [70] synthesized an A<sub>4</sub>B<sub>6</sub> cyclo-oligomer from 1,3,5-benzene tricarboxylic acid (A) and (R,R)-1,2-diaminocyclohexane (B). This synthesis yielded two isomers with tetrahedral or D<sub>2</sub> symmetry. Further studies proved the capability of the D<sub>2</sub>-symmetrical isomer to recognize selectively L-amino acids (70–90% ee), and to discriminate between different side chains of amino acids according to their size (phenyl > benzyl, ethyl > methyl) [71].

Ishida et al. [72] prepared a cyclopeptide c(Ala-Aba)<sub>3</sub>, starting from 2-aminobenzoic acid (Aba). This amino acid restricts conformational freedom of the macrocyclus. *p*-Nitrophenylphosphate was found by UV spectroscopy to bind to this macrocyclus with an association constant of  $K_{\text{ass}} = 1.2 \times 10^6 \text{ M}^{-1}$  in a host-guest stoichiometry of 1:1.

Cyclic amides have been prepared by reaction of isophthalic acid chloride with 1,2-phenyldiamine in toluene [73]. Complex formation between the cyclic hexamer and CaCl<sub>3</sub><sup>-</sup> or Ca<sub>2</sub>Cl<sub>4</sub> could be shown by X-ray crystallography.



**Figure 10.5.** Macrocyclic pseudopeptide.

In the past years, new recognition sites have been designed which bear resemblance to the recognition site of antibodies [35]. Hamuro et al. [36] coupled four conformationally restrained peptide loops to a calix[4]arene. The peptide loops were based on a cyclic hexapeptide in which two amino acid residues have been replaced by a 3-aminobenzoyl-dipeptide analogon [74, 75]. This dipeptide analogon contained a 5-amino substituent for attachment to the scaffold, and the high affinity of this antibody mimetic towards cytochrome c was demonstrated by affinity chromatography. Other macrocyclic receptors have been reviewed elsewhere [76–82].

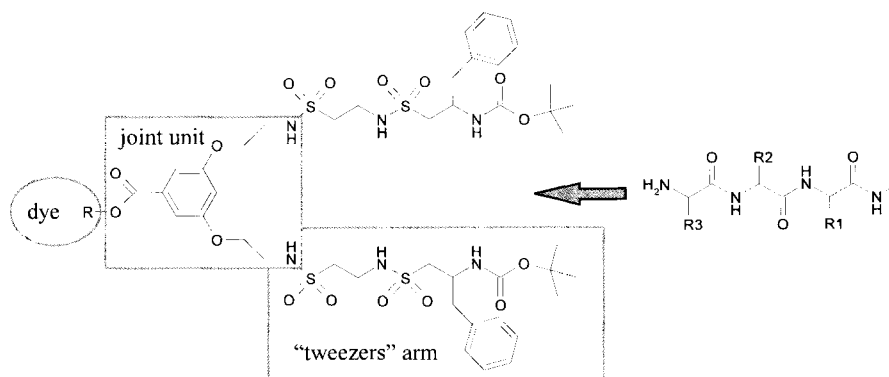
These few examples exhibit cavity-forming structures, making them suitable for receptor library developments.

## 10.4 Combinatorial Receptor Libraries

Although the design and synthesis of receptor molecules for target molecules has made tremendous progress, it is still difficult to design *de novo* a host for a distinct guest. Recent reports have used a biomimetic approach by preparing receptor libraries using combinatorial chemistry. This approach follows the combinatorial strategy which nature demonstrates successfully in the immune system.

On the basis of this concept, Gennari et al. [83] and Löwik et al. [84] developed ‘tweezers-like’, two-armed synthetic receptors made of peptid-sulfonamides as peptidomimetics. Their aim was to investigate the affinity of libraries of these ‘tweezers-like’ receptors to peptides or other biomolecules. The structure of a two-armed receptor, which is shown in Figure 10.6, consists of three areas: First, there is a joint unit to which the arms of the tweezers are connected; then, a dye molecule or a resin bead is attached to the joint unit; and finally, there are the two arms of the tweezers which carry peptidic or peptidomimetic interaction sites. If a ligand with a chromophore has bound to this receptor, the colored resin beads can be selected and characterized.

Boyce et al. [85] designed a library of peptidosteroidal receptors. In this library, the arms of the tweezers were composed of different peptides with a steroid scaffold as joint



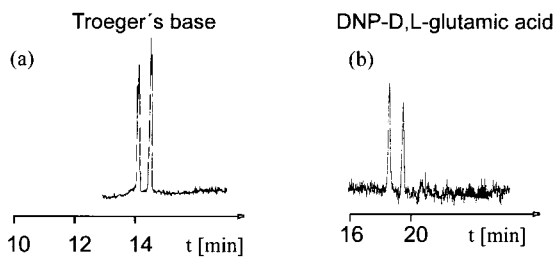
**Figure 10.6.** Structure of a 'tweezers-like' receptor [82].

unit. The binding of 10000 of these receptors to four enkephalines has been checked, and different affinities have been found. In order to find more selective receptors, Still and co-workers investigated the influence of the joint group using steroidal scaffolds carrying linear peptide libraries as tweezers arms. A library of  $10^4$  receptors has been screened with a dye-labeled enkephaline showing remarkable selectivity, with only ten receptors being bound to the enkephaline substrate. Each of these ten receptors had the same amino acid residue in distinct positions [86]. In recent work, Löwik et al. [87] found that they could increase affinity towards tripeptides without losing selectivity by variation of the joint unit. To obtain selective molecular interactions, however, it was not sufficient to restrict flexibility of the joint unit. Nestler [88] found selective receptors to enkephalines in a tweezers library with pentamethylene groups as joint units.

Goodman et al. [89] used metal ions as template for two-armed receptor libraries, and showed that octahedral Ru(II)-complexes with two terpyridine ligands could discriminate between two different ammonium derivatives.

In the group of Jung [90], a new approach has been developed for the use of selector libraries for chiral separations. Complex peptide libraries consisting of  $18^3$  different cyclohexapeptides in each sublibrary have been used as chiral additive in capillary electrophoresis. Three libraries separated enantiomers of  $\alpha$ -amino acid derivatives as well as, e.g., Tröger's base (Fig. 10.7). Meanwhile, our approach has also been applied successfully by Chiari et al. [91], who described the deconvolution of cyclopeptide libraries down to single selectors.

Combinatorial libraries of carbohydrates have been prepared by Patterson et al. [92] by attaching carbohydrates to polyallylamines. Mixtures of carboxylic acids and boric acid which was fluorescence-labeled with anthracene were coupled to polyallylamines. Boric acid served as the anchor position for the carbohydrates. Different receptors were attached to the polyallylamine backbone in close proximity to the boric acid anchors. If a carbohydrate was binding to the anchor unit, the fluorescence due to anthracene was increased. Moreover, if there was also interaction with the receptors, then fluorescence was quenched by photo-induced electron transfer (PET) of close tertiary amines. The response to N-acetylneuraminic acid was more intense than that to either glucose or fructose.



**Figure 10.7.** Enantiomeric resolution of Tröger's base (a) and DNP-D,L-glutamic acid (b) with *c*(DFXXXa). Conditions: 10 mM cyclopeptide library *c*(DFXXXa) in phosphate buffer pH 7.4 (20 mM), capillary 50  $\mu\text{m}$  i.d.  $\times$  50 cm effective length, 20 kV.

Combinatorial approaches are not restricted to studies of host–guest interactions of organic molecules, but can also be applied for the development of selective ligands for metal ions. Burger and Still [93] demonstrated that introduction of peptide chains in macrocyclic tetra-amines led to a variation of binding affinity to  $\text{Co}^{2+}$  and  $\text{Cu}^{2+}$  ions depending on the amino acid sequence of the peptide chain. However, no ligand could be isolated from the libraries with higher affinity to  $\text{Co}^{2+}$  than to  $\text{Cu}^{2+}$ .

Francis and co-workers [94] prepared a library of linear peptides by split-mix synthesis, and a turn-inducing building block was introduced into the peptide chain. In this way, the peptide chain was able to complex metal ions. In an assay with different transition metals, each metal was able to bind selectively to a number of library components.

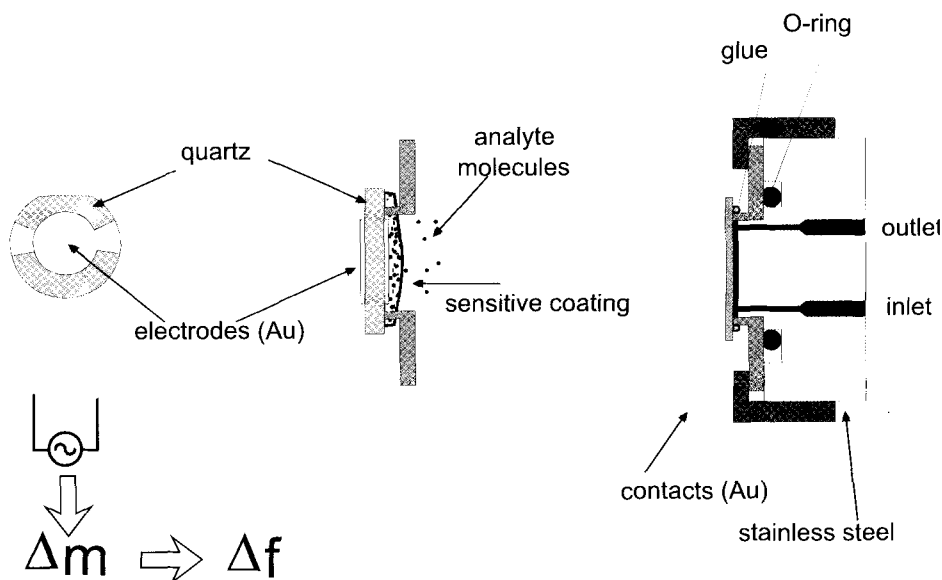
## 10.5 Cyclopeptides as Supramolecular Recognition Sites for Chemosensors

So far, host–guest interactions have been investigated only in homogeneous phase. However, application in chemosensing devices requires covalent attachment of the receptors to the surface. Valinomycin [95] and gramicidin [96] were successfully applied as sensors for potassium ions in bilayers. The bilayers have been covalently attached to gold electrodes.

In this section, we wish to discuss the results of recent studies on cyclopeptides carried out by our group [97–99]. Cyclopeptides as molecular receptors have been immobilized on different sensor surfaces, and can be prepared in high diversity by parallel solid-phase peptide synthesis (SPPS) [100]. However, functional groups for the attachment to the surface must be available. On the other hand, the interaction with the analyte should be as specific as possible, and for this reason appropriate side chains must be selected.

### 10.5.1 Quartz Microbalance Measurements in the Liquid Phase

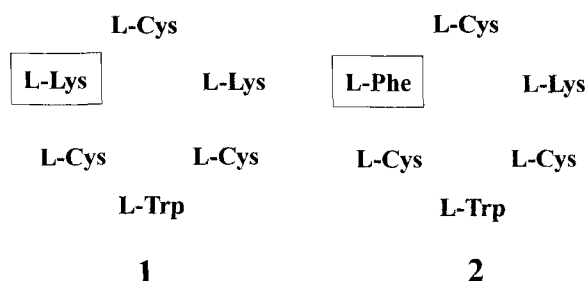
Quartz microbalances are mass-sensitive sensors consisting of a quartz slide with gold electrodes on each side of the slide (Fig. 10.8). To investigate cyclopeptides as new su-



**Figure 10.8.** Principle of analyte measurement in a flow cell [101].

pramolecular recognition sites, the peptides must be attached to the gold electrodes. This can be achieved simply by the thiol group in the side chain of cysteine. Due to the piezoelectronic properties of the quartz slide, application of alternating potential to the electrodes initiates vibrations. If there is interaction between the cyclopeptide and the analyte, the resonance frequency of the quartz changes. Additional mass decreases the resonance frequency.

We have investigated a large variety of cyclopeptides and analytes using quartz sensors in gas and liquid phases. Here, we report on the interaction of cyclopeptides **1** and **2** (Fig. 10.9) with proteinogenic amino acids in aqueous solution. Both peptides contain three thiol groups for the covalent and conformationally restricting attachment to the gold electrodes. They differ in only one amino acid (L-Lys in **1**, L-Phe in **2**).

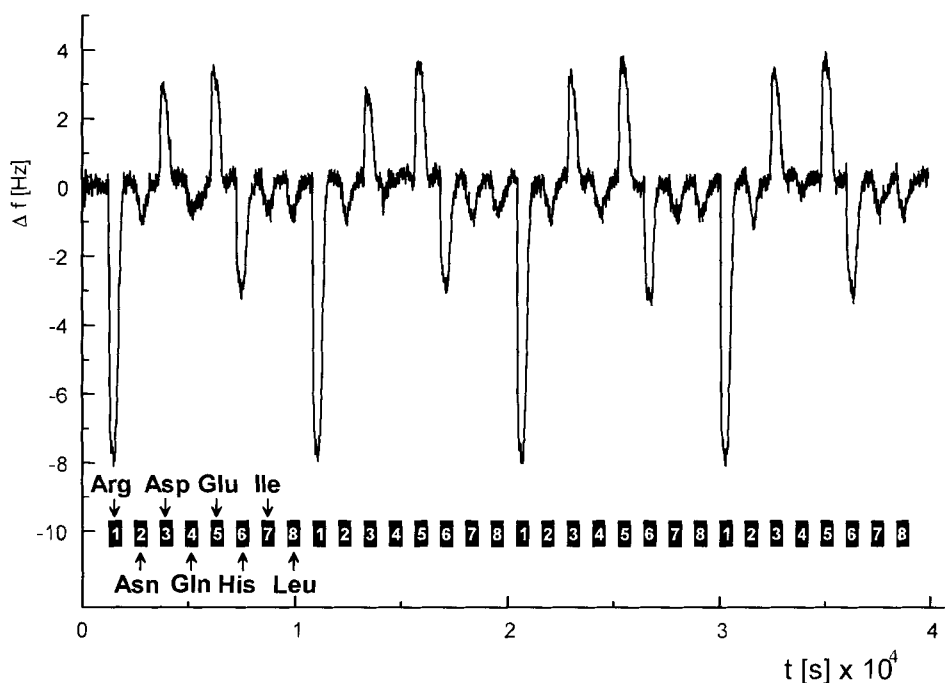


**Figure 10.9.** Amino acid sequences of cyclopeptides **1** and **2**, each attached via three cysteine thiol groups to gold surfaces [99].

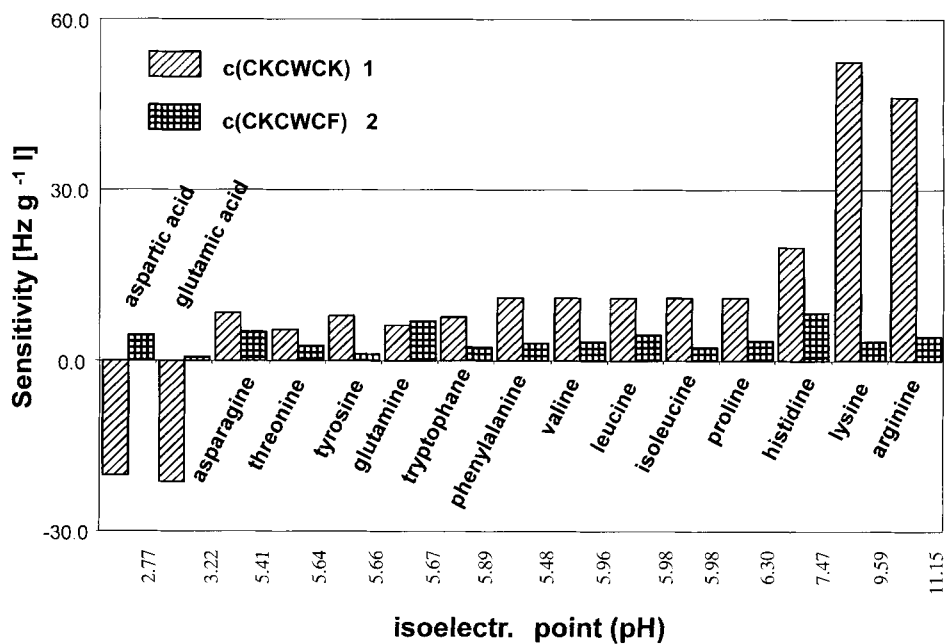
Acidic, neutral or basic amino acids have been used as analytes dissolved in phosphate buffer (0.1 M, pH 7.4). Under these conditions, basic groups like the side chains of arginine, lysine and to some extent also histidine are protonated and therefore cationic. Acidic groups like the side chains of aspartic or glutamic acid are deprotonated and therefore anionic. Tryptophan was included in the macrocyclus because its indole system offers  $\pi$ -cation interactions or hydrophobic interactions.

Repeated measurements with **1** and eight different amino acids are shown in Figure 10.10. Frequency increase by L-glutamic and L-aspartic acid can be explained by pushing away hydrated phosphate ions from the surface. Altogether, this effect can reduce the adsorbed mass.

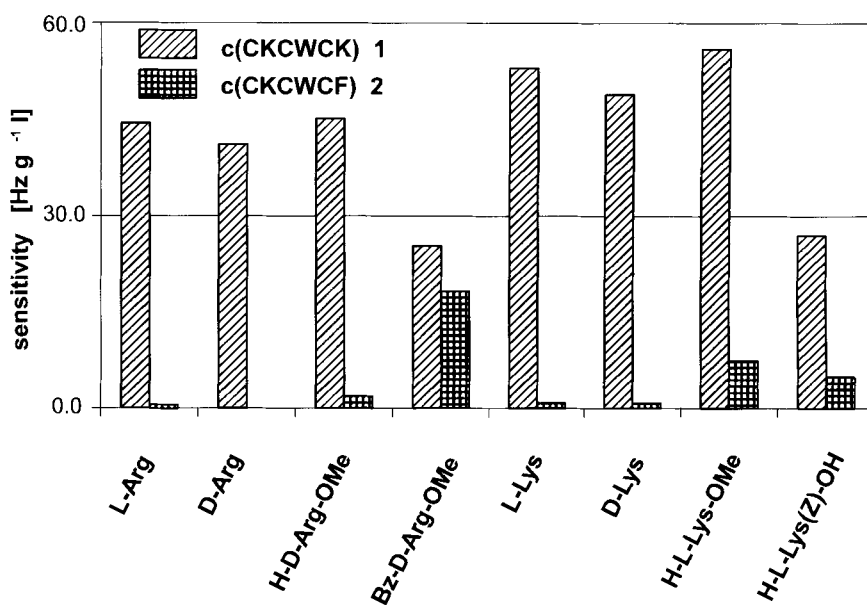
Figure 10.11 summarizes the measurements for 15 different amino acids. The order of the amino acids is according to their isoelectric points. Interestingly, cyclopeptide **2** is almost not sensitive, whereas the coating with cyclopeptide **1** can discriminate significantly between amino acids of acidic, neutral or basic character. L-Lysine and L-arginine show the strongest interactions, but the response is different between these two amino acids. This strong interaction could be due to direct ionic interactions between amino acid, cyclopeptide **1** and phosphate ions. When we used protected derivatives of lysine and arginine, we were able to confirm that the interaction between **1** and lysine or arginine is mainly dependent on the basic groups of these amino acids (Fig. 10.12). While esterifica-



**Figure 10.10.** Frequency shifts of a quartz microbalance coated with cyclopeptide **1** measured for eight different amino acids (1 mmol) in phosphate buffer (0.1 mol/l, pH 7.4). The whole measurement was repeated four times [99].

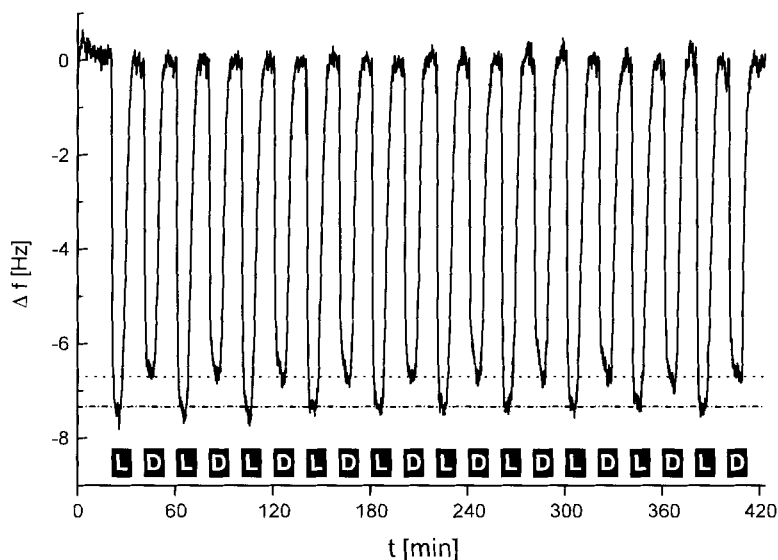


**Figure 10.11.** Sensitivity of quartz microbalances coated with cyclopeptides **1** or **2** for 15 different L-amino acids (1 mmol/l) in phosphate buffer (0.1 mol/l, pH 7.4) [99].



**Figure 10.12.** Sensitivity of quartz microbalances coated with cyclopeptides **1** or **2** for different arginine or lysine derivatives (1 mmol/l) in phosphate buffer (0.1 mol/l, pH 7.4); Bz = benzoyl, Z = benzyloxycarbonyl [99].





**Figure 10.13.** Frequency shifts of a quartz microbalance coated with cyclopeptide **2** exposed to solutions of L- and D-arginine (1 mmol/l in 0.1 mol/l phosphate buffer at pH 7.4) [99].

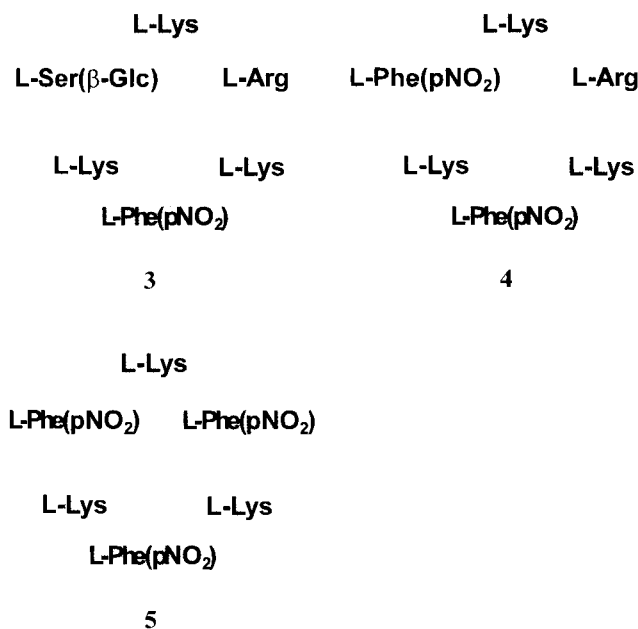
tion of the carboxy group only has minor effects, blocking of the  $\alpha$ -amino group of arginine or of the  $\epsilon$ -amino group of lysine reduces sensitivity by ca. 50%.

Interestingly, sensitivity of a quartz microbalance coated with **2** significantly increased when this quartz microbalance was exposed to a solution containing arginine with benzoyl-protected  $\alpha$ -amino group. This could be due to a stronger interaction of this protecting group with the two aromatic side chain residues in **2**.

The capability of quartz microbalances for the analysis of enantiomers has been demonstrated successfully by Bodenhöfer et al. [102]. Cyclopeptides have already been successful as chiral selectors in enantiomer separation by capillary electrophoresis [90, 91]. Here, we showed that cyclopeptide **1** exhibits different affinities to D- and L-arginine [99]. The differences in sensitivity of the quartz microbalances result in a chiral separation factor of 1.12 [76]. Sensitivity differences of the quartz microbalance coated with **1** for the two enantiomers of arginine are shown in Figure 10.13.

### 10.5.2 Reflectometric Interference Spectroscopy (RIFS) in the Liquid Phase

Cyclohexapeptide monolayers on quartz microbalances are able to discriminate between different analytes in the liquid phase (see Section 10.5.1). On the basis of these results, we also have immobilized cyclopeptides (Fig. 10.14) on glass transducers. In this case, the interaction between cyclopeptides and analytes was monitored by reflectometric interference spectroscopy (RIFS) [98]. RIFS is an optical detection method in which the phenomenon of reflection and interference of light at phase boundaries is used to measure changes in optical thickness (refractive index  $\times$  layer thickness) of transparent films

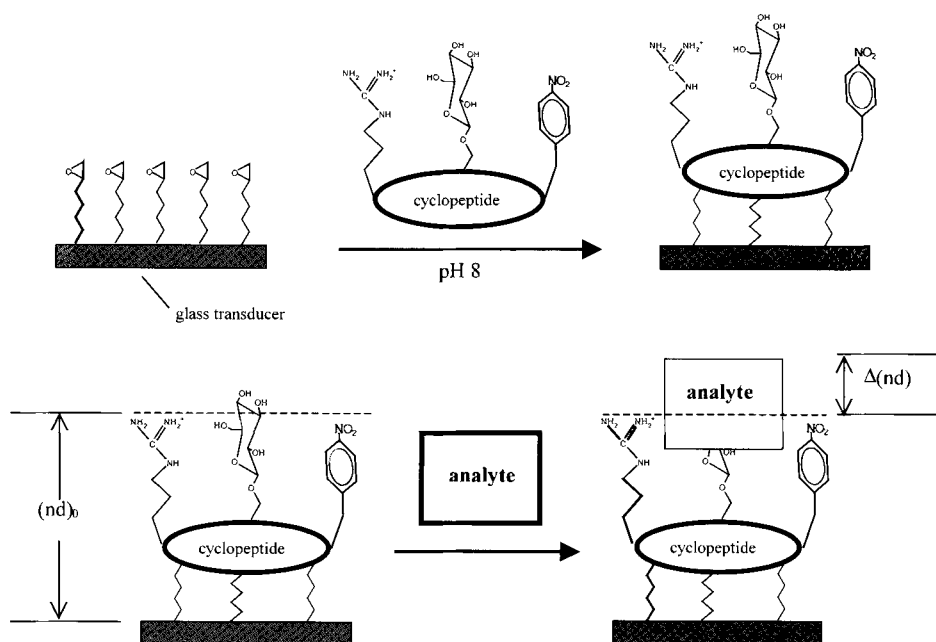


**Figure 10.14.** Amino acid sequences of examples of cyclopeptides used for attachment via lysine side chains to glass surfaces [97].

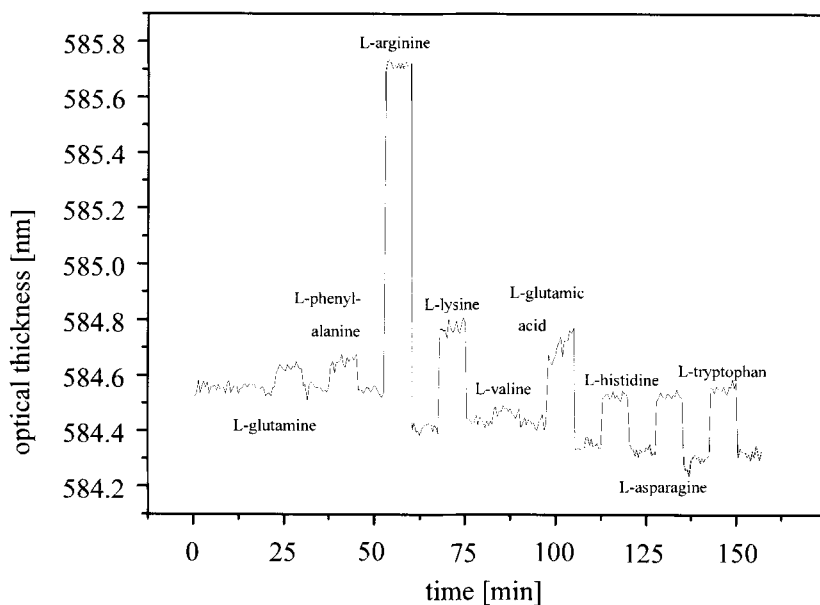
by shift of the interference pattern [103]. Adsorption and binding processes of thin cyclopeptide layers can be monitored online with this technique. Furthermore, numerous analytes and receptors can be measured within a relatively short time. Among the variety of different cyclopeptides investigated, the immobilization of a monolayer of cyclohexapeptides **3**, **4** and **5** (Fig. 10.14) by amino groups of the side chain of lysine on a transducer surface presenting epoxy groups is shown in Figure 10.15.

L-Lysine residues in positions 1, 3 and 5 have been chosen as anchor residues. In positions 2, 4 and 6 L-arginine, 4-nitro-L-phenylalanine and O-glycosidically modified serine have been used. It was known from capillary electrophoresis experiments that L-arginine exhibits important effects [90]. In the tricyclic peptide antibiotic vancomycin, which is used as a chiral stationary phase, sugar residues and aromatic residues play an important role [58]. The interactions between amino acids and cyclopeptides were investigated in neutral water without buffer, cyclopeptide layers having been exposed to solutions of amino acids. After reaching an equilibrium the resulting optical layer thickness was determined. A time-resolved measurement of different amino acids is shown in Figure 10.16.

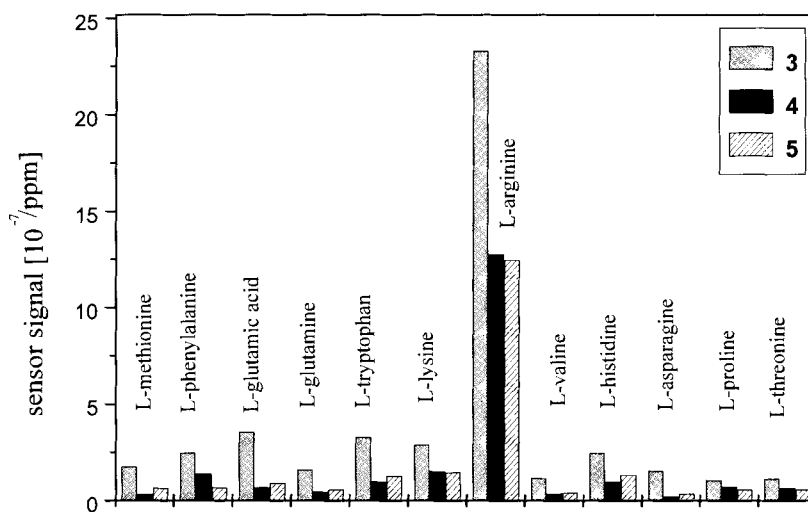
The results for the three surface-bound cyclopeptides and 12 L-amino acids are summarized in Figure 10.17. The sensor signals show that all three cyclopeptides interact strongly with L-arginine, in contrast to L-lysine which is also a basic amino acid. Signals of peptide layer **3** are always more intense than signals of layers of **4** or **5**, which means that the glucose residue has an amplifying effect. For arginine, the signal of **3** is doubled compared with the signal of **4** or **5**.



**Figure 10.15.** Immobilization of a cyclohexapeptide on a RIFs transducer, and schematic representation of the resulting change in optical thickness ( $\Delta(nd)$ ) on interaction with an analyte [97, 98].



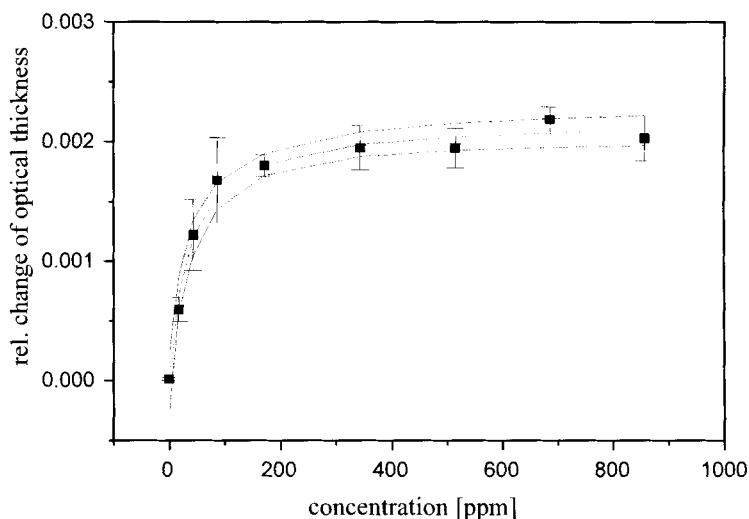
**Figure 10.16.** Changes in optical thickness measured with a transducer coated with cyclohexapeptide 3 after contact with solutions of amino acids in water (1000 ppm). 1 ppm corresponds to 1 mg of amino acid per liter [97].



**Figure 10.17.** Comparison of sensor signals of the three cyclopeptide receptor layers sensitive for amino acids in buffer-free neutral water [97].

Chiral discrimination between D- and L-arginine with the layer of **3** was within experimental error. Figure 10.18 shows a concentration-dependent curve for L-arginine. In order to quantify these experimental results, they have been fitted as Langmuir adsorption [104],

$$Y = A + \frac{(b \cdot c \cdot x)}{(1 + b \cdot x)}$$



**Figure 10.18.** Calibration curve of a layer of cyclopeptide **3** for L-arginine and a fitted Langmuir curve [97].

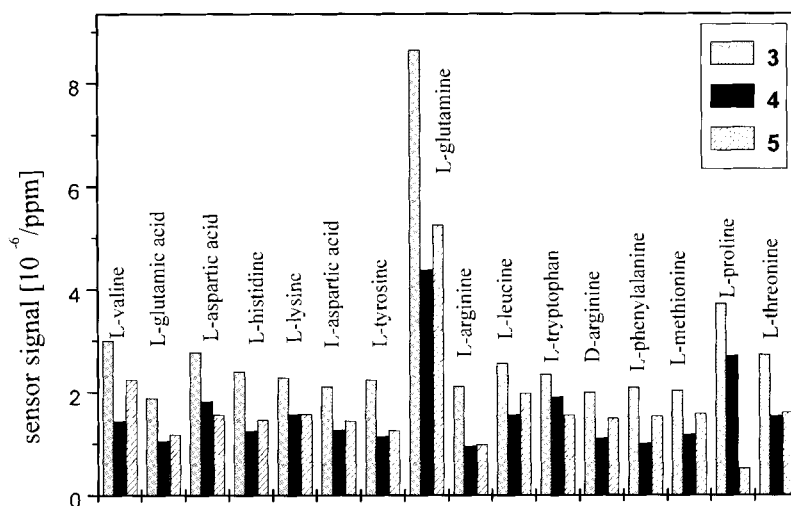
where  $Y$  is signal intensity,  $x$  is concentration,  $A$  is the axis intercept,  $b$  is the relationship of adsorption and desorption constant, and  $c$  is the saturation signal. With this equation, the detection limit for arginine can be calculated by setting  $Y$  as three-fold the noise signal of the baseline. The calculated  $x$ -values for L-arginine in Table 10.2 show that a monolayer of cyclopeptide 3 can detect a concentration of ca. 15  $\mu\text{mol/l}$  for arginine in neutral, buffer-free water.

However, sensor signals strongly depend on the matrix: if amino acid solutions in phosphate-buffered saline (PBS), pH 7.4 are used, the selectivity pattern changes.  $\alpha$ -Amino acids are mainly zwitterionic in PBS. Among a set of 14 amino acids, L-glutamine yields the highest signal and can be discriminated from the very similar L-asparagine (Fig. 10.19). Again, the sugar residue in cyclopeptide 3 has an amplifying effect. Although cyclopeptides 4 and 5 differ in only one amino acid residue, the signal pattern for the amino acids is different.

This example shows that combinatorial variations of cyclopeptides and variation of the matrix allow the search for selectivity for other amino acids. The principle of cyclopeptide

**Table 10.2.** Detection limits of all three cyclopeptides for L-arginine (1  $\mu\text{mol/l}$  corresponds to 0.174 mg/l).

Cyclohexapeptide	L-arginine [ $\mu\text{mol/l}$ ]
3	14.4
4	40.2
5	20.1



**Figure 10.19.** Signals of layers of cyclopeptides 3, 4 and 5 for different amino acids in PBS using RIfS measurements [97].

layers as selective recognition sites for small-molecular weight compounds has also been successfully applied to volatile organic compounds in the gaseous phase. Again, measurements were either carried out on quartz microbalances with thiolated cyclopeptides [105], or on RIfS transducers with cyclopeptides covalently attached by amino groups of lysine [98].

## 10.6 Outlook

In the last section, we demonstrated that cyclopeptides are valuable molecular receptors for chemosensors. By using combinatorial chemistry, cyclopeptide libraries (virtual library:  $10^{16}$  cyclopeptides from commercially available natural and non natural amino acids) are accessible and may be immobilized as sensitive layers. The exceptional diversity of cyclopeptides allows screening for selectivities for different analytes. Combinatorial approaches to molecular receptors are particularly versatile using the cyclopeptide approach due to the large number of available building blocks. The multitude of cyclopeptides offers the preparation of sensor arrays with systematic variations. Their properties can be influenced by choice of the sensitive layers and sensor data can be evaluated using established methods of pattern recognition. Although in these experiments shown we have used cyclopeptide *monolayers*, cyclopeptides may also be attached to three-dimensional matrices in order to increase sensitivity.

## 10.7 References

- [1] U. E. Spichiger-Keller (ed.), *Chemical Sensors and Biosensors for Medical and Applications*, VCH-Wiley, Weinheim, **1997**.
- [2] W. Göpel, Ch. Ziegler, H. Breer, D. Schild, R. Apfelbach, J. Joerges, R. Malaka, *Biosens. Bioelectron.* **1998**, *13*, 479–493.
- [3] Ch. Ziegler, W. Göpel, H. Hämmerle, H. Hatt, G. Jung, L. Laxhuber, H.-L. Schmidt, S. Schütz, F. Vögtle, A. Zell, *Biosens. Bioelectron.* **1998**, *13*, 539–571.
- [4] A. J. Ricco, R. M. Crooks, *Acc. Chem. Res.* **1998**, *31*, 200–202.
- [5] J. Janata, M. Josowicz, P. Vanysek, M. D. DeVaney, *Anal. Chem.* **1998**, *70*, 179–209.
- [6] W. Göpel, J. Hesse, J. N. Zemel (ed.), *Sensors – A Comprehensive Survey*, VCH-Wiley, Weinheim, **1991**.
- [7] W. Göpel, *Chemische Sensoren: Übersicht, Ionensensitive Elektroden, Optochemische Sensoren, Massensensitive Sensoren, Kalorimetrische Sensoren, Multikomponentenanalyse für die chemische Sensorik*, in: R. Grabowski (ed.), *Sensoren und Aktoren*, VDE-Verlag, Berlin (FRG), **1991**.
- [8] U. Weimar, S. Vaihinger, K.D. Schierbaum, W. Göpel, *Multicomponent Analysis in Chemical Sensing*, in: N. Yamazoe (ed.), *Chemical Sensor Technology*, Vol. III, Kodansha Ltd., Tokyo, **1991**.
- [9] K.D. Schierbaum and W. Göpel, *Chemische Sensoren: Grundlagen und Anwendungen in Arrays*, in: H. Ahlers (ed.), *Multisensorik Praxis*, Springer, Berlin, **1996**.
- [10] H. Baltes, W. Göpel, and J. Hesse, *Sensors Update: Sensor Technology – Applications – Markets*, (Series eds.), VCH-Wiley, Weinheim, Vol. 1, **1996**, Vol. 2, **1996**.
- [11] F. L. Dickert, O. Schuster, *Chemie in unserer Zeit* **1994**, *28*, 147–152.
- [12] C. Gennari, H. P. Nestler, U. Piarulli, B. Salom, *Liebigs Ann./Recueil* **1997**, 637–647.
- [13] J. Janata, M. Josowicz, P. Vanysek, M. D. DeVaney, *Anal. Chem.* **1998**, *70*, 179–209.
- [14] J. Mitrovics, H. Ulmer, U. Weimar, W. Göpel, *Acc. Chem. Res.* **1998**, *31*, 307–316.
- [15] J. Rickert, W. Göpel, W. Beck, G. Jung, P. Heiduschka, *Biosens. Bioelectron.* **1996**, *11*, 757–768.
- [16] C. Striebel, A. Brecht, G. Gauglitz, *Biosens. Bioelectron.* **1994**, *9*, 139–146.

- [17] A. Brecht, G. Gauglitz, J. Polster, *Biosens. Bioelectron.* **1993**, *8*, 387–392.
- [18] G. Gauglitz, A. Brecht, G. Kraus, W. Nahm, *Sens. Act. B* **1993**, *11*, 21–27.
- [19] G. G. Guibault, A. Sufeiman, *Am. Biotech. Lab.* **1990**, *8*, 28–32.
- [20] J. Rickert, T. Weiss, W. Kraas, G. Jung, W. Göpel, *Biosens. Bioelectron.* **1996**, *11*, 591–598.
- [21] B. Liedberg, C. Nylander, I. Lundström, *Sens. Act. B* **1983**, *4*, 299–304.
- [22] W. Knoll, *Rev. Phys. Chem.* **1998**, *49*, 569–638.
- [23] R. Cush, J. M. Cronin, W. J. Stewart, C. H. Maule, J. Molloy, N. J. Goddard, *Biosens. Bioelectron.* **1993**, *8*, 347–364.
- [24] P. M. Nellen, W. Lukosz, *Sens. Act. B* **1990**, *1*, 592–596.
- [25] A. Brandenburg, R. Polzius, F. Bier, U. Bilitewski, E. Wagner, *Sens. Act. B* **1996**, *30*, 55–59.
- [26] J. Hartmann, R. Auge, S. Lucklum, S. Rösler, P. Hauptmann, *Sens. Act. B* **1996**, *34*, 305–311.
- [27] J.-M. Jehn, *Supramolecular Chemistry: Concepts and Perspectives*, Wiley-VCH, Weinheim, **1995**.
- [28] B. Dietrich, P. Viont, J.-M. Lehn, *Macrocyclic Chemistry: Aspects of Organic and Inorganic Supramolecular Chemistry*, Wiley-VCH, Weinheim, **1997**.
- [29] V. Böhmer, *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 713–744.
- [30] L. C. Groenen, D. N. Reinhoudt, *NATO ASI Ser.* **1992**, *371*, 51–70.
- [31] A. Ikeda, S. Shinkai, *Chem. Rev.* **1997**, *97*, 1713–1734.
- [32] J. Schmitz, F. Vögtle, M. Nieger, K. Gloe, H. Stephan, O. Heitzsch, H.-J. Buschmann, W. Hasse, K. Cammann, *Chem. Ber.* **1993**, *126*, 2483–2491.
- [33] Z. Asfari, S. Wenger, J. Vicens, *Pure & Appl. Chem.* **1995**, *67*, 1037–1045.
- [34] R. J. W. Lugtenberg, Z. Brzecz, K. Gloe, H. Stephan, O. Heitzsch, H.-J. Buschmann, W. Hasse, *Anal. Chim. Acta* **1995**, *310*, 263–267.
- [35] J. O. Magrans, A. R. Ortiz, M. A. L. P. H. P. Molins, J. Sanchez-Quesada, P. Prados, M. Pons, F. Cago, J. de Mendoza, *Angew. Chem. Int. Ed. Engl.* **1996**, *108*, 1712–1715.
- [36] Y. Hamuro, M. C. Calama, H. S. H. A. D. Park, A. D. Hamilton, *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2680–2683.
- [37] M. C. Calama, R. Hulst, R. Fokkens, N. M. M. Nibbering, P. Timmerman, D. N. Reinhoudt, *Chem. Commun.* **1998**, *9*, 1021–1022.
- [38] P. Timmerman, D. N. Reinhoudt, *Adv. Mater.* **1999**, *11*, 71–74.
- [39] J. Szejtli, *Chem. Rev.* **1998**, *98*, 1743–1753.
- [40] A. R. Hedges, *Chem. Rev.* **1998**, *98*, 2035–2044.
- [41] K. Uekama, F. Hirayama, T. Irie, *Chem. Rev.* **1998**, *98*, 2045–2076.
- [42] V. Schurig, H. P. Nowotny, *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 939–952.
- [43] W. A. König, *Trends Anal. Chem.* **1993**, *12*, 130–142.
- [44] V. Schurig, M. Jung, S. Mayer, S. Negura, M. Fluck, H. Jakubetz, *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2222–2224.
- [45] V. Schurig, in: *Chromatographic separations based on molecular recognition* (K. Jinnó, ed.), Wiley, New York, **1997**.
- [46] A. R. Khan, P. Forgo, K. J. Stine, V. T. D'Souza, *Chem. Rev.* **1998**, *98*, 1977–1996.
- [47] D. J. Cram, R. H. Bauer, *J. Am. Chem. Soc.* **1959**, *81*, 5983–5987.
- [48] M. E. Garcia, J. A. Gavin, N. Deng, A. A. Andrievsky, T. E. Malouk, *Tetrahedron Lett.* **1996**, *37*, 8313–8316.
- [49] C. J. Pedersen, *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 1021–1026.
- [50] C. J. Pedersen, *J. Am. Chem. Soc.* **1967**, *89*, 7017–7036.
- [51] H.-J. Schneider, U. Schneider, *J. Inclusion Phenom. Mol. Recognit. Chem.* **1994**, *19*, 67–83.
- [52] E. U. Thoden van Velzen, J. F. J. Engbersen, D. N. Reinhoudt, *J. Am. Chem. Soc.* **1994**, *116*, 3597–3598.
- [53] W. Göpel, *Synth. Metals* **1991**, *41*, 1087–1093.
- [54] F. H. Moser, A. L. Thomas, *The phthalocyanines – properties*, CRC Press, Boca Raton, **1983**.
- [55] R. M. Izatt, J. S. Bradshaw, K. Pawlak, R. L. Bruening, B. J. Tarbet, *Chem. Rev.* **1992**, *92*, 1261–1354.
- [56] L. Stryer (ed.): *Biochemistry*, 4<sup>th</sup> edition, Freeman, New York, **1995**.
- [57] D. W. Armstrong, Y. Tang, S. Chen, Y. Zhou, C. Bagwill, J. R. Chen, *Anal. Chem.* **1994**, *66*, 1473–1475.
- [58] D. W. Armstrong, K. Rundlett, J. R. Chen, *Chirality* **1994**, *6*, 496–497.
- [59] C. Deber, E. R. Blout, *J. Am. Chem. Soc.* **1974**, *96*, 7566–7568.

- [60] V. Madison, C. Deber, E. R. Blout, *J. Am. Chem. Soc.* **1977**, *99*, 4788–4798.
- [61] H. Miyake, K. Shibata, Y. Kojima, T. Yamashita, A. Ohsuka, *Makromol. Chem., Rapid Commun.* **1990**, *11*, 667–671.
- [62] Y. Koshima, T. Yamashita, M. Washizawa, A. Ohsuka, *Makromol. Chem., Rapid Commun.* **1989**, *10*, 121–125.
- [63] H. Miyake, Y. Kojima, T. Yamashita, A. Ohsuka, *Chem. Soc. Jpn.* **1992**, *65*, 917–919.
- [64] H. Miyake, Y. Kojima, T. Yamashita, A. Ohsuka, *Makromol. Chem.* **1993**, *194*, 1925–1933.
- [65] T. Yamashita, J. Maruo, A. Fujimoto, K. Shibata, Y. Kojima, A. Ohsuka, *Makromol. Chem.* **1990**, *191*, 1261–1268.
- [66] H. Kataoka, T. Hanawa, T. Katagi, *Chem. Pharm. Bull.* **1992**, *40*, 570–574.
- [67] H. Kataoka, T. Hanawa, T. Katagi, *Tetrahedron* **1987**, *43*, 4519–4530.
- [68] H. Miyake, Y. Kojima, *Coord. Chem. Rev.* **1996**, *148*, 301–314.
- [69] H. Miyake, Y. Kojima, T. Yamashita, H. Tsukube, *Tetrahedron Lett.* **1995**, *36*, 7669–7672.
- [70] S. S. Yoon, W. C. Still, *Tetrahedron Lett.* **1994**, *35*, 8557–8560.
- [71] M. Torneiro, W. C. Still, *J. Am. Chem. Soc.* **1995**, *117*, 5887–5888.
- [72] H. Ishida, M. Suga, K. Donowaki, K. Ohkubo, *J. Org. Chem.* **1995**, *60*, 5374–5375.
- [73] Y. H. Kim, J. Calabrese, C. McEwen, *J. Am. Chem. Soc.* **1996**, *118*, 1545–1546.
- [74] M. Nigam, C. M. Seong, Y. Qian, M. A. Blaskovich, A. D. Hamilton, S. M. Sebt, *J. Biol. Chem.* **1993**, *268*, 20695–20698.
- [75] A. C. Bach, C. J. Eyerman, J. D. Gross, M. J. Bower, R. L. Harlow, P. C. Weber, W. F. DeGrado, *J. Am. Chem. Soc.* **1994**, *116*, 3207–3219.
- [76] F. P. Schmidtchen, M. Berger, *Chem. Rev.* **1997**, *97*, 1609–1646.
- [77] J. A. Semlyen (ed.), *Large Ring Molecules*, Wiley & Sons, New York, **1996**.
- [78] M. M. Morgan, J. Rebeck, *Chem. Rev.* **1997**, *97*, 1647–1668.
- [79] J. C. Ma, D. A. Dougherty, *Chem. Rev.* **1997**, *97*, 1303–1324.
- [80] P. Wallimann, T. Marti, A. Fürer, F. Diederich, *Chem. Rev.* **1997**, *97*, 1567–1608.
- [81] X. X. Zhang, J. S. Bradshaw, R. M. Izatt, *Chem. Rev.* **1997**, *97*, 3313–3361.
- [82] J. Dowden, P. D. Edwards, S. S. Flack, J. D. Kilburn, *Chem. Eur. J.* **1999**, *5*, 79–89.
- [83] C. Gennari, H. P. Nestler, B. Salom, W. C. Still, *Angew. Chem. Int. Ed.* **1995**, *34*, 1765–1768.
- [84] D. W. P. M. Löwik, S. J. E. Mulders, Y. Cheng, Y. Shao, R. M. J. Liskamp, *Tetrahedron Lett.* **1996**, *37*, 8253–8256.
- [85] R. Boyce, G. Li, H. P. Nestler, T. Suenaga, W. C. Still, *J. Am. Chem. Soc.* **1994**, *116*, 7955–7956.
- [86] Y. Cheng, T. Suenaga, W. C. Still, *J. Am. Chem. Soc.* **1996**, *118*, 1813–1814.
- [87] D. W. P. M. Löwik, M. D. Weingarten, M. Broekma, A. J. Brouwer, W. C. Still, R. M. J. Liskamp, *Angew. Chem. Int. Ed. Engl.* **1998**, *37*, 1846–1850.
- [88] H. P. Nestler, *Mol. Diversity* **1996**, *2*, 35–40.
- [89] M. S. Goodman, V. Jubian, B. Linton, A. D. Hamilton, *J. Am. Chem. Soc.* **1995**, *117*, 11610–11611.
- [90] G. Jung, H. Hofstetter, S. Feiertag, D. Stoll, O. Hofstetter, K.-H. Wiesmüller, V. Schurig, *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2148–2150.
- [91] M. Chiari, V. Desperati, E. Manera, R. Longhi, *Anal. Chem.* **1998**, *70*, 4967–4973.
- [92] S. Patterson, B. D. Smith, R. E. Taylor, *Tetrahedron Lett.* **1998**, *39*, 3111–3114.
- [93] M. T. Burger, W. C. Still, *J. Org. Chem.* **1995**, *60*, 7382–7383.
- [94] M. B. Francis, N. S. Finney, E. N. Jacobsen, *J. Am. Chem. Soc.* **1996**, *118*, 8983–8984.
- [95] B. Raguse, V. Braach-Maksvytis, B. A. Cornell, L. G. King, P. D. J. Osman, R. J. Pace, L. Wiczorek, *Langmuir* **1998**, *14*, 648–659.
- [96] B. A. Cornell, V. Braach-Maksvytis, L. G. King, P. D. J. Osman, B. Raguse, L. Wiczorek, R. J. Pace, *Nature* **1997**, *387*, 580–583.
- [97] D. Leipert, D. Nopper, M. Bauser, G. Gauglitz, G. Jung, *Angew. Chem. Int. Ed.* **1998**, *37*, 3308–3311.
- [98] D. Leipert, F. Rathgeb, M. Herold, J. Mack, G. Gauglitz, G. Jung, *Anal. Chim. Acta.*, in press.
- [99] T. Weiß, D. Leipert, G. Jung, W. Göpel, *Adv. Mater.* **1999**, in press.
- [100] G. Jung (ed.), *Combinatorial peptide and nonpeptide libraries*, VCH-Wiley, Weinheim, **1996**.
- [101] J. Rickert, A. Brecht, W. Göpel, *Anal. Chem.* **1997**, *69*, 1441–1448.
- [102] K. Bodenhöfer, A. Hierlemann, J. Seemann, G. Gauglitz, B. Koppenhoefer, W. Göpel, *Nature* **1997**, *387*, 577–580.
- [103] A. Brecht, J. Piehler, G. Lang, G. Gauglitz, *Anal. Chim. Acta.* **1995**, *311*, 289–300.
- [104] *Comprehensive Polymer Science*, Pergamon, Oxford, **1989**.
- [105] D. Leipert, T. Weiß, J. Mack, W. Göpel, G. Jung, *Biosensors* **1998**, Berlin, *Proceedings*, 443.



This Page Intentionally Left Blank

# 11 Peptide Libraries in T-Cell-Mediated Immune Response

*B. Fleckenstein, G. Jung and K.-H. Wiesmüller*

## 11.1 Introduction

T lymphocytes have adopted a central position in the induction and regulation of immune response against infectious agents and tumor cells. Insufficient regulation and activity of T lymphocytes can lead to autoimmune response and tumor growth. The principles of antigen recognition by T lymphocytes have to be elucidated in order to understand the specificity of T-cell-mediated immunity. Non self peptides are recognized in the context of self major histocompatibility complex (MHC) molecules on the surface of cells, leading to the destruction of the target cells or to the production of protective antibodies. The rules for peptide selection by MHC molecules are based on sequences of single peptides, and natural peptide libraries have been eluted from MHC proteins. The analysis of mutations in sequences of known T-cell epitopes on peptide binding to MHC molecules and on T-cell responses, crystal structure analyses of defined peptide–MHC complexes, as well as the investigation of peptide libraries were important steps for the *current understanding of cellular immune response*.

### 11.1.1 Generation and Presentation of Antigens

Intracellular proteins are marked for proteolytic degradation by ubiquitinylation and are cleaved by the proteasome complex in the cytosol [1]. Following degradation, the generated octa- to hexadecapeptides are translocated into the endoplasmic reticulum (ER) by the Transporter Associated with Antigen Processing (TAP) localized in the ER-membrane [2, 3]. After translocation and assembly of MHC class I and  $\beta_2$ -microglobulin, peptides are loaded to MHC class I molecules and the complex migrates to the cell surface for recognition by  $CD8^+$  T cells.

Extracellular proteins are taken up by antigen-presenting cells through receptor-mediated endocytosis, phagocytosis or pinocytosis. As an alternative to MHC class I, MHC class II molecules interact with peptides generated in this endocytotic pathway. Protein degradation occurs in endosomes and lysosomes [4, 5] and MHC class II molecules are loaded with peptides in the MHC class II compartment. The MHC–peptide complex migrates to the cell surface for recognition by  $CD4^+$  T cells.

### 11.1.2 Combinatorial Peptide Libraries

Large collections of synthetic peptides were introduced in 1986 to identify antigens. By means of solid-phase synthesis, more than one amino acid were incorporated at selected coupling cycles, yielding a degenerate sequence position. Such peptide mixtures can be described by the number and the position(s) of so-called defined sequence positions, where only one amino acid was coupled [6]. Soluble synthetic peptide libraries have been applied to search for an antigen [7, 8] or to identify streptavidin-binding peptides by screening a pentapeptide library immobilized on single resin beads followed by micro-sequencing of active beads [9]. Numerous applications of the peptide library approach in basic research and for the identification of peptide lead structures for drug development have boosted the field of combinatorial organic chemistry [10, 11].

To synthesize highly diverse libraries, carrying more than six degenerated sequence positions (X), several methods using premixed mixtures of amino acids for the introduction of X positions in the sequence have been described [12]. Obviously, the frequency of each amino acid in a X position is influenced by the composition of the premixed amino acid mixture, the chemistry applied for coupling as well as by the nature of the solid support. The synthetic peptide libraries have been analyzed by electrospray mass spectrometry and pool sequencing [13, 14]. Due to their stoichiometry, peptides from extremely diverse combinatorial collections are only partially accessible for biological assays [12].

### 11.1.3 Peptide Libraries for the Investigation of TAP

The TAP is a heterodimeric complex formed by TAP1 and TAP2, each containing one nucleotide binding and one transmembrane domain [3, 15]. TAP is localized in the ER-membrane [16] and supplies peptides to many different class I molecules [17]. Totally degenerate peptide libraries were investigated in a bimolecular assay for the determination of equilibrium binding affinity constants  $K_D$  [18] and the length selectivity of human TAP was defined for peptides from 8 to 16 amino acids. These results were confirmed by translocation studies using peptides with radio-iodination-sites [19].

In comparison to a reporter peptide, the affinity of human TAP for a totally degenerate nonapeptide library was found to be 17-fold lower [20]. Single amino acid substitutions in the high-affinity reporter peptides resulted in only two- to three-fold differences in transport rates, but in up to 80-fold differences as revealed by the use of synthetic peptide libraries.

A complete map of the fine specificity of human TAP was obtained by testing nonapeptide sublibraries [20]. The most pronounced effects for side-chain substitutions were found for the C-terminus where aromatic, aliphatic, and positively charged residues were preferred (Table 11.1). The amino-terminal region (positions one to three) also proved to be important.

These results verified the results obtained with totally randomized undecapeptide libraries and pentadecapeptide libraries characterized by one position occupied by D-amino acids [21]. Interestingly, TAP can accommodate peptides with bulky or non natural side chains, and even branched peptides [21, 22].

**Table 11.1.** Ranking of amino acids unfavorable (a) or favorable (b) for peptide recognition by human transporter associated with antigen processing (TAP), as identified by screening of O/X<sub>n</sub> peptide libraries [20].

Position	1	2	3	4	5	6	7	8	9
(a)	D	P	D	N	T		D		D
	E		E				E		E
	F		G						G
									N
									S
(b)	R	R	W			R	I		V
	K	Q	Y						Y
	N	I							R
									L
									F

The C-terminal amino acid residues preferred by TAP match with those generated by the proteolytic activities of the proteasome [23], as well as with C-terminal anchors used by most human class I molecules (database SYFPEITHI: [www.medizin.uni-tuebingen.de/sfb510/index.html](http://www.medizin.uni-tuebingen.de/sfb510/index.html)) [24].

#### 11.1.4 MHC Class I and Class II Molecules

MHC class I or MHC class II (in humans, these are termed HLA) peptide complexes are recognized on the cell surface by the T-cell receptor (TCR) of a T cell. MHC molecules are optimized for binding a wide spectrum of different peptides. The MHC-bound peptides are located in a groove (MHC class I) or a cleft (MHC class II) formed by a  $\beta$ -sheet platform and two  $\alpha$ -helices. The peptide-binding groove of class I molecules is closed at both ends and the conformational space is suited for binding of octa- to undeca-peptides [25]. The cleft formed by MHC class II molecules has open ends for binding of longer peptides protruding out from the cleft [26]. A large number of highly polymorphic different antigens (natural peptide library) can be presented at the cell surface by MHC molecules. From the several hundred alleles existing in a whole population, individuals carry up to six different MHC class I and six different MHC class II molecules. This polymorphism plays a key role in allograft rejection and for allele-specific susceptibility for certain autoimmune diseases and cancer; this highlights the importance of understanding MHC-peptide interaction for targeted therapeutic intervention.

MHC class I molecules are folded by support of a variety of chaperones, *e.g.*, calnexin, calreticulin, tapasin. The class II pathway has to exclude the presentation of peptides of intracellular origin. Therefore, in the ER MHC class II molecules remain associated with the invariant chain Ii, where the CLIP region of Ii occupies the peptide-binding cleft [27]. CLIP is removed from the MHC in the MHC class II compartment MIIC [28]. MHC class II molecules adopt a conformational change upon loading with peptides that allows the sorting to the cell surface.

#### 11.1.4.1 Peptide Ligands of MHC Class I Molecules

Strategies for the determination of natural T-cell epitopes include mostly the identification of epitopes from source proteins by molecular genetic techniques. Truncation mutants of the original genes were analyzed in bioassays and further characterized with overlapping peptides [29]. Alternatively, natural T-cell epitopes are eluted from cell lines that restrict the T-cell response. The eluted peptides are purified and sequenced by Edman degradation [30].

These approaches are time-consuming and limited by the amount of biological materials. A fast method to define recognition patterns of MHC molecules has been introduced by screening synthetic combinatorial peptide libraries [31].

The synthetic generation of millions of peptides can supplement the search for natural epitopes. Synthetic, as well as natural, epitopes can be modified to obtain potential drugs with patient specific properties [32].

The rules for peptide selection by MHC class I molecules were defined by the characterization of peptide mixtures extracted from class I complexes [33], and furthermore, by the effects of collections of different peptides on binding to MHC molecules and by the response of cytotoxic T cells on these presented peptides [34]. Crystal structure analyses of defined MHC-peptide and MHC-peptide-T-cell receptor (TCR) complexes [35–38] gave detailed information on the molecular interaction between peptides and MHC proteins. MHC class I ligands are mainly octa- or nonapeptides and obey allele-specific sequence motifs carrying prominent anchor residues [24]. The peptide-binding groove offers specific pockets to interact with these anchor residues [35].

Complex synthetic peptide libraries have been used for further analysis of MHC-peptide interaction, especially to define the contributions of primary, secondary, and non anchor residues in class I ligands and to investigate the response of cytotoxic T cells. To study peptide binding to the murine MHC class I molecules H-2K<sup>b</sup> and H-2L<sup>d</sup>, and to define the contribution of each individual amino acid in a given sequence position O, 152 octapeptide libraries O/X<sub>7</sub>, each representing a collection of about  $900 \times 10^6$  individual peptides, were investigated. Thereby class I-binding was defined by a standard stabilization assay for in-vitro detection of peptide-MHC complexes. The results were independent from restrictions by intracellularly preselected, naturally processed peptides. Defined amino acids in the eight positions of the library O/X<sub>7</sub> were classified due to their contribution for MHC I-binding (Table 11.2) and due to the recognition of the resulting complexes by cytotoxic lymphocytes (CTL) [31, 39, 40].

Prominent primary anchor residues in anchor positions determined by pool sequencing (e.g., positions 5 and 8 for H-2K<sup>b</sup>, positions 2 and 8 for H-2L<sup>d</sup>) were ascertained by screening octapeptide libraries [39]. Unfavorable effects on binding can be induced also by amino acids in non anchor positions. Conformational analysis of several individual peptides bound to MHC class I molecules showed that the position of the peptide backbone in the binding groove, as well as the orientation of side chains in other sequence positions, is influenced by individual amino acid side chains [36]. These interactions strongly determine the accessibility of the peptide on the MHC I surface and thus the response of the CTL effector cells. Anchor residues promote efficient binding to MHC I, but notably suppress the CTL response to the complex peptide libraries [39].

**Table 11.2.** Ranking of amino acids according to results from stabilization assays of O/X<sub>7</sub> libraries on MHC H-2K<sup>b</sup> molecules [31]. Amino acids O are classified for their influence on stabilization (destabilizing: upper part (a); stabilizing: lower part (b)). Motif amino acids identified by sequence analysis of natural peptide libraries [33] are underlined.

Position	1	2	3	4	5	6	7	8
(a)	D	D	E	G	D	D	D	N
	P	E	D		E	W	W	K
	W	K	K		K	<u>E</u>		D
	N	G	G		I			G
	H	F	A		N			R
	Y	P	S		G			Y
	E	T	Q		M			W
	R				R			P
					Q			S
								E
(b)	<u>Δ</u>	A	<u>P</u>	L	V	A	L	T
	V	I	L	<u>K</u>	<u>E</u>	P	<u>K</u>	F
	<u>S</u>	W	<u>I</u>	P	<u>Y</u>	R	M	A
	I	S	F	Q			H	I
		V	<u>Y</u>	V			S	<u>M</u>
		<u>N</u>					<u>Q</u>	<u>V</u>
							R	<u>L</u>
							<u>N</u>	
							P	

Nonapeptide libraries have been applied to determine the peptide-binding motifs of three rat MHC class I molecules which were incubated with a completely randomized sequence. Proper folding and peptide loading in the presence of the X<sub>6</sub> library, isolation of peptides from the MHC-peptide complex, and pool sequencing revealed an almost identical motif as obtained by classical procedures with acid-eluted peptides [41]. The motif obtained with the X<sub>6</sub> library is not strongly affected by cellular processing and transport and thus, it is similar to the stabilization experiments using the O/X<sub>8</sub> libraries. These experiments can serve as a measure for the conformational space offered by the binding groove of the respective MHC molecule.

#### 11.1.4.2 Peptide Ligands of MHC Class II Molecules

Compared to MHC class I molecules, class II proteins bind longer naturally processed peptides with 10–25 amino acids and thereby no apparent restriction on peptide length has been observed [42, 43]. Although they also show allele-specific motifs [30, 44], numerous different peptides are capable of binding to a MHC molecule encoded by a defined haplotype, which can be considered as a natural peptide library [45]. They bind by core regions of nine amino acid residues carrying several allele-specific anchor positions. The understanding of the molecular interactions which determine peptide binding to MHC class I and class II molecules is of great importance for the prediction and identifi-

cation of T-cell epitopes, especially for the design of synthetic vaccines and for understanding immunopathology. The fine specificity of peptide binding to MHC class II molecules has been analyzed by using different approaches including binding assays with substitution analogues of known ligands [46]. Isolation and subsequent sequencing of class II-associated single self-peptides [47] or pool sequencing of isolated natural ligands revealed allele-specific motifs [30]. Quite different to these methods is the search for MHC ligands using M13 phage display libraries [44, 48].

Crystal structure analysis of the human MHC class II molecule HLA-DR1, complexed with a tridecapeptide from influenza virus, shows several interaction sites or pockets within the peptide-binding cleft of HLA-DR1, five of which accommodate hydrophobic side chains of the bound influenza virus peptide. Many of the residues forming these pockets are highly polymorphic. This polymorphism is thought to be responsible for the different peptide specificities of different class II proteins. There are 11 core residues of the influenza virus peptide interacting with the DR1 molecule [49].

### 11.1.5 T-Cell Response

More than  $10^7$  different antigens can be recognized specifically by different TCR molecules expressed on different T cells [50]. The accessory molecules CD8 or CD4 allow recognition of peptide-MHC I or peptide-MHC II complexes, respectively. CD8<sup>+</sup> cytotoxic T cells recognize target cells presenting non self peptides in complex with MHC class I molecules, and induce cell lysis by secretion of pore-forming proteins [51] or by inducing apoptosis. CD4<sup>+</sup> helper T cells recognize non self peptides bound to class II molecules followed by lymphokine release [52].

#### 11.1.5.1 Recognition of MHC Class I-Bound Peptides by TCR

A prerequisite for CTL-mediated immune response is the formation of the MHC-class I-peptide complex and subsequent recognition by the T-cell repertoire, which can be analyzed in cell lysis assays with <sup>51</sup>Cr-loaded target cells. Chromium release is a measure for peptide-induced cell lysis by CTL, and indicates the potency of the peptide to serve as an allele-specific epitope. A synthetic epitope has been identified with the peptide library approach to elucidate the molecular basis for the observed cross-recognition of two ligands by a single receptor [53].

Screening of 152 O/X<sub>7</sub> peptide libraries on cytolysis of target cells by several CTL clones showed clone-specific Recognition Patterns which allow the deduction of synthetic T-cell epitopes effective in the concentrations also typical for naturally occurring epitopes [54, 55]. The degree of degeneracy is characteristic for a certain CTL clone and is also represented in the Recognition Pattern of the peptide library.

The approach succeeded in the construction of functional mimics of tumor epitopes. One peptide was identified that induces a CTL response *in vivo* and thus represents the functional mimics of a CTL epitope. The value of peptide libraries for the identifica-

tion of synthetic epitopes as potential tumor vaccines was practically confirmed by adoptive transfer of a CTL line from mice immunized with the epitope. The transferred CTL inhibited the growth of an ascetic tumor *in vivo* [32].

Combinatorial peptide libraries, either completely random or characterized by one or several defined positions, are useful tools for the identification of the critical features of MHC class I binding peptides and of natural and synthetic epitopes. The complete Recognition Pattern of an O/X<sub>7</sub> library with 152 individual peptide mixtures represents the influence of amino acid residues on mediating contact to the MHC class I molecules and of residues crucial for recognition by the TCR [54, 55].

#### 11.1.5.2 Recognition of MHC Class II-Bound Peptides by TCR

Protein antigens are taken up by APC and are degraded by cellular proteases, especially in late endosomes and phagosomes. However, little is still known about the enzymes responsible for the processing of antigens, including their specificity. Recently, a 15-residue synthetic peptide library and a native protein were used as substrates to study cleavage specificity of the aspartic proteases cathepsin D and E [56]. By combining automated N-terminal sequencing with a new method for C-terminal sequencing of peptide pools generated by these endopeptidases, the cleavage motif of cathepsin E was confirmed and extended. Pool-sequencing analysis provided evidence that both enzymes share the same cleavage motif with a hydrophobic amino acid in positions P1 and P1' [56].

Susceptibility to a series of autoimmune diseases is strongly associated with particular HLA class II alleles, and identification of antigenic epitopes of T-cell clones is critical for understanding the etiology of these diseases. Peptide library-based methods which include all possible peptide sequences turned out to be a powerful tool for the detection of cross-reactive antigenic peptides recognized by CD4<sup>+</sup> T cells. In a recent approach, the number of peptides per mixture was reduced by utilizing the known binding motif of peptides for HLA-DRB1\*0405 [57]. Thus, a strong proliferative response of unprimed peripheral blood lymphocytes (PBL) from HLA-DRB1\*0405-positive donors was measured, and the library detected antigenic peptides that activated three antigen-specific T-cell lines restricted by DRB1\*0405, with different specificities. This motif-based approach is useful for monitoring T cells in large heterogeneous T-cell populations and for the identification of the mimic peptide epitopes of T-cell lines and clones [57]. In a similar approach for identification of CD4<sup>+</sup> T-cell epitopes, peptide libraries are also designed to bind to the HLA class II restriction molecule of the T-cell clone of interest [58]. Screening is based on three selection rounds using partial release of 14-mer peptides from synthesis beads and subsequent sequencing of the remaining peptide attached to the bead. With this approach, two peptides were identified that stimulated a beta cell-reactive T-cell clone isolated from an insulin-dependent diabetes mellitus patient. This system should be capable of determining epitopes for autoreactive CD4<sup>+</sup> T-cell clones with previously unknown peptide specificity [58].

A major hindrance for transplantation arises from the high frequency of alloreactive T cells, although the molecular basis of alloreactivity remains elusive. Recently, a combinatorial peptide library was used to study I-E<sup>P</sup>-alloreactivity of a hemoglobin (64-76)/I-E<sup>K</sup>-



specific murine T cell clone [59], and a highly stimulatory alloepitope mimic was identified. Moreover, combinatorial peptide libraries were applied as immunogens [60] and allele-specific libraries were used for identification of T-cell mimotopes [61].

## 11.2 Methods

### 11.2.1 Synthesis of Peptides and Peptide Libraries

The synthetic peptides and the peptide libraries and sublibraries were prepared by solid-phase peptide synthesis using Fmoc/tBu chemistry. Solvents, amino acids and coupling reagents were handled by a robot for organic chemistry (Syro, MultiSynTech, Bochum, Germany). The syntheses were carried out on Rink amide MBHA polystyrene resin or on Wang-polystyrene resin (Novabiochem, Switzerland). Fmoc-L-amino acids were used with the following side chain-protecting groups: *tert*.-butyl ethers for Ser, Thr, Tyr; *tert*.-butyl esters for Asp and Glu; trityl for His, Asn and Gln; *tert*.-butyloxycarbonyl for Lys and Trp; 2,2,5,7,8-pentamethylchroman-6-sulfonyl (Pmc) for Arg.

For coupling of amino acids in degenerate sequence positions (X), double couplings were performed with equimolar mixtures of Fmoc-L-amino acids and Fmoc-Gly (cysteine was replaced by  $\alpha$ -aminobutyric acid) which were used in an equimolar ratio with respect to the coupling sites on the resins. The resin was distributed in 30 mg aliquots (15  $\mu$ mol) to filter tubes, which were positioned in the format of a microtiter plate on valve blocks. Fmoc deprotections were carried out twice for 7 min each, with 30% piperidine in dimethylformamide (DMF) (220  $\mu$ l). Nine washing steps were carried out with DMF (300  $\mu$ l). Coupling reagent diisopropylcarbodiimide [1.5 M in DMF:dichloromethane (DCM) (1:2, v/v), 50  $\mu$ l] and Fmoc-amino acids (200  $\mu$ l) or Fmoc-amino acid mixtures (200  $\mu$ l) were distributed to the reaction vessels. Fmoc-amino acids (0.5 M) were dissolved with HOBt (0.5 M) in DMF. These solutions were used for coupling of defined positions. A premixed equimolar mixture of Fmoc-amino acids [total 0.075 M in DMF:DCM (1:7, v/v), 200  $\mu$ l] was distributed for coupling of X positions. Double couplings (3 h each) were carried out in open tubes. After 2 h coupling time, diisopropylethylamine (DIEA) [1 M in DMF:DCM (1:1, v/v), 20  $\mu$ l] was added. Coupling reagents were filtered off and the resins were washed three times with DMF. Acetylation was carried out with a 10-fold excess of acetic acid anhydride and diisopropylethylamine for 30 min.

The peptides and peptide mixtures were cleaved from the resins and the side chains were deprotected with trifluoroacetic acid:phenol:ethanedithiol:thioanisole:water (96:2:1:2:1, v/w/v/v/v) (1 ml) within 3 h. The products were filtered from the resins, which were washed once with acetic acid (0.3 ml), and were precipitated at  $-20^{\circ}\text{C}$  by the addition of cold n-heptane:diethylether (1:1, v/v) (5 ml). The precipitates were washed twice by sonication with n-heptane:diethylether (1:1, v/v) and were lyophilized from acetic acid:water:*tert*.-butyl alcohol (1:10:50, v/v/v), yielding individual peptides, the completely randomized peptide amide libraries  $X_n$  and 220 undecapeptide amide sublibraries  $O/X_{10}\text{-NH}_2$ . The biotinylated peptide sequence (see below) was obtained by elongating the ApoB 100 peptide 2877–2894 with two spacer amino acids (Ahx), followed by biotin using the coupling method described above.

### 11.2.2 Analytical Characterization of Peptide Libraries

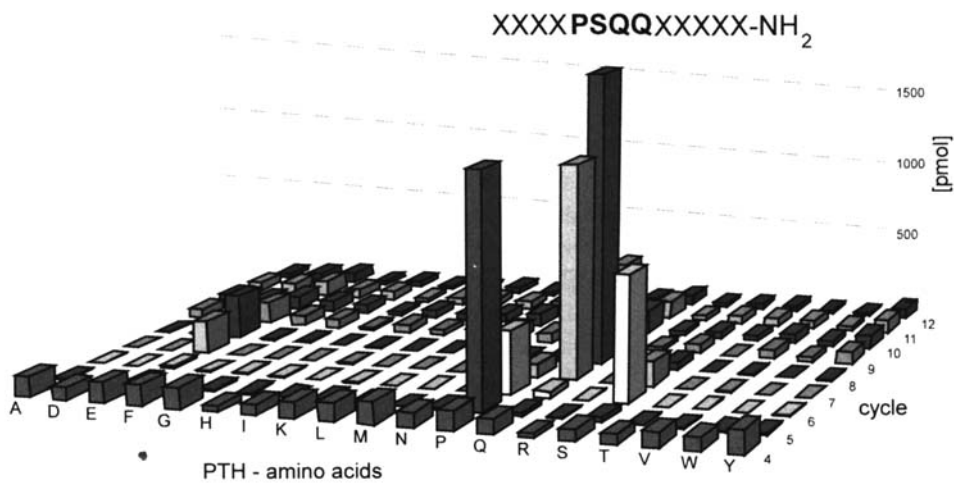
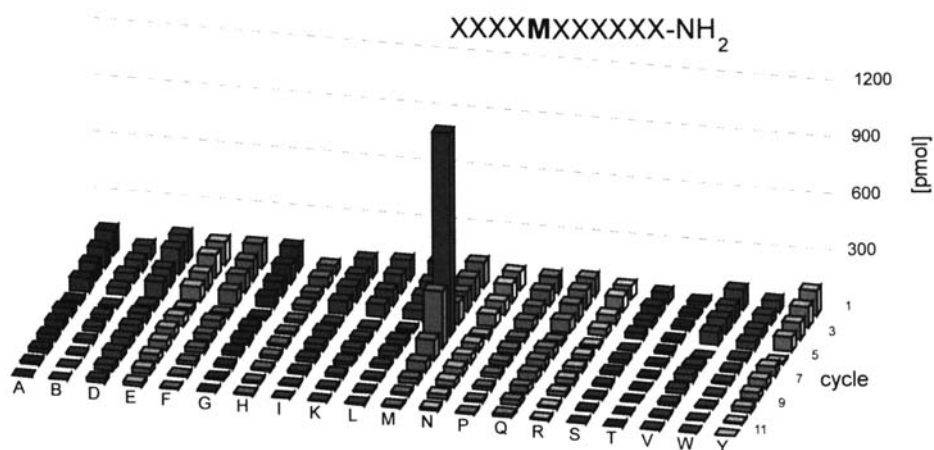
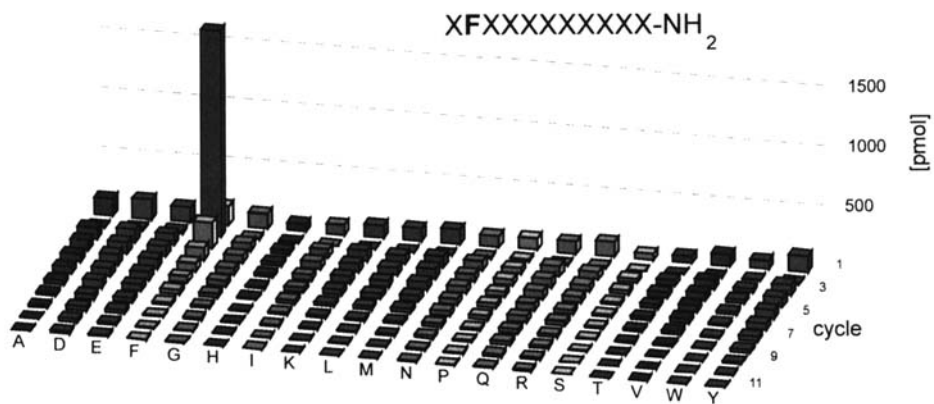
The identity of the defined peptides was confirmed by electrospray mass spectrometry [13, 62] and the purity of defined peptides was higher than 80 % as determined by HPLC. The amino acid composition of selected peptide libraries and of the sublibraries was determined by pool sequencing [14] (Fig. 11.1) and amino acid analysis. Deviations from equimolar representation of the amino acids in randomized sequence positions were found to be within the error limits of the analytical methods [12]. Electrospray mass spectrometry for pattern analysis of peptide libraries differing in the position of the defined amino acid was carried out with a triple quadrupole mass spectrometer (Fisons, Manchester, UK) (Fig. 11.2).

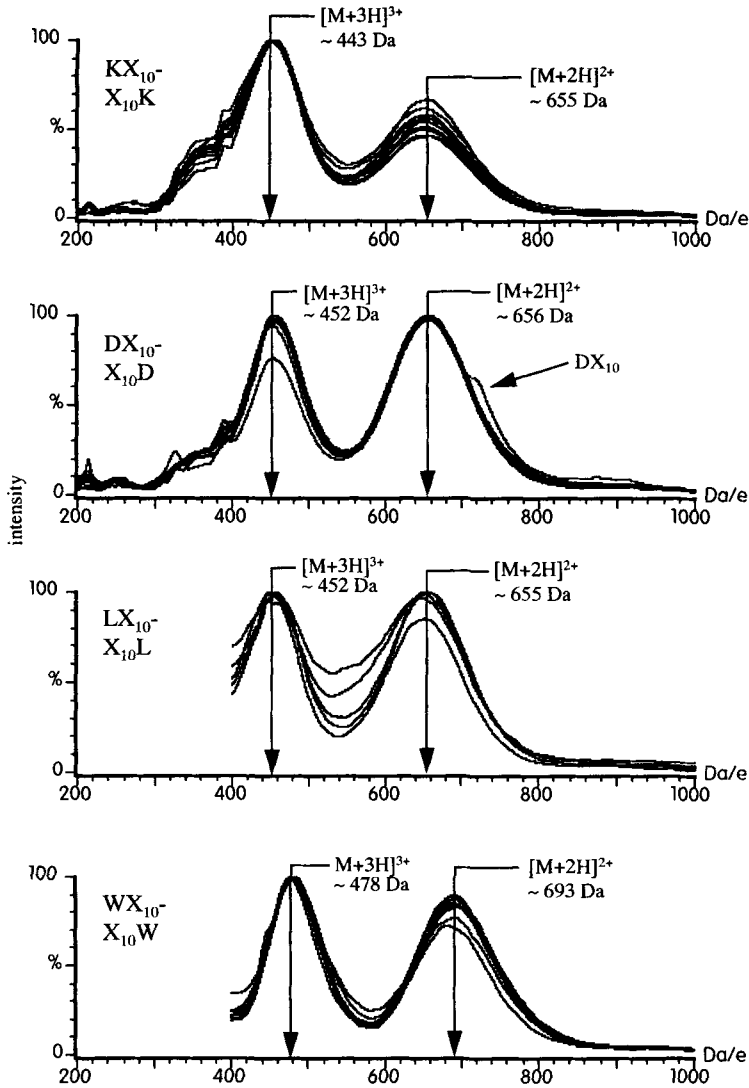
### 11.2.3 HPSEC-competition assay

High-performance size-exclusion chromatography (HPSEC) was used for competition studies. Solubilized HLA-DR1 (0.13  $\mu\text{M}$ ) was incubated for 48 h at 37 °C with the N-terminally 7-amino-4-methylcoumarin-3-acetic acid (AMCA)-labeled IM-(19-31)-peptide dissolved in 150 mM sodium phosphate, pH 5.5, containing 15 % (v/v) acetonitrile, 0.1 % (w/v) Zwittergent-12 (Calbiochem) and a cocktail of protease inhibitors (0.2 mM PMSF, 5  $\mu\text{M}$  leupeptin, 10  $\mu\text{M}$  pepstatin, and 1  $\mu\text{M}$  chymostatin). Competition assays were performed in a 1.5  $\mu\text{M}$  solution of AMCA-peptide; as competitors, different peptides, the peptide library or peptide sublibraries were added in concentrations ranging from 10.3  $\mu\text{M}$  to 41.3  $\mu\text{M}$ . All samples were analyzed on a Pharmacia Superdex 75 HR 5/20 high-performance gel filtration column, essentially as described previously [63]. The column was operated at a flow rate of 0.4 ml/min (250 p.s.i.) using the HPSEC buffer, pH 6.0. The effluent passed through a Merck fluorescence spectrophotometer (350/450 nm) and a Merck UV detector (214 nm) set up in series. Fluorescence and UV signals eluting with HLA-DR1 dimers were recorded by a model D-2500 integrator (Merck-Hitachi).

### 11.2.4 Isolation of HLA Class II Molecules

Epstein–Barr virus (EBV)-transformed homozygous human B-cell lines were used as a source for isolating HLA class II molecules. In the case of isolation of HLA-DR1, WT-100 cell pellets were lysed by NP-40, and DR1 was isolated from homogenates by affinity chromatography with the monoclonal antibody L243 essentially as described [64, 65]. The purity of the preparation was analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), HPSEC and Western blotting.





**Figure 11.2.** Analysis of undeca-peptide libraries  $O/X_{10}\text{-NH}_2$  by electrospray mass spectrometry. Spectra of four groups of libraries are superimposed. Each group is expected to show identical mass patterns. The libraries are characterized by one defined amino acid (K, D, L, W) in different sequence positions. The positively charged libraries  $K/X_{10}\text{-NH}_2$  show, as expected, striking intensities of triply charged ions. Additional experiments are necessary to explain the shift in the mass pattern for  $DX_{10}\text{-NH}_2$  characterized by the negatively charged amino acid D in the N-terminal position.

◀ **Figure 11. 1.** Results from pool sequencing (477 A, Applied Biosystems, Weiterstadt) of undeca- and tri-decapeptide libraries with one or four defined amino acid residues in different sequence positions. Bars indicate detected amounts of PTH-amino acids in each cycle. Edman degradation was carried out as described [14].

## 11.3 Results and Discussion

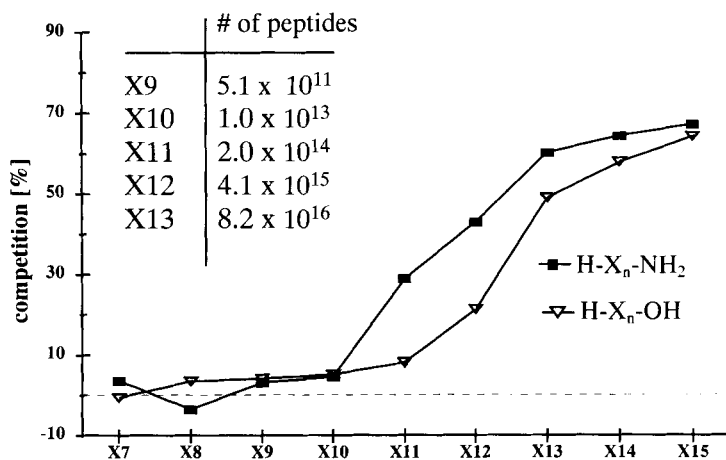
### 11.3.1 Optimal Length of Peptides for MHC Class II Interaction

Completely randomized synthetic peptide amide libraries from 7 to 15 amino acids ( $X_7$ – $X_{15}$ ) were synthesized and applied to elucidate an ideal peptide length applicable in competition studies with HLA-DR1 (Fig. 11.3). A minimal length of peptide amides comprising 11 amino acids was found to be promising in assays with DR1 molecules [66].

### 11.3.2 Activity Pattern Describing DR1-Peptide Interaction

An undecapeptide amide library was designed for the elucidation of ‘Activity Patterns’ of amino acids for the interaction with MHC class II molecules. This consists of a completely randomized  $X_{11}$  library and 220 sublibraries (11 positions  $\times$  20 amino acids). Sublibraries contained 10 randomized positions (X) and one defined amino acid (O) moving across the 11 sequence positions. In the X-positions and the O-positions, 19 proteinogenic L-amino acids were used and cysteine was replaced by its chemical analogue  $\alpha$ -aminobutyric acid (B). Each of the 220 O/ $X_{10}$  sublibraries comprised  $20^{10}$  ( $1 \times 10^{13}$ ) different peptides. The  $X_{11}$  library and the 220 sublibraries were employed in competition assays as described above [66] and competition C of a given sublibrary was compared to that of the  $X_{11}$  library resulting in a ‘relative competition’ rel C ( $\text{rel C} = C_{\text{sublib}}/C_{X_{11}}$ ).

The activity of a sublibrary is only governed by the amino acid residue in the defined sequence position, whereas the remaining 10 randomized positions provide an averaged contribution to binding. Thus, by summarizing the rel C values of the 220 sublibraries in an ‘Activity Pattern’ [66], the contribution of any amino acid side chain in any sequence



**Figure 11.3.** Influence of peptide length of completely randomized peptide libraries on binding to HLA-DR1, shown for peptide amides and free acids.

position to DR1-binding is directly obvious. Based on the crystal structure analysis of the HLA-DR1 protein complexed with the influenza hemagglutinin (306–318)–peptide [49], amino acid side chains in sequence positions 2, 5, 7, 8, 10 and 11 of the library are expected to interact with pockets in the DR1-binding cleft. In contrast, side chains in positions 1, 3, 4, 6 and 9 should be orientated away from the binding site. In positions interacting with the DR1 molecule we identified residues strongly contributing to MHC-binding as well as residues with an unfavorable effect on the interaction (Table 11.3). In all other positions, however, defined amino acid residues showed a more or less neutral influence on MHC-binding. In summary, in all sequence positions of undecapeptides a wide variety of side chains is accepted by the DR1 protein, but especially the allele-specific pockets in the binding cleft show favorable interactions to several amino acid residues. This clearly demonstrates the strongly degenerated peptide binding to class II molecules.

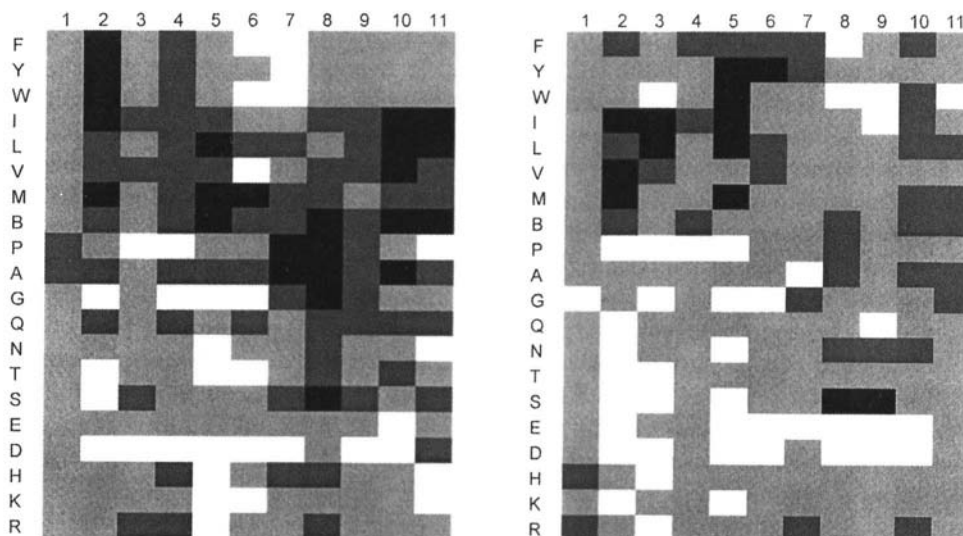
A very clear presentation of this two-dimensional data collection describing peptide binding to DR1 and to DRB1\*1501 is given in Figure 11.4. Relative competition-values of the 220 sublibraries are indicated by squares of different black intensities resulting in a ‘fingerprint’ of the studied class II molecule.

After evaluation of this approach, ‘Activity Patterns’ for several HLA class II haplotypes associated with autoimmune diseases have been generated. They serve currently as the basis for the search and identification of new autoantigens recognized by autoreactive T cells. Thus, peptide binding to the class II molecule HLA-DRB1\*1501 (DR2b), associated with multiple sclerosis, was also studied. (DR2b was isolated from the EBV-transformed homozygous human B lymphoblastoid cell line HTC-LAN, essentially as described for HLA-DR1.)

Our interpretation of results from competition studies DR2b and with peptide libraries allowed us to describe a new interesting phenomena, called translational invariance: e.g., sublibraries carrying aliphatic residues in position 2 showed comparable competitions as those with aliphatic side chains in position 3 (Fig. 11.4). Similar results were obtained for

**Table 11.3.** ‘Activity Pattern’ of an undecapeptide amide library. A total of 220 undecapeptide sublibraries O/X<sub>10</sub> each composed of ten degenerate positions and one defined position, were screened for binding to HLA-DR1. Unfavorable (a) and favorable (b) amino acids at every sequence position are indicated. Amino acids with intermediate contribution to binding are not shown; (B:  $\alpha$ -aminobutyric acid).

Position	1	2	3	4	5	6	7	8	9	10	11
(a)		D S G T	P D	P G D	D N K T H R G	W T D V K F G	W Y F D		D	E D	N P H K
(b)		Y M I F W			B L M	M	P A	S B P A G		I A L V B	L I B



**Figure 11.4.** Schematic representation of allele-specific 'Activity Patterns' of combinatorial undecapeptide libraries for two DR-molecules.

aromatic residues in position 5 and 6, as well as for N and S in positions 8 and 9. Thus, positive binding contribution of a defined residue O in position  $i$  is often also found in position  $i + 1$  (Fig. 11.4). This can be explained by the facts that the class II binding cleft is open at both ends, and that the core region of a peptide binding to class II molecules preferably consists of nine rather than 11 amino acids. Therefore, undecapeptides might protrude out at the ends of the binding cleft. This might be reflected by an intrinsic inaccuracy in the competition data obtained from undecapeptide amide libraries characterized by only one defined amino acid.

In parallel, an efficient screening procedure for the identification of high-affinity HLA-DRB1\*0301 (DR 17)-ligands has been established to characterize peptide-binding specificity for this haplotype associated with the autoimmune disease myasthenia gravis (HLA-DRB1\*0301 protein was isolated from the human cell hybrid transfectant T2DR3 [67]). A total of 209 synthetic undecapeptide amide sublibraries  $O/X_{10}\text{-NH}_2$  representing collections of  $20^{10}$  individual peptides were applied in a competition enzyme-linked immunosorbent assay (ELISA) using the biotinylated natural ligand ApoB(2877-2894) (Table 11.4).

In this study, identified favorable amino acids in DRB1\*0301 ligands and positively ranked amino acid residues were listed as a database for random combination of new undecapeptide ligands (Table 11.4). This novel approach for the design of ligands was introduced to compensate for the inaccuracy induced by the translational invariance of amino acids in longer  $O/X_n$  peptide libraries characterized by one defined amino acid. An alignment of newly defined ligands with the highest affinities allows for the elucidation of amino acids essential in peptide sequences binding to DRB1\*0301 [68]. A second, more

**Table 11.4.** Activity Pattern of an undecapeptide library defined by competition experiments. Influence of amino acid residues unfavorable (a) or favorable (b) on DRB1\*0301 binding correlated with its position in peptide sublibraries [68]. Residue-position combinations were ranked as active or inactive (indifferent ranked amino acids are not shown).

Position	1	2	3	4	5	6	7	8	9	10	11
(a)	F K M	K L I M	D G K P R S T	D E G H K N P T	F G H K N P Q T	G H P Q R S	A N S	A G Q S	G H I L Q W	H P	E G H
(b)	P A E G H S T W	Q G A E H T W	I L M W F V Y	I W V	D V I L S	D L M W	E H I K P Q V	N E H K P R W	Y A M P	S A F K M	R A I N P Q S T V Y

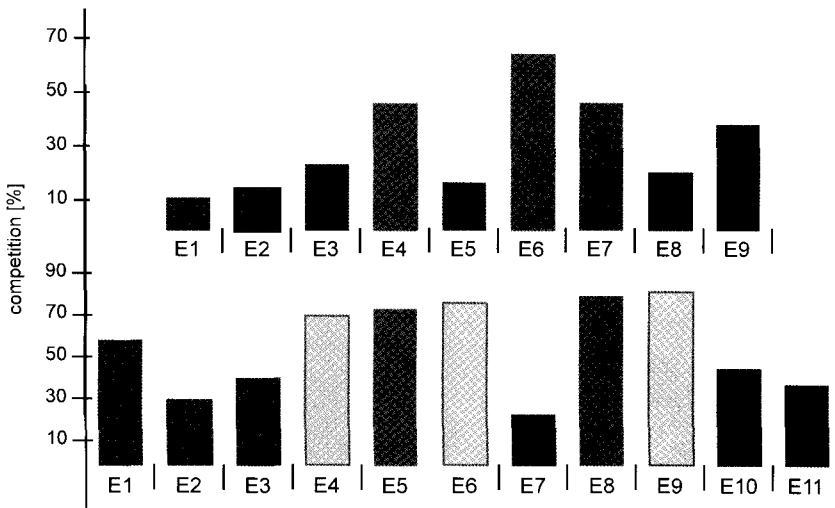
directly deduced, set of 24 peptides was obtained by combination of the most favorable residues in each position. All individual peptides were investigated in the competition assay. The most active HLA DRB1\*0301 ligands were obtained by random selection of favorable amino acids and a significant number of them (Table 11.5) showed improved affinity compared to the model ligand  $\alpha$ AChR 310–325 [68].

In order to avoid translational invariance when studying peptide binding to class II molecules with peptide libraries, two approaches would be possible. First, the incorporation of more than one defined residue in the undecapeptide library, *e.g.*, a common hydrophobic residue in the first anchor position of DR-ligands, would prevent translational invariance. However, universal applicability of the peptide library approach would be lost.

**Table 11.5.** Determination of the  $IC_{50}$  values of five newly designed peptides for binding to DRB1\*0301. The sequences are based on the 'Activity Pattern' and were investigated in comparison to the standard binder  $\alpha$ AChR 310–325. Amino acids in anchor positions are in bold text.

Peptide	Sequence											$IC_{50}(\mu M)$			
1				W	Q	<b>L</b>	I	V	<b>D</b>	V	<b>K</b>	A	S	<b>A</b>	5
2				H	T	<b>L</b>	V	V	<b>D</b>	F	<b>R</b>	P	M	<b>A</b>	7
3				H	Q	<b>M</b>	W	Y	<b>D</b>	E	<b>R</b>	Y	T	<b>R</b>	11
4				E	W	<b>L</b>	I	D	<b>D</b>	K	<b>K</b>	Y	S	<b>Q</b>	13
5				W	A	<b>L</b>	I	V	<b>D</b>	P	<b>K</b>	A	A	<b>A</b>	14
$\alpha$ AChR 310–25	N	W	V	R	K	<b>V</b>	F	I	<b>D</b>	T	<b>I</b>	P	N	<b>I</b>	M F 17





**Figure 11.5.** HLA-DQ(A1\*0501, B1\*0201): Comparison of the results from competition experiments carried out with nonapeptide sublibraries Ac-E/X<sub>8</sub>-NH<sub>2</sub> and undecapeptide sublibraries H-E/X<sub>10</sub>-NH<sub>2</sub>.

because additional defined amino acids have to be introduced with respect to the corresponding class II allele. A second possibility is to reduce peptide length, *e.g.*, to nonapeptides covering exactly the length of the peptide core interacting with the binding cleft. Especially N-terminally acetylated nonapeptide amides will be capable to form all hydrogen bonds from the peptide's backbone to side chains of the HLA molecule.

In competition assays with an acetylated nonapeptide amide library, Activity Patterns for two DQ molecules were generated and used to predict DQ(A1\*0501/B1\*0201) (= DQ $\alpha$ 5)-specific ligands from tissue transglutaminase as potential T-cell epitopes in celiac disease [69]. As glutamic acid can serve as anchor amino acid in two positions of DQ $\alpha$ 5-ligands, H-E/X<sub>10</sub>-NH<sub>2</sub> libraries were compared to Ac-E/X<sub>8</sub>-NH<sub>2</sub>-mixtures and, as expected, the phenomenon described as translational invariance [68] was observed only in experiments utilizing the undecapeptide libraries (Fig. 11.5). Thus, depending on the MHC-haplotype to be screened, acetylated nonapeptide amides could be more suitable for studying class II-peptide binding. However, a much higher excess of competitors is necessary and optimal binding conditions have to be evaluated before carefully.

### 11.3.2.1 Tolerance to Amino Acid Variations in a HLA Class II Ligand

Tolerance to amino acid variations for binding of undecapeptides to HLA DR1 molecules was determined by calculating the sum of the absolute values of the  $\ln(\text{relC})$  for all 20 O/X<sub>10</sub> sublibraries of one sequence position. For sequence positions that allow close contact of the amino acids to the DR1 molecule, low tolerance with respect to biological activity to amino acid variations was expected. In contrast, amino acids variations should be more tolerated in positions projecting away from the binding cleft. The X-ray structure

of an influenza virus peptide in complex with DR1 shows the orientation of the central 11 residues of the viral peptide [49]. This orientation was confirmed by the results of the studies with undecapeptide libraries: sequence positions 2, 5, 7, 8, 10 and 11 show low tolerance to variations of amino acids. Positions 1, 3 and 9 are most tolerant, whereas positions 4 and 6 show intermediate tolerance [66].

#### 11.3.2.2 Competition of Peptides Deduced from the 'Activity Pattern' of the Undecapeptide Library

Screening results from 220 sublibraries were clustered to an 'Activity Pattern' of the peptide library. Six sequence positions have strong influence on DR1-binding (positions 2, 5, 7, 8, 10, 11) (Fig. 11.4). In these positions the most potent amino acids (relative competition >2.80) as well as unfavorable amino acids were selected to design new undecapeptide amides with expected high or low DR1-affinity. In positions 1, 3, 4, 6 and 9, which are more tolerant with respect to amino acid exchanges, one amino acid with a relative competition value near 1 was introduced.

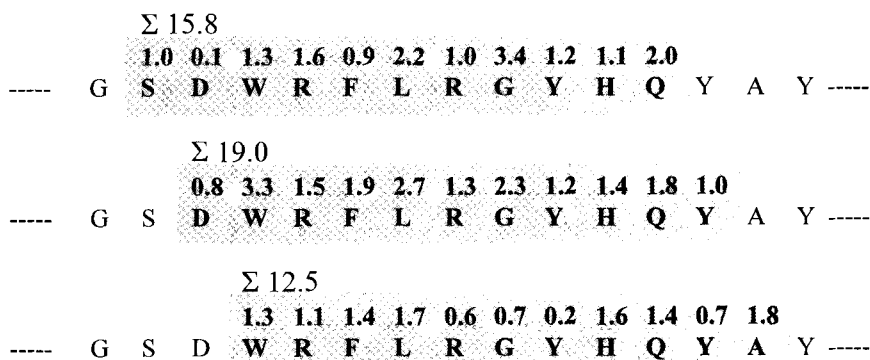
All possible combinations within the expected favorable and unfavorable amino acids were made, and two sets of peptides were synthesized: one set comprising 27 undecapeptide amides with sequences expected to be highly competitive and a second set with 27 sequences expected to show no competition. Peptides showing a combination of favorable amino acids in the appropriate positions exhibited competitions in the range of 35% to 92%. In contrast, peptides from the unfavorable group showed no or very low competition ( $\leq 8.4\%$ ), despite being tested in a 4-fold higher concentration. Peptides with glycine in position 8 were consistently less effective competitors (35–55%) when compared with most other favorable peptides showing competition higher than 80%.

The newly designed ligand SFHTMSAAKLI-NH<sub>2</sub> was titrated and the IC<sub>50</sub> value of 1.2  $\mu$ M found to be about 60 times lower than the IC<sub>50</sub> value of the X<sub>11</sub> library [66]. In contrast, for peptide SSHTKSYDKDN-NH<sub>2</sub>, synthesized as negative control with unfavorable amino acids selected from the 'Activity Pattern', no competition could be measured, even at a concentration up to 900  $\mu$ M.

#### 11.3.2.3 Predictions of HLA Class II-Ligands and T-Cell Epitopes by the Algorithm 'Actipat'

The rules for predicting HLA class II-ligands and T-cell epitopes of CD4<sup>+</sup> T cells are limited and are mainly based on preferences for certain amino acid residues in certain positions of the ligands. It is apparent though, that binding can be influenced by the amino acids in all sequence positions of the peptide. Thus, allele-specific 'Activity Patterns' represent the ideal database for new prediction methods, because they reflect quantitatively the binding contribution of every residue in a ligand.

Initially, the DR1-specific 'Activity Pattern' of the undecapeptide amide library was used to predict DR1-ligands in various proteins. Thereby, protein sequences were covered by a shifted window of 11 amino acids and the DR1-binding features of the correspond-



**Figure 11.6.** Calculation of the binding affinity of undecapeptide amides to HLA class II molecules by the prediction algorithm 'Actipat'. The contribution of every amino acid residue according to its sequence position is indicated by the value taken from the allele-specific 'Activity Patterns' of the peptide library.

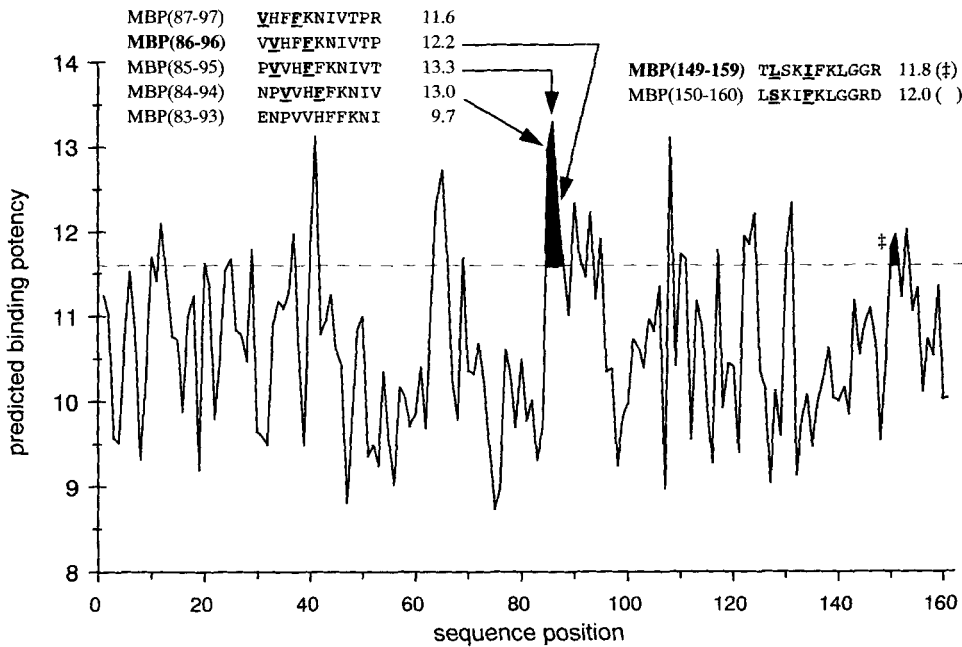
ing undecapeptide amide were calculated (Fig. 11.6). For this purpose, the values from the 'Activity Pattern' which correspond to the respective amino acid in the respective sequence position were summarized and assigned to the selected undecapeptide in the window. The obtained value was attached to the N-terminal amino acid of the selected undecapeptide. Shifting the window step-by-step by one amino acid through the protein sequence results in the predicted binding probability of undecapeptides to DR1 [70].

Comparing predicted ligands from the nucleoprotein of the measles virus with experimental data from competition assays, it transpired that all high-affinity ligands were predicted by the new approach. Furthermore, 86 % of all known T-cell epitopes and 83 % of DR1-ligands known from literature were predicted by 'Actipat'. This prediction efficacy could be confirmed also for other class II-alleles on the basis of the corresponding 'Activity Patterns' (DRB1\*1501, DQ(A1\*0501,B1\*0201), DQ(A1\*0201,B1\*0202)).

Compared to epitope predictions based on the presence of certain amino acid residues in anchor positions, this approach is superior with respect to its sensitivity. In a standard protein sequence, about 13–16 % of all undecapeptides are predicted, thus reducing the number of peptides to be tested tremendously when compared to standard epitope mapping (Fig. 11.7). However, among these predicted ligands about 50 % show marginal or no affinity to the respective class II allele. Taking into account the cleavage specificities of intracellular proteases [56] involved in antigen processing, will further improve the specificity of the predictions.

#### 11.3.2.4 Peptide Libraries and the Antigen Recognition of CD4<sup>+</sup> T Cells

The antigen recognition of autoreactive CD4<sup>+</sup> T-cell clones (TCC) was dissected by the response to the set of 220 undecapeptide sublibraries [71]. TCC 5G7 and TCC 5F6 were generated from PBL of multiple sclerosis patients by a limiting dilution split-well technique with MBP protein as described [72]. Both TCC 5G7 and TCC 5F6 are DR2b-

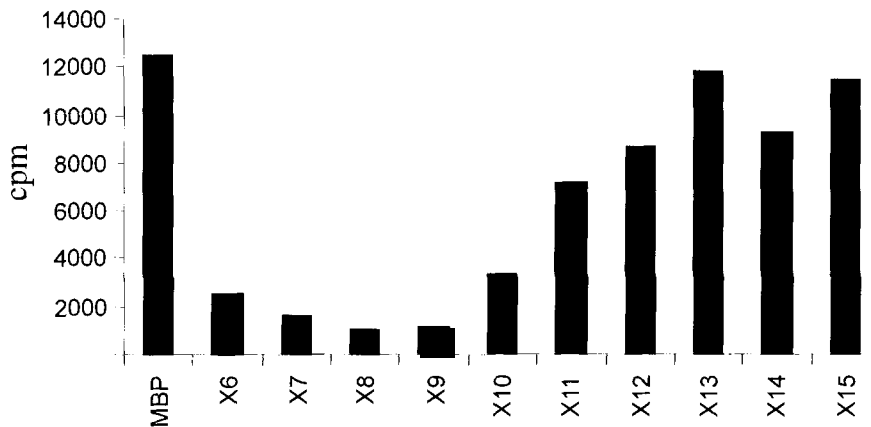


**Figure 11.7.** 'Actipat'-generated prediction profile for DRB1\*1501-binding peptides from the human myelin basic protein (MBP), the major autoantigen in multiple sclerosis.

(DRB1\*1501)-restricted and specific for the immunodominant peptide MBP(83-99) and MBP(86-96). T-cell proliferation was measured by standard [ $^3\text{H}$ ]thymidine incorporation as described [73]. TCC were added to wells of 96-well U-bottom plates containing  $5 \times 10^4$  irradiated (3000 Rad) PBL and various concentrations of peptides or peptide libraries. Cells were cultured for 48 h at  $37^\circ\text{C}$ , and during the last 6 h of culture  $1 \mu\text{Ci}$  [ $^3\text{H}$ ]thymidine was added to each well. Cells were harvested, and the incorporated radioactivity was measured by scintillation counting.

#### 11.3.2.5 T-Cell Response to Completely Randomized Peptide Libraries

The TCC 5G7 was investigated for proliferative response to completely randomized peptide amide libraries ranging from hexa- ( $X_6$ ) to pentadecapeptides ( $X_{15}$ ).  $X_6$  to  $X_9$  induced no proliferative response, whereas  $X_{10}$  to  $X_{15}$  induced significant proliferation of the TCC (Fig. 11.8). The observation that the TCC responds to a mixture of  $2 \times 10^{14}$  different peptides ( $X_{11}$ ) almost as strongly as to the reporter peptide MBP(86-96), suggests a high degree of degeneracy in antigen recognition.



**Figure 11.8.** Proliferative response of T-cell clone 5G7 to MBP(86-96) and to completely randomized peptide amide libraries spanning the length from 6 ( $X_6$ ) to 15 amino acids ( $X_{15}$ ).

### 11.3.2.6 Complete Dissection of the Epitope for TCC 5G7

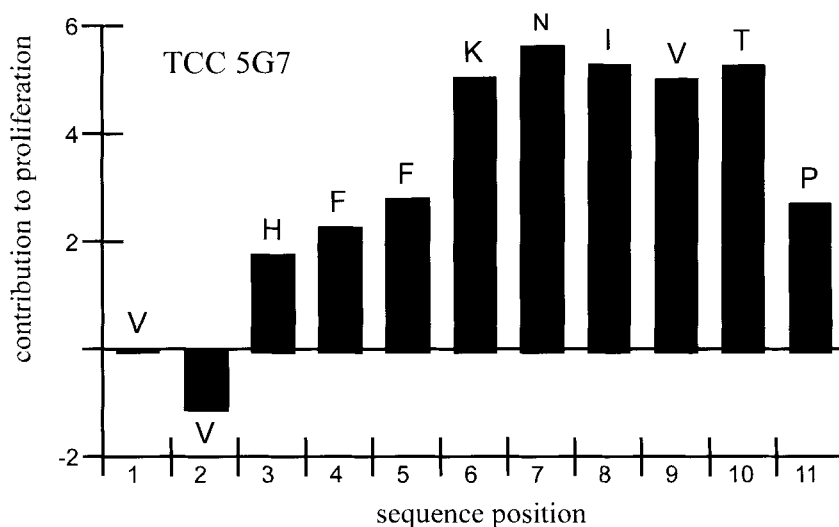
To define the proliferative response of TCC 5G7, different concentrations of the 220 peptide amide sublibraries were tested [71]. Differences in the proliferative response were found mainly between 100 and 250  $\mu\text{g}/\text{ml}$ . The results can be expressed as relative stimulation indices ( $\text{rel SI} = \text{SI}_{\text{sublib}}/\text{SI}_{X_{11}}$ ) and summarized in a 'Recognition Pattern' for TCC 5G7. No significant differences were observed in position 1 (P1). Sublibraries with defined residues I, L, M, V, and F in position 2 induced an increased response in comparison to  $X_{11}$ . The same amino acids were found for P3. In P4 residues F, M, Q, I, L, and W and in P5 F and Y were identified. In P6 amino acids K and R and in P7 residue N elicited increased responses. V and B in P8 and V, I, and B in P9 were positive. Finally, sublibraries with amino acids I, L, and V in P10 and K in P11 induced a proliferation. To summarize, amino acid residues in positions 2, 5, and 10 are thought to interact with the HLA-DR2b molecule. In accordance with binding studies, only hydrophobic aromatic and aliphatic residues were tolerated in position 2, and only the aromatic, uncharged amino acids F and Y in position 5. For P10, only aliphatic amino acids were favorable for T-cell proliferation. In the expected TCR contact positions, Q is preferred in P3 and the hydrophobic aromatic and aliphatic amino acids are favorable in P4. In contrast, P6 tolerated the positively charged amino acids K and R, and N was preferred in P7. P8 and P9 tolerate aliphatic amino acids. As described above, no clear conclusions can be drawn for P1 and P11. Interestingly, only six amino acid residues of the MBP(86-96) sequence matched with optimal library predictions for P2-P10.

### 11.3.2.7 Evaluation of TCR Contact Positions

In analogy to calculations based on the 'Activity Patterns' derived from competition assays, the sum of the absolute values of the  $\ln(\text{relSI})$  for all 20  $O/X_{10}$  sublibraries of a given defined position  $i$  ( $\sum |\ln \text{relSI}|$ ) represents its tolerance to amino acid variations with respect to recognition of the peptide-MHC complexes by the complementary TCR. The obtained  $11 \sum |\ln \text{relSI}|$ -values and the  $11 \sum |\ln \text{relCI}|$ -values were each normalized to the corresponding value of position 1. To eliminate effects of peptide-DR2b-binding, the normalized  $\sum |\ln \text{relCI}|$ -values were subtracted from the normalized  $\sum |\ln \text{relSI}|$ -values of the same sequence position (Fig. 11.9). These processed data showed that five positions (P6 to P10) were responsible for T-cell recognition of TCC 5G7 and TCC 5F6, whereas positions 2 to 5 are more tolerant to amino acid exchanges. These results reflect the recognition of the DR2b-peptide complex by the TCR.

### 11.3.2.8 Superagonists and Prediction of Epitopes

Based on the proliferative response of TCC to 220 peptide amide sublibraries, individual (defined) undecapeptide amides potentially acting as synthetic epitopes were synthesized. From investigation of these peptides in proliferation assays it became evident that TCC 5G7 responded to all selected peptides at much lower antigen concentration than MBP (86-96). The synthetic epitopes induced maximum stimulation at a concentration of 1 ng/ml, whereas 100  $\mu\text{g/ml}$  of the natural epitope MBP(86-96) was necessary to induce comparable proliferation.



**Figure 11.9.** Tolerance to amino acid variations in the 11 sequence positions with respect to proliferation of the TCC 5G7. Data were calculated as described from the allele-specific DR2b-'Activity Pattern' and from the 'Recognition Pattern' of TCC 5G7. Letters indicate amino acids from MBP (85-95).

To identify epitopes derived from natural proteins, the Swissprot database was searched for peptide sequences fitting the prerequisites for epitope residues established by screening peptide libraries. In scanning the human entries, several candidate self-peptides were identified and, interestingly, the MBP sequence was among them. A second search using BLITZ-search for the entire Swissprot database revealed several sequences of microbial antigens. Candidate self and foreign peptides, which matched at least six amino acid residues with the library prediction for P2–P10, were synthesized and tested in proliferation assays with TCC 5G7. Several peptides derived from human and microbial proteins were stimulatory for the TCC [71]. In addition, self and foreign peptides were shown to be superagonists as they were more potent agonists for TCC 5G7 than MBP(86-96).

## 11.4 Conclusions

Natural peptide libraries are involved in some of the most fundamental events in mammalian immune response, and are recognized in complex with MHC class I and II molecules. The receptors show high degeneracy with respect to their peptide interactions. Synthetic peptide libraries are promising tools to investigate the conformational space of MHC molecules, as well as the T-cell response induced by the recognition of the MHC-peptide complex. The influence of all amino acid residues on MHC binding can be summarized in an Activity Pattern of the peptide library as a source for the prediction and the design of numerous new natural or synthetic epitopes or ligands.

Results from testing combinatorial libraries will support the design of cocktails of peptides applicable in prophylactic or therapeutic tumor vaccination or the construction of immunomodulating peptides for therapy of autoimmune diseases. Peptide libraries are also one of the most promising tools in the development of therapeutic approaches based on the actual status of an individual's immune response.

## 11.5 Acknowledgements

This work has been supported by the Deutsche Forschungsgemeinschaft (SFB 510, projects C5 Jung-Wiesmüller and C3 Sommer-Jung).

## 11.6 References

- [1] Rock KL, Gramm C, Rothstein L, Clark K, Stein R, Dick L, Hwang D, Goldberg AL (1994) Inhibitors of the proteasome block the degradation of most cell-proteins and the generation of peptides presented on MHC class-I molecules. *Cell* 78:761–771.
- [2] Niedermann G, Butz S, Ihlenfeldt H-G, Grimm R, Lucchiari M, Hoshützky H, Jung G, Maier B, Eichmann K (1995) Contribution of proteasome-mediated proteolysis to the hierarchy of epitopes presented by the major histocompatibility complex class I molecules. *Immunity* 2:289–299.
- [3] Koopmann JO, Hämmerling GJ, Momburg F (1997) Generation, intracellular-transport and loading of peptides associated with MHC class-I molecules. *Curr Opin Immunol* 9:80–88.
- [4] Gosselin EJ, Wardwell K, Gosselin DR, Alter N, Fisher JL, Guyre PM (1992) Enhanced antigen presentation using human Fc-g receptor-specific immunogens. *J Immunol* 149:3477–3481.

- [5] Vidard L, Rock KL, Benacerraf B (1992) Diversity in MHC class II ovalbumin epitopes generated by distinct proteases, *J Immunol* 149:498–504.
- [6] Geysen HM, Rodda SJ, Mason TJ (1986) A priori delineation of a peptide which mimics a discontinuous antigenic determinant, *Mol Immunol* 23:709–715.
- [7] Houghten RA, Pinilla C, Blondelle SE, Appel JR, Dooley CT, Cuervo JH (1991) Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery, *Nature* 364:84–86.
- [8] Pinilla C, Appel JR, Blanc P, Houghten RA (1992) Rapid identification of high-affinity peptide ligands using positional scanning synthetic peptide combinatorial libraries, *BioTechniques* 13:901–905.
- [9] Lam KS, Salmon SE, Hersh EM, Hruby VJ, Kazmierski WM, Knapp RJ (1991) A new type of synthetic peptide library for identifying ligand-binding activity, *Nature* 364:82–84.
- [10] Gallop M, Barrett RW, Dower WJ, Fodor SPA, Gordon EM (1994) Applications of combinatorial technologies to drug discovery. 1. Background and peptide combinatorial libraries, *J Med Chem* 37:1233–1251.
- [11] Früchtel JS, Jung G (1996) Organic Chemistry on solid supports, *Angew Chem* 108:19–46, *Angew Chem Int Ed Engl* 35:17–42.
- [12] Wiesmüller K-H, Feiertag S, Fleckenstein B, Kienle S, Stoll D, Herrmann M, Jung G (1996) Peptide and cyclopeptide libraries: Automated synthesis, analysis and receptor binding assays. In: Jung G (ed) *Combinatorial Peptide and Nonpeptide Libraries - A Handbook for the Search of Lead Structures*, pp 203–246, Verlag Chemie, Weinheim, Germany.
- [13] Metzger JW, Kemper C, Wiesmüller K-H, Jung G (1994) Electrospray mass spectrometry and tandem mass spectrometry of synthetic multicomponent peptide mixtures: determination of composition and purity, *Anal Biochem* 219:261–277.
- [14] Stevanovic S, Jung G (1993) Multiple sequence analysis: Pool sequencing of synthetic and natural peptide libraries, *Anal Biochem* 212:212–220.
- [15] Müller KM, Ebensperger C, Tampé R (1994) Nucleotide binding to the hydrophilic C-terminal domain of the transporter associated with antigen processing (TAP), *J Biol Chem* 269:14032–14037.
- [16] Kleijmeer M, Kelly A, Geuze HJ, Slot JW, Townsend A, Trowsdale J (1992) Location of MHC-encoded transporters in the endoplasmic reticulum and *cis*-golgi, *Nature* 357:342–344.
- [17] Momburg F, Roelse J, Hämmerling GJ, Neeffjes JJ (1994) Peptide size selection by the major histocompatibility complex-encoded peptide transporter, *J Exp Med* 179:1613–1623.
- [18] van Endert PM, Tampé R, Meyer TH, Tisch R, Bach JF, McDevitt HO (1994) A sequential model for peptide binding and transport by the transporters associated with antigen processing, *Immunity* 1:491–500.
- [19] Koopmann JO, Post M, Neeffjes JJ, Hämmerling GJ, Momburg F (1996) Translocation of long peptides by transporters associated with antigen-processing (TAP), *Eur J Immunol* 26:1720–1728.
- [20] Uebel S, Kraas W, Kienle S, Wiesmüller K-H, Jung G, Tampé R (1997) Recognition principle of the TAP-transporter disclosed by combinatorial peptide libraries, *Proc Natl Acad Sci USA* 94:8976–8981.
- [21] Uebel S, Meyer TH, Kraas W, Kienle S, Jung G, Wiesmüller KH, Tampé R (1995) Requirements for peptide binding to the human transporter associated with antigen-processing revealed by peptide scans and complex peptide libraries, *J Biol Chem* 270:18512–18516.
- [22] Grommé M, van der Valk R, Shiedregt K, Vernie L, Liskamp R, Hämmerling GJ, Koopmann J-O, Momburg F, Neeffjes JJ (1997) The rational design of TAP inhibitors using peptide substrate modifications and peptidomimetics, *Eur J Immunol* 27:898–904.
- [23] Ehring B, Meyer TH, Eckerskorn C, Lottspeich F, Tampé R (1996) Effects of MHC-encoded subunits on the peptidase and proteolytic activities of human 20S proteasomes - Cleavage of proteins and antigenic peptides, *Eur J Biochem* 235:404–415.
- [24] Rammensee HG, Friede T, Stevanovic S (1995) MHC ligands and peptide motifs: First listing, *Immunogenetics* 41:178–222.
- [25] Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC (1987) Structure of the human class I histocompatibility antigen HLA-A2, *Nature* 329:506–512.
- [26] Brown JH, Jardetzky T, Gorga JC, Stern LJ, Strominger JL, Wiley DC (1993) The three-dimensional structure of the human class II histocompatibility antigen HLA-DR1, *Nature* 364:33–39.
- [27] Cresswell P (1992) Chemistry and functional role of the invariant chain, *Curr Opin Immunol* 4:87–92.



- [28] Sherman MA, Weber DA, Jensen PE (1995) DM enhances peptide binding to class II MHC by release of invariant chain-derived peptide, *Immunity* 3:197–205.
- [29] Gavin MA, Gilbert MJ, Ridell SR, Greenberg RD, Bevan MJ (1993) Alkali hydrolysis of recombinant proteins allows for the rapid identification of class I MHC-restricted CTL epitopes, *J Immunol* 151:3971–2894.
- [30] Falk K, Rötzschke O, Stevanovic S, Jung G, Rammensee HG (1994) Pool sequencing of natural HLA-DR, DQ and DP ligands reveals detailed peptide motifs, constraints of processing, and general rules, *Immunogenetics* 39:230–242.
- [31] Udaka K, Wiesmüller K-H, Kienle S, Jung G, Walden P (1995) Tolerance to amino acid variations in peptides binding to the MHC class I protein H-2K<sup>b</sup>, *J Biol Chem* 270:24130–24136.
- [32] Blake J, Johnston JV, Hellström KE, Marquardt H, Chen L (1996) Use of combinatorial libraries to construct functional mimics of tumor epitopes recognized by MHC class I-restricted cytotoxic T lymphocytes, *J Exp Med* 184:121–130.
- [33] Falk K, Rötzschke O, Stevanovic S, Jung G, Rammensee HG (1991) Allele-specific motifs revealed by sequencing of self peptides eluted from MHC molecules, *Nature* 351:290–296.
- [34] Chen W, McCluskey J, Rodda S, Carbone FR (1993) Changes at peptide residues buried in the major histocompatibility complex (MHC) influence T cell recognition: a possible role for indirect conformational alterations in the MHC class I or bound peptide in determining T cell recognition, *J Exp Med* 177:869–873.
- [35] Fremont DH, Matsumura M, Stura EA, Peterson PA, Wilson IA (1992) Crystal structure of two viral peptides in complex with murine MHC class I H-2K<sup>b</sup>, *Science* 257:919–927.
- [36] Madden DR, Gorga JC, Strominger JL, Wiley DC (1992) The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC, *Cell* 70:1035–1048.
- [37] Garboczi DN, Ghosh P, Utz U, Fan QR, Biddison WE, Wiley DC (1996) Structure of the complex between human T cell receptor, viral peptide and HLA-A2, *Nature* 384:134–141.
- [38] Garcia KC, Degano M, Stanfield RL, Brunmark A, Jackson MR, Peterson PA, Teyton L, Wilson IA (1996) An  $\alpha\beta$  T cell receptor structure at 2.5 Å and its orientation in the TCR-MHC complex, *Science* 274:209–219.
- [39] Udaka K, Wiesmüller K-H, Kienle S, Jung G, Walden P (1995) Decrypting the structure of MHC-I restricted CTL epitopes with complex peptide libraries, *J Exp Med* 181:2097–2108.
- [40] Pridzun L, Wiesmüller K-H, Kienle S, Jung G, Walden P (1996) Amino acid preferences in the octapeptide subunit of the major histocompatibility complex class I heterotrimer H-2L<sup>d</sup>, *Eur J Biochem* 236:249–253.
- [41] Stevens J, Wiesmüller K-H, Barker PJ, Walden P, Butcher G, Joly E (1998) Efficient generation of MHC class I-peptide complexes using synthetic peptide libraries, *J Biol Chem* 274:2874–2884.
- [42] Rudensky AY, Preston-Hurlburt P, Al-Ramadi BK, Rothbard J, Janeway CA (1992) Truncation variants of peptides isolated from MHC class II molecules suggest sequence motifs, *Nature* 359:429–431.
- [43] Chicz RM, Urban RG, Lane WS, Gorga JC, Stern LJ, Vignali DAA, Strominger JL (1992) Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size, *Nature* 358:764–768.
- [44] Hammer J, Valsasini P, Tolba K, Bolin D, Higelin J, Takacs B, Sinigaglia F (1993) Promiscuous and allele-specific anchors in HLA-DR-binding peptides, *Cell* 74:197–203.
- [45] Jung G (1996) Natural peptide libraries of microbial and mammalian origin. In: Jung G (ed) *Combinatorial peptide and nonpeptide libraries - A handbook for the search of lead structures*, Verlag Chemie, Weinheim, Germany.
- [46] Malcherek G, Gnau V, Stevanovic S, Rammensee H-G, Jung G, Melms A (1994) Analysis of allele-specific contact sites of natural HLA-DR17 ligands, *J Immunol* 153:1141–1149.
- [47] Chicz RM, Urban RG, Gorga JC, Vignali DA, Lane WS, Strominger JL (1993) Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles, *J Exp Med* 178:27–47.
- [48] Hammer J, Takacs B, Sinigaglia F (1992) Identification of a motif for HLA-DR1 binding peptides using M13 phage display libraries, *J Exp Med* 176:1007–1013.
- [49] Stern LJ, Brown JH, Jardetzky TS, Gorga JC, Urban RG, Strominger JL, Wiley DC (1994) Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide, *Nature* 368:215–221.

- [50] Townsend A, Bodmer H (1989) Antigen recognition by class I-restricted T lymphocytes, *Ann Rev Immunol* 7:601–624.
- [51] Schwartz RH (1985) T lymphocyte recognition of antigen in association with gene products of the major histocompatibility complex, *Ann Rev Immunol* 3:261–327.
- [52] Singer A, Hodes RJ (1983) Mechanism of T cell B cell interaction, *Ann Rev Immunol* 1:211–241.
- [53] Udaka K, Wiesmüller K-H, Kienle S, Jung G, Walden P (1996) Self MHC-restricted peptides recognized by an alloreactive T lymphocyte clone, *J Immunol* 157:670–678.
- [54] Gundlach B, Wiesmüller K-H, Junt T, Kienle S, Jung G, Walden P (1996) Determination of T cell epitopes with random peptide libraries, *J Immunol Meth* 192:146–155.
- [55] Gundlach B, Wiesmüller K-H, Junt T, Kienle S, Jung G, Walden P (1996) Specificity and degeneracy of minor histocompatibility antigen-specific MHC-restricted cytotoxic T lymphocytes, *J Immunol* 156:3645–3651.
- [56] Arnold D, Keilholz W, Schild H-J, Dumrese T, Stevanovic S, Rammensee H-G (1997) Substrate specificity of cathepsins D and E determined by N-terminal and C-terminal sequencing of peptide pools, *Eur J Biochem* 249:171–179.
- [57] Tana T, Kamikawaji N, Savoie CJ, Sudo T, Kinoshita Y, Sasazuki T (1998) An HLA-binding-motif aided peptide epitope library: a novel library design for the screening of HLA-DR4-restricted antigenic peptides recognized by CD4<sup>+</sup> T cells, *J Hum Genet* 43:14–21.
- [58] Hiemstra HS, Duinkerken G, Benckhuijsen WE, Amons R, de Vries RR, Roep BO, Drijfhout JW (1997) The identification of CD4<sup>+</sup> T cell epitopes with dedicated synthetic peptide libraries, *Proc Natl Acad Sci USA* 94:10313–10318.
- [59] Daniel C, Horvath S, Allen PM (1998) A basis for alloreactivity: MHC helical residues broaden peptide recognition by the TCR, *Immunity* 8:543–552.
- [60] Estaquier J, Ameisen J-C, Auriault C, Boutillon C, Gras-Masse H, Tartar A (1998) Combinatorial peptide library as an immunogen, *Methods Mol Biol* 87:281–296.
- [61] Gavin MA, Bevan MJ (1998) Major histocompatibility complex allele-specific peptide libraries and identification of T-cell mimotopes, *Methods Mol Biol* 87:235–248.
- [62] Metzger JW, Wiesmüller K-H, Kienle S, Brünjes J, Jung G (1996) Mass Spectrometric Analysis of Peptide Libraries, In: Jung G (ed) *Combinatorial Peptide and Nonpeptide Libraries – A Handbook for the Search of Lead Structures*, pp 247–285, Verlag Chemie, Weinheim, Germany.
- [63] Max H, Halder T, Kalbus M, Gnau V, Jung G, Kalbacher H (1994) A 16-residue peptide of the human autoantigen calreticulin is a most prominent HLA-DR4Dw4-associated self-peptide, *Human Immunol* 41:39–45.
- [64] Klohe EP, Watts R, Bahl M, Alber C, Yu W-Y, Anderson R, Silver J, Gregersen PK, Karr RW (1988) Analysis of the molecular specificities of anti-class II monoclonal antibodies by using L cell transfectants expressing HLA class II molecules, *J Immunol* 141:2158–2164.
- [65] Gorga JC, Horejsi V, Johnson DR, Raghupathy R, Strominger JL (1987) Purification and characterization of class II histocompatibility antigens from a homozygous human B cell line, *J Biol Chem* 262:16087–16094.
- [66] Fleckenstein B, Kalbacher H, Muller CP, Stoll D, Halder T, Jung G, Wiesmüller K-H (1996) New ligands binding to the human leukocyte antigen class II molecule DRB1\*0101 based on the activity pattern of an undecapeptide library, *Eur J Biochem* 240:71–77.
- [67] Riberdy JM and Cresswell P (1992) The antigen-processing mutant T2 suggests a role for MHC-linked genes in class II antigen presentation, *J Immunol* 148:2586–2592.
- [68] Jung C, Kalbus M, Fleckenstein B, Malcherek G, Melms A, Jung G, Wiesmüller K-H (1998) New ligands for HLA DRB1\*0301 by random selection of favourable amino acids ranked by competition studies with undecapeptide amide sublibraries, *J Immunol Meth* 219:139–149.
- [69] Terreaux et al., unpublished.
- [70] Fleckenstein B, Jung G, Wiesmüller K-H (1997) Prediction and Design of new MHC class II-ligands based on the Activity Pattern of a Synthetic Undecapeptide Library, In: Brown F, Burton D, Doherty PC, Mekananos J, Norrby E (eds) *Vaccines 97*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.
- [71] Hemmer B, Fleckenstein B, Vergelli M, Jung G, McFarland H, Martin R, Wiesmüller KH (1997) Identification of high potency microbial and self ligands for a human autoreactive class II-restricted T cell clone, *J Exp Med* 185:1651–1659.

- [72] Martin R, Utz U, Coligan JE, Richert J, Flerlage M, Robinson E, Stone R, Biddison WE, McFarlin DE, McFarland HF (1992) Diversity in fine specificity and T cell receptor usage of the human CD4<sup>+</sup> cytotoxic T cell response specific for the immunodominant myelin basic protein peptide 87–106, *J Immunol* 148:1359–1366.
- [73] Vergelli M, Hemmer B, Utz U, Vogt AB, Kalbus M, Tranquill L, Conlon P, Ling N, Steinmann L, McFarland HF, Martin R (1996) Differential activation of human autoreactive T cell clones by altered peptide ligands derived from myelin basic protein peptide (87–99), *Eur J Immunol* 26: 2624–2634.

# 12 Combinatorial Biosynthesis of Microbial Metabolites

*Andreas Bechthold and José Antonio Salas Fernández*

## 12.1 Introduction

Advances in recombinant DNA technology have enabled the cloning and expression of many biosynthetic gene clusters from different biological sources. Gene clusters encoding proteins catalyzing the biosynthesis of different natural bioactive products have been isolated and characterized from plants, fungi, and bacteria. Since many natural bioactive compounds are produced by actinomycetes, most of them by streptomycete species, many research groups have focused their research on the cloning of biosynthetic genes of these organisms. One important landmark in this field was the pioneer research work by David Hopwood and co-workers at the John Innes Institute (Norwich, UK) on the development of *Streptomyces* genetic tools for the cloning and expression of streptomycete genes. From the isolation of the first antibiotic biosynthetic genes from a streptomycete [1–3] and the cloning and expression of entire biosynthetic pathways of antibiotic-producing streptomycetes [4, 5], much progress has been achieved. Functions have been assigned to many gene products in the biosynthesis of different bioactive compounds. More recently, researchers in this area – including molecular biologists and bioorganic chemists – have begun to use these different genes to alter the structure of natural compounds by genetic engineering, or to combine genes from different biosynthetic pathways. This new technology, named combinatorial biosynthesis [6, 7], might result in the formation of novel natural products. The goal of this chapter is to review the ‘state-of-the-art’ and prospects in this field of research.

## 12.2 Cloning of Biosynthetic Gene Clusters

The cloning of biosynthetic gene clusters started with the shotgun cloning of random fragments of DNA from a *Streptomyces* strain, producing a given antibiotic, into a mutant of the strain blocked at a step in the biosynthesis of the antibiotic. Fragments restoring antibiotic production were used for further investigations. For example, this approach has been successfully performed to clone the gene clusters of actinorhodin (**1**) [4], tetracenomycin (**2**) [5], or avermectin (**3**) [8] biosynthesis. In another approach, resistance genes of antibiotic producers were cloned in a heterologous host. This led to the identification of linked biosynthetic genes as demonstrated for the erythromycin- (**4**) [9, 10], oxytetracycline- (**5**) [11], and spiramycin- (**6**) [12] biosynthetic gene clusters. A few gene clusters

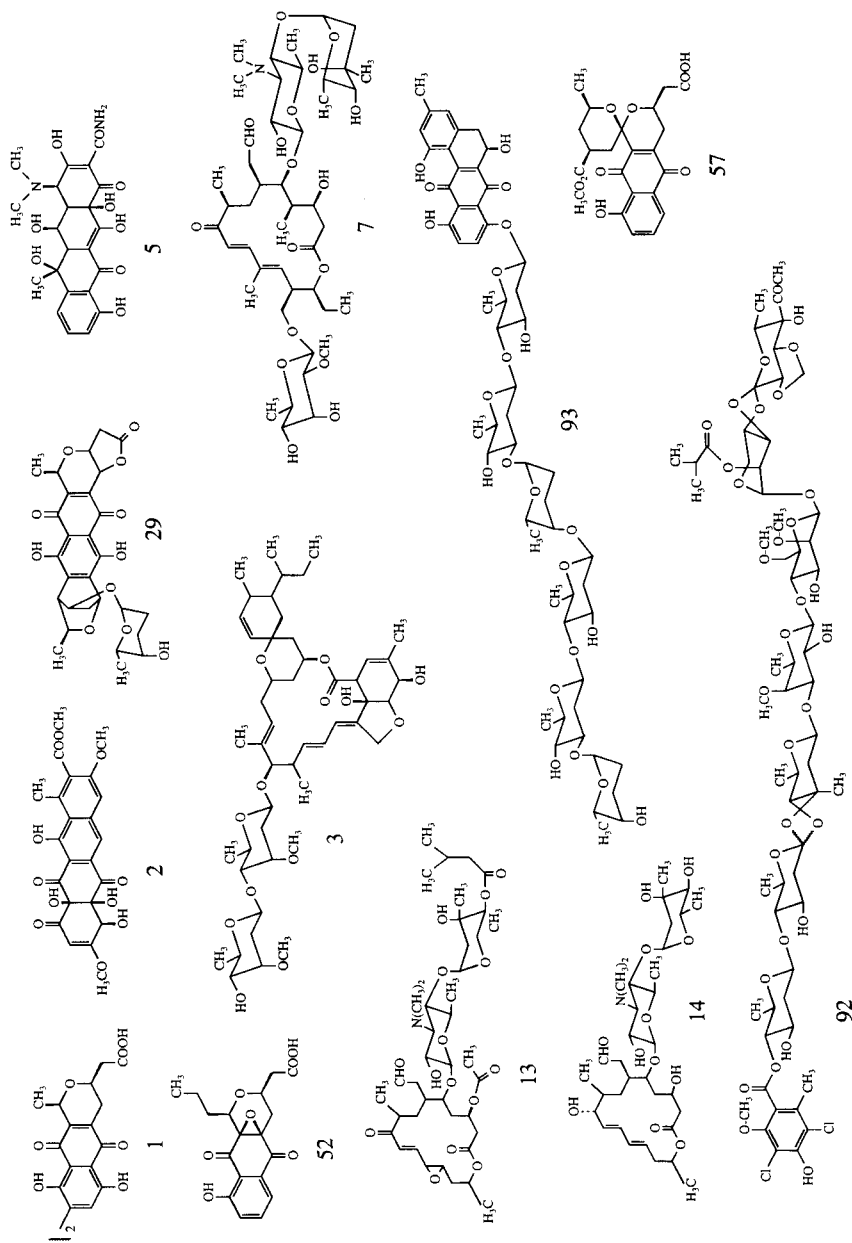
**Table 12.1.** Examples of antibiotic biosynthetic pathways from which the gene clusters have been cloned and characterized.

Strain	Antibiotic	Gene	Reference
<i>S. coelicolor</i> A3 (2)	actinorhodin ( <b>1</b> )	<i>act</i>	[4]
<i>S. glaucescens</i> Tü49	tetracenomycin ( <b>2</b> )	<i>tcm</i>	[5]
<i>S. violaceoruber</i> Tü22	granaticin ( <b>29</b> )	<i>gra</i>	[23]
<i>S. rimosus</i>	oxytetracycline ( <b>5</b> )	<i>oct</i>	[11]
<i>S. peucetius</i> ATCC 29050	daunomycin ( <b>35</b> )	<i>dps</i>	[32]
<i>S. olivaceus</i> Tü353	clloramycin ( <b>30</b> )	<i>elm</i>	[24]
<i>S. griseus</i> N2-3-11	griseusin ( <b>57</b> )	<i>gris</i>	[47]
<i>S. fradiae</i> Tp2717	urdamycin ( <b>18</b> )	<i>urd</i>	[19]
<i>S. cyanogenus</i> S136	landomycin ( <b>93</b> )	<i>lan</i>	manuscript in preparation
<i>S. roseofulvus</i>	frenolicin ( <b>53</b> )	<i>fren</i>	[47]
<i>S. argillaceus</i>	mithramycin ( <b>34</b> )	<i>mtm</i>	[68, 69]
<i>Saccharopolyspora erythraea</i>	erythromycin ( <b>4</b> )	<i>ery</i>	[9, 10]
<i>S. fradiae</i> T59235	tylosin ( <b>7</b> )	<i>tyl</i>	[13]
<i>S. ambofaciens</i>	spiramycin ( <b>6</b> )	<i>srm</i>	[12]
<i>S. mycarofaciens</i>	midecamycin ( <b>14</b> )	<i>mdm</i>	[18]
<i>S. thermotolerans</i>	carbomycin ( <b>13</b> )	<i>car</i>	[17]
<i>Streptomyces</i> sp. MA6548	FK506 ( <b>10</b> )	<i>fkf</i>	[16]
<i>S. avermitilis</i>	avermectin ( <b>3</b> )	<i>ave</i>	[8]
<i>S. viridochromogenes</i> Tü57	avilamycin A ( <b>92</b> )	<i>avi</i>	[67]
<i>Bacillus subtilis</i>	surfactin ( <b>82</b> )	<i>srf</i>	[58]
<i>Amycolatopsis orientalis</i> C329.4	vancomycin ( <b>20</b> )	<i>gtf</i>	[1]

have been cloned after purifying proteins involved in the biosynthesis of the antibiotic and cloning the gene via reverse genetics. As an example, an O-methyltransferase (TylF) involved in tylosin biosynthesis (**7**) was purified and partially sequenced. Based on the deduced DNA sequence information about *tylF*, the entire cluster was cloned [13]. Genes with known functions were used as probes in Southern hybridization to identify homologous genes in other organisms. The most popular gene probes were *actI* and *actIII* from *S. coelicolor* and *strD* from *S. griseus*. Several biosynthetic gene clusters were isolated using these gene probes [14]. A selection of antibiotics whose biosynthetic gene clusters have been cloned and some representative structures are given in Table 12.1 and Figure 12.1.

### 12.3 New Drugs by Genetic Engineering

Approaches to produce new compounds with altered structures involved targeted gene disruption experiments, expression of single genes or gene clusters, and the rational design. In this chapter, we summarize the currently available strategies to create novel molecules.



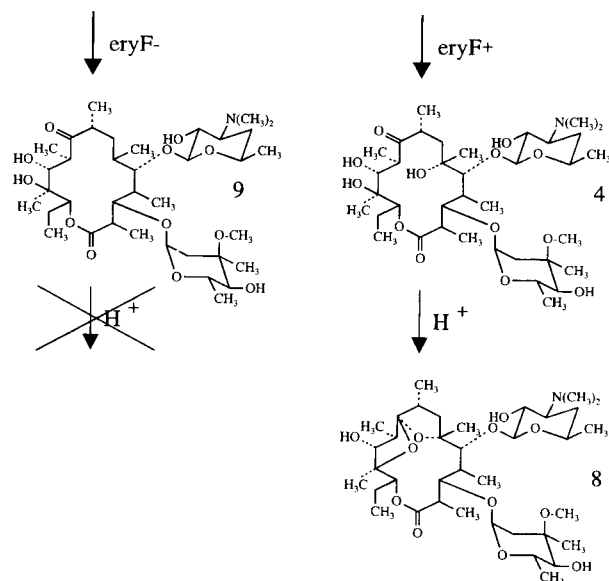
**Figure 12.1.** Structures of selected natural products derived from polyketide syntheses.

### 12.3.1 New Drugs by Targeted Gene Disruption

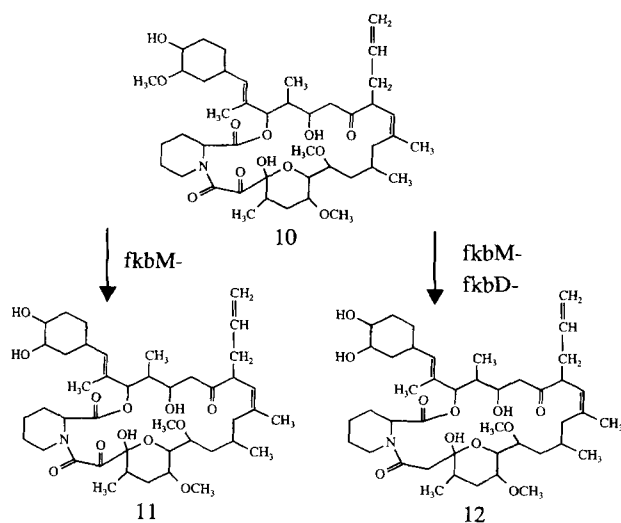
Inactivation of specific selected genes is now a very usual methodology for demonstrating the involvement of a gene product in catalyzing a given function. It can be also used to generate novel structural variations of known natural products. This approach can be taken even if a detailed understanding of the biochemistry and genetics involved in the formation of the desired compound is not available, although this knowledge can greatly improve the efficiency of this approach. A few examples demonstrate the potentiality of using this method to create specifically altered molecules.

Erythromycin (**4**) is a macrolide antibiotic clinically useful in the treatment of infections by Gram-positive bacteria. A hydroxyl group at carbon-6 of the erythronolide macrolactone is involved in acidic inactivation in the stomach when administered orally through its conversion into anhydroerythromycin (**8**). Erythromycin derivatives lacking this hydroxyl group escape the acidic inactivation, and are therefore interesting from therapeutic and pharmacological points of view. Leonard Katz and his group at Abbott Laboratories inactivated the *eryF* gene which encodes a cytochrome  $P_{450}$  mono-oxygenase responsible for the introduction of this hydroxyl group into the macrolactone and the mutant produced 6-deoxyerythronolide B (**9**) (Fig. 12.2). This is a much more acid-stable antibiotic (six times more stable than erythromycin) and, although slightly less active, oral administration is as efficient as erythromycin due to its higher stability [15].

Analogues of the immunosuppressant FK506 (**10**) were also produced by gene inactivation. This drug is used in the prevention of graft rejection and in the treatment of auto-



**Figure 12.2.** Production of 6-deoxyerythronolide B (**8**) after disruption of *eryF* in *Saccharopolyspora erythraea*.



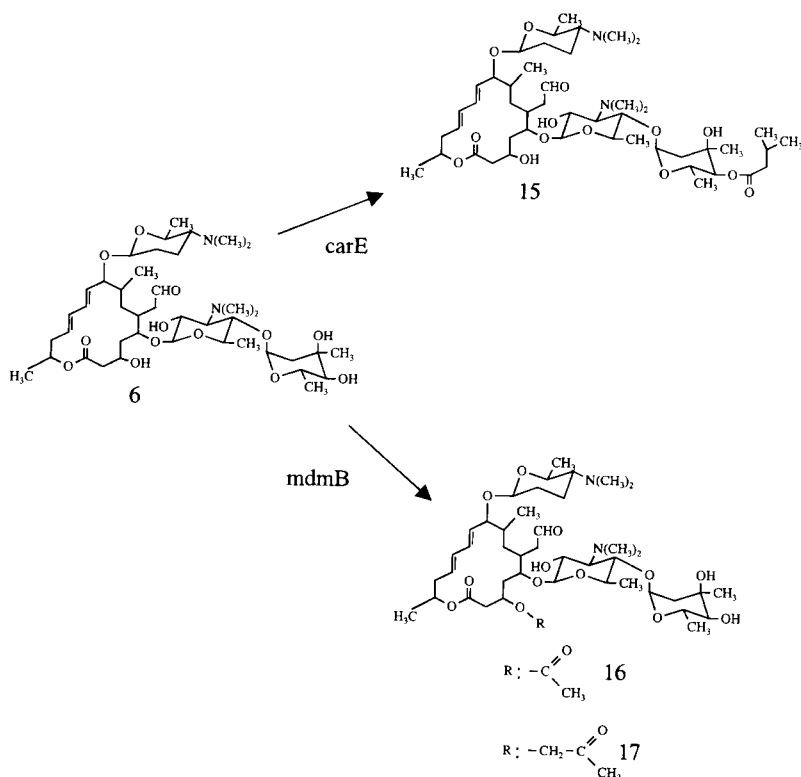
**Figure 12.3.** Production of 31-demethyl FK506 (**11**) and 9-deoxy-31-O-demethyl FK506 (**12**) after disruption of a methyltransferase gene (*fkbM*) and a hydroxylase gene (*fkbD*) in the producer of FK506 (**10**).

immune diseases, but shows high neurotoxicity and nephrotoxicity. Disruption of the *fkbD* and *fkbM* genes which code for 31-O-demethyl-FK506 methyltransferase and a cytochrome P<sub>450</sub>-dependent hydroxylase resulted in the accumulation of two new biologically active metabolites, 31-demethyl FK506 (**11**) and 9-deoxy-31-O-demethyl FK506 (**12**) (Fig. 12.3). These two compounds were difficult to obtain by classical medicinal chemistry, and may serve as starting materials for further chemical modification aimed at imparting an improved therapeutic index compared to the parent compound [16].

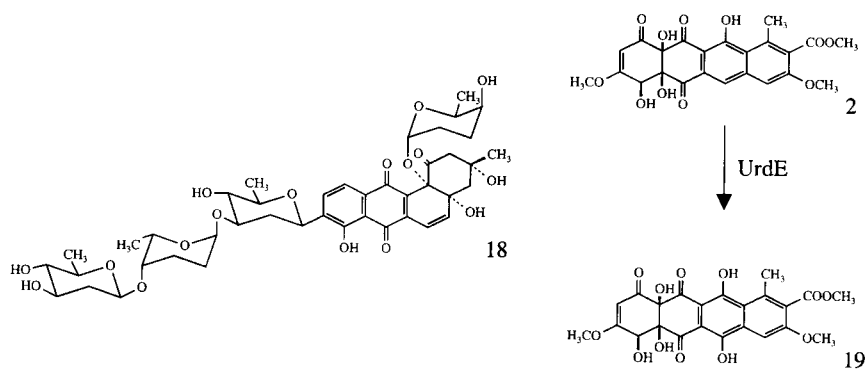
### 12.3.2 New Drugs by Expression of Single Genes

The expression of single genes in heterologous hosts has been used successfully to create a number of hybrid compounds. The potentiality of this approach to incorporate new functional groups into a given structure can be shown by a few different examples. Carbomycin (**13**), midecamycin (**14**) and spiramycin (**6**) are macrolide antibiotics. Expression of the *carE* gene from *S. thermotolerans* (carbomycin producer) or the *mdmB* from *S. mycarofaciens* (midecamycin producer) in *S. ambofaciens* (spiramycin producer) led to the production of new antibiotics, 4''-O-isovalerylspiramycin (**15**), 3-O-acetyl- (**16**) and 3-O-propionylspiramycin (**17**) which were not synthesized by the wild-type strain (Fig. 12.4) [17, 18]. Heterologous expression of a putative oxygenase gene from the urdamycin (**18**) gene cluster in *S. glaucescens* GLA.O, producer of tetracenomycin (**2**), caused the production of a hybrid antibiotic 6-hydroxy tetracenomycin C (**19**) [19] (Fig. 12.5). This experiment indicated a relaxed substrate specificity of this oxygenase.





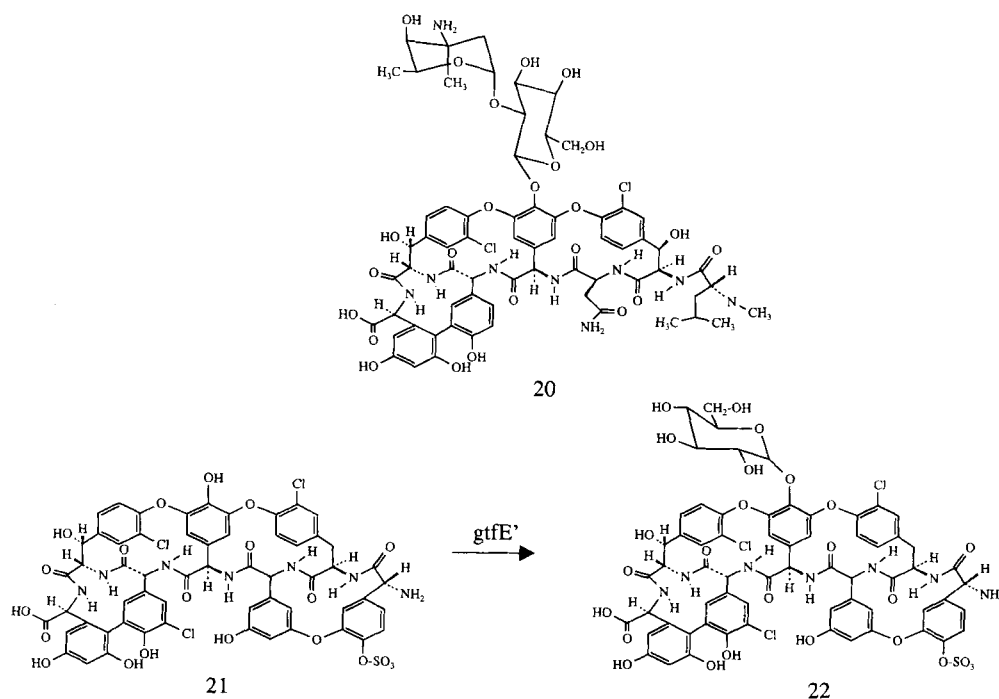
**Figure 12.4.** Production of 4''-*O*-isovaleryspiramycin (**15**), 3'-*O*-acetylspiramycin (**16**) and 3'-*O*-propionylspiramycin (**16**) after expression of *carE* from *S. thermotolerans* or *mdmB* from *S. mycarofaciens* in *S. ambofaciens* (spiramycin (**6**) producer).



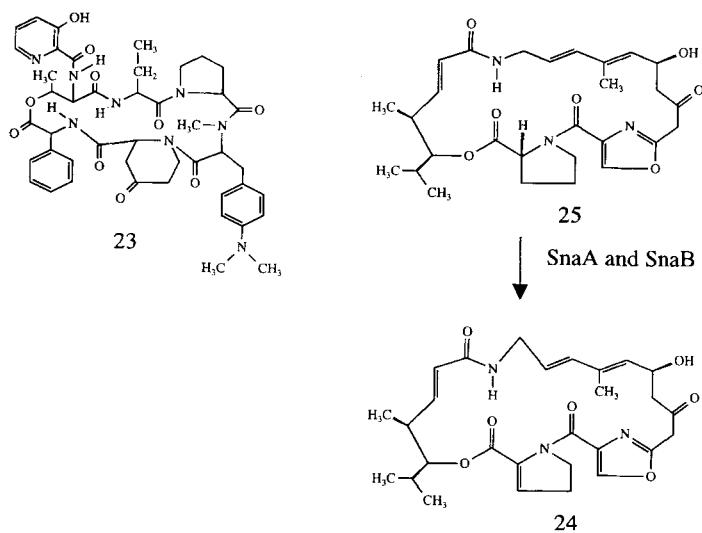
**Figure 12.5.** Conversion of tetraenomycin (**2**) into 6-hydroxy tetraenomycin (**19**) after expression of an oxygenase gene (*urdE*) from the urdamycin (**18**) producer in *S. glaucescens*.

Single gene expression in another host can also be used for the transfer of different sugars. The *gtfE'* gene from *Amycolatopsis orientalis* encodes a TDP-glucose:aglycosyl vancomycin transferase involved in the transfer of glucose to the aglycone of vancomycin (**20**). In *S. toyocaensis*, producer of the glycopeptide A47934 (**21**), the conversion of A47934 into glycosyl-A47934 (**22**) was observed [20] (Fig. 12.6).

The expression of two genes has been used to overcome a metabolic bottleneck in the production of the antibiotic pristinamycin. This is produced by *S. pristinaespiralis*, and is a mixture of two types of macrocyclic lactone peptolides, pristinamycin I (PI) (**23**), a branched cyclic hexadepsipeptides of the streptogramin B group, and pristinamycin II (PII), a polyunsaturated cyclic peptolides of the streptogramin A group. PI and PII each inhibit the growth of bacteria, displaying only bacteriostatic activity. In combination they demonstrate a synergistic bactericidal activity, even at low concentration. The mixture is active against the dreaded *Streptococcus* and *Staphylococcus*, which makes it a valuable antibiotic. The PII component of pristinamycin is produced mainly in two forms, called PIIA (**24**) (80%) and PIIB (**25**) (20%). A water-soluble derivative of pristinamycin, named RP59500, was obtained by chemical modification of PIIA. To generate a PIIA-specific producer strain two genes, *snaA* and *snaB*, were isolated from the biosynthetic gene cluster. The enzymes encoded by *snaA* and *snaB* catalyze the conversion of PIIB to PIIA (Fig. 12.7). Both genes were placed under the transcription control of a strong promoter (*ermEp*) and were cloned into an integrative, non replicative and non conjugative



**Figure 12.6.** Conversion of A47934 (**21**) into glycosyl A47-934 (**22**) after expression of a glycosyltransferase gene (*gtfE*) from the vancomycin (**20**) producer in *S. toyocaensis*.

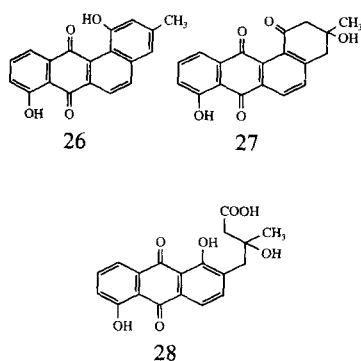


**Figure 12.7.** Structures of pristinamycin I (PI) (**23**), pristinamycin PIIB (**25**) and pristinamycin PIIA (**24**). The conversion of PIIB into PIIA is catalyzed by SnaA and SnaB (PIIA synthase). The overexpression of *snaA* and *snaB* in *S. pristinaespiralis* resulted in the complete conversion of PIIB to PIIA at high yield.

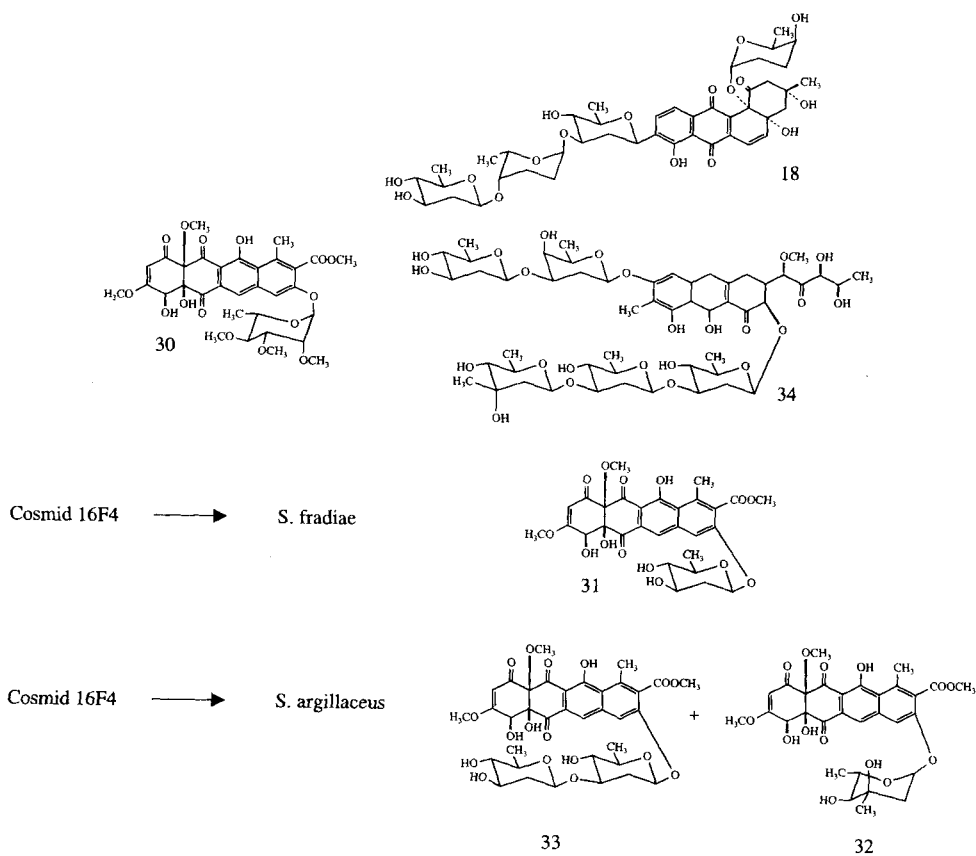
vector. The integration of this vector into the chromosome of the producer strain resulted in the production of 100% PIIA, and this was achieved in high concentrations [21]. In addition, the recombinant strain showed 100% stability and can be maintained without antibiotic selection.

### 12.3.3 New Drugs by Expression of Gene Clusters

Biosynthetic gene clusters of many natural compounds are located on the chromosome of producer strains extending within a range between 20 and nearly 100 kb. Large DNA fragments can be isolated by preparing cosmid libraries using chromosomal DNA of the producer strains, and single cosmids containing DNA fragments can be expressed in heterologous *Streptomyces* strains (e.g., *S. lividans* TK24, *S. coelicolor* CH999). These recombinant strains can produce either the entire antibiotic, or derivatives of it. *S. rimosus* NRRL 3016 produces the angucycline antibiotics tetrangulol (**26**) and tetrangomycin (**27**). Cosmids containing DNA of this strain were expressed in *S. lividans*, resulting not only in the production of tetrangulol and tetrangomycin but also in the production of larger amounts of fridamycin E (**28**) (Fig. 12.8) [22]. Possibly, this compound was produced because of a relaxed specificity of an oxygenase gene in the foreign host. Heterologous expression of another complete pathway, that leads to granaticin (**29**), was achieved using the cosmid pOJ446-22-24 containing DNA of *S. violaceoruber* Tü22 in *S. coelicolor* CH999 [23]. This compound contains a sugar side chain which is connected to a polyketide moiety. This example clearly describes that *S. coelicolor* CH999, which does not accumulate any known glycosylated compound, is able to produce sugar-containing derivatives.



**Figure 12.8.** Structures of tetragulol (26), tetragomycin (27), and fridamycin E (28). These compounds were produced after expression of DNA (cosmid) from *S. rimosus* in *S. lividans*.

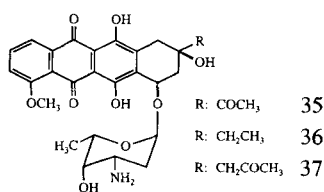


**Figure 12.9.** Production of glycosylated tetragenomycin C derivatives (31, 32, 33) after expression of the elloramycin (30) biosynthetic gene cluster in *S. fradiae* (urdamycin (18) producer) and *S. argillaceus* (mithramycin (34) producer).

The expression of a 25 kb DNA fragment containing genes of the elloramycin (**30**) biosynthetic gene cluster in the urdamycin producer *S. fradiae* resulted in the production of the hybrid antibiotic 8-D-oliviosyl-8-demethyl tetracenomycin C (**31**) [24]. This compound clearly contains structural elements of elloramycin (**30**) (polyketide moiety) and of urdamycin (**18**) (sugar moiety) (Fig. 12.9). In a similar manner, two new glycosylated tetracenomycins, 8-D-mycarosyl-8-demethyl tetracenomycin C (**32**) and 8-D-dioliviosyl-8-demethyl tetracenomycin C (**33**) (Fig. 12.9) were produced by expressing the same DNA fragment of the elloramycin producer in *S. argillaceus*, the producer of the antitumor drug mithramycin (**34**) [25]. The latter two examples indicate a broad substrate specificity of glycosyltransferases.

#### 12.3.4 New Drugs by Variation of the Starter Units

*S. coeruleorubidus* produces different daunomycin (**35**) derivatives. Various blocked mutants have been prepared by UV and chemical mutation. One of these mutants produced the anthracyclines feodomycin A (**36**) and B (**37**) which were not produced by the wild-type strain. The structure of the new compounds indicated that the starter carbon units in the formation of these new compounds were acetate, butyrate, and acetoacetate instead of propionate which is used by the native proteins [26] (Fig. 12.10). Similar results have been obtained in others studies [27, 28] (see also Section 12.3.5.1.2).



**Figure 12.10.** Structures of daunomycin (**35**)-derivatives (**36**, **37**) produced by blocked mutants of *S. coeruleorubidus*.

#### 12.3.5 New Drugs by Recombinant Assembly of Enzymatic Subunits

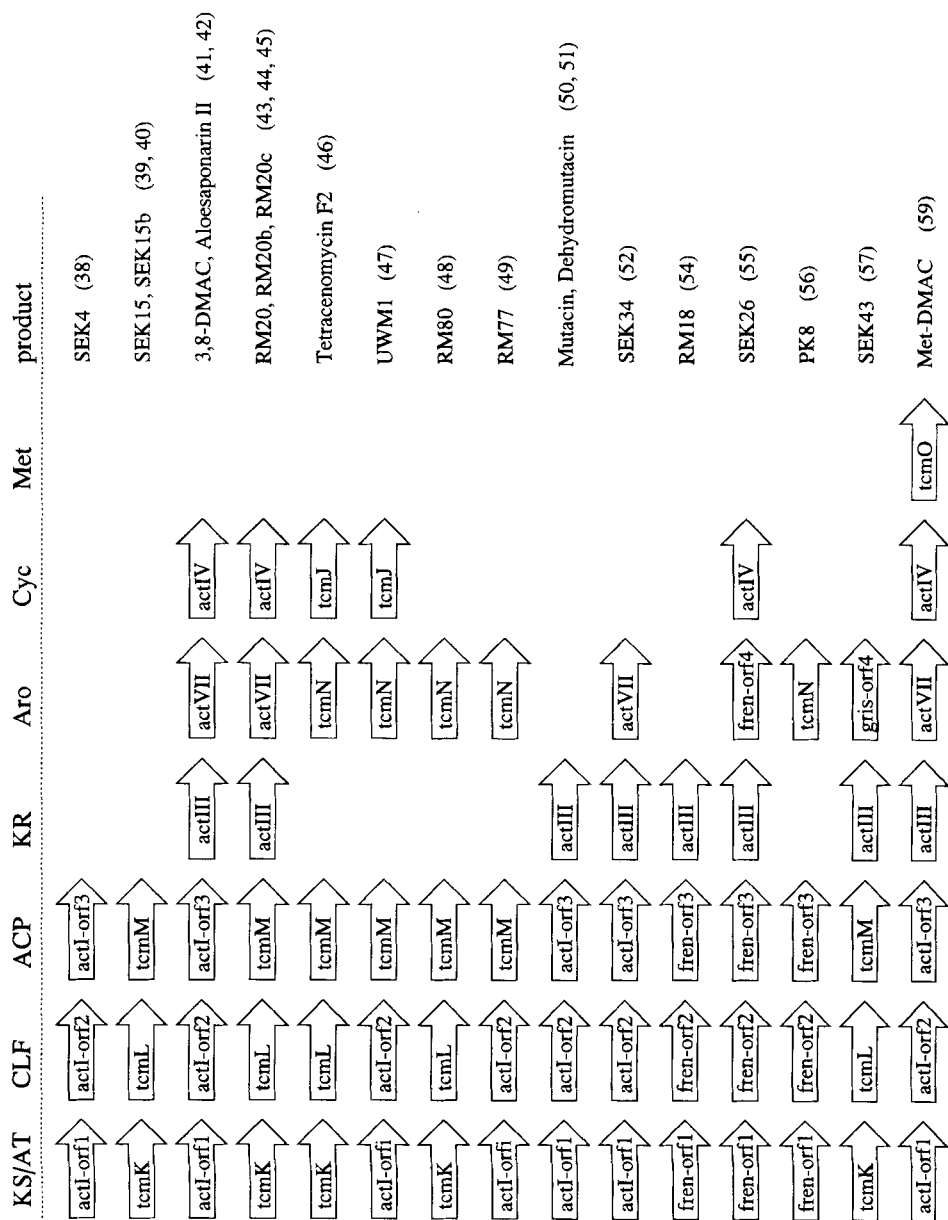
This chapter mainly summarizes the work of Professors D.A. Hopwood (Norwich, UK), C. Khosla (Stanford, USA), C.R. Hutchinson (Madison, USA), L. Katz (Abbott Park, USA) and P.F. Leadlay (Cambridge, UK), and their co-workers. These scientists developed systems for the expression of genes from different biosynthetic pathways in heterologous hosts, making possible the so-called ‘mix-and-match’ approach. This was leading to the biosynthesis of a number of new compounds, the structure of many of these being quite different from the parental compounds. This is an indication about the high potential of the mix-and-match approach for the development of new lead drugs in the future.

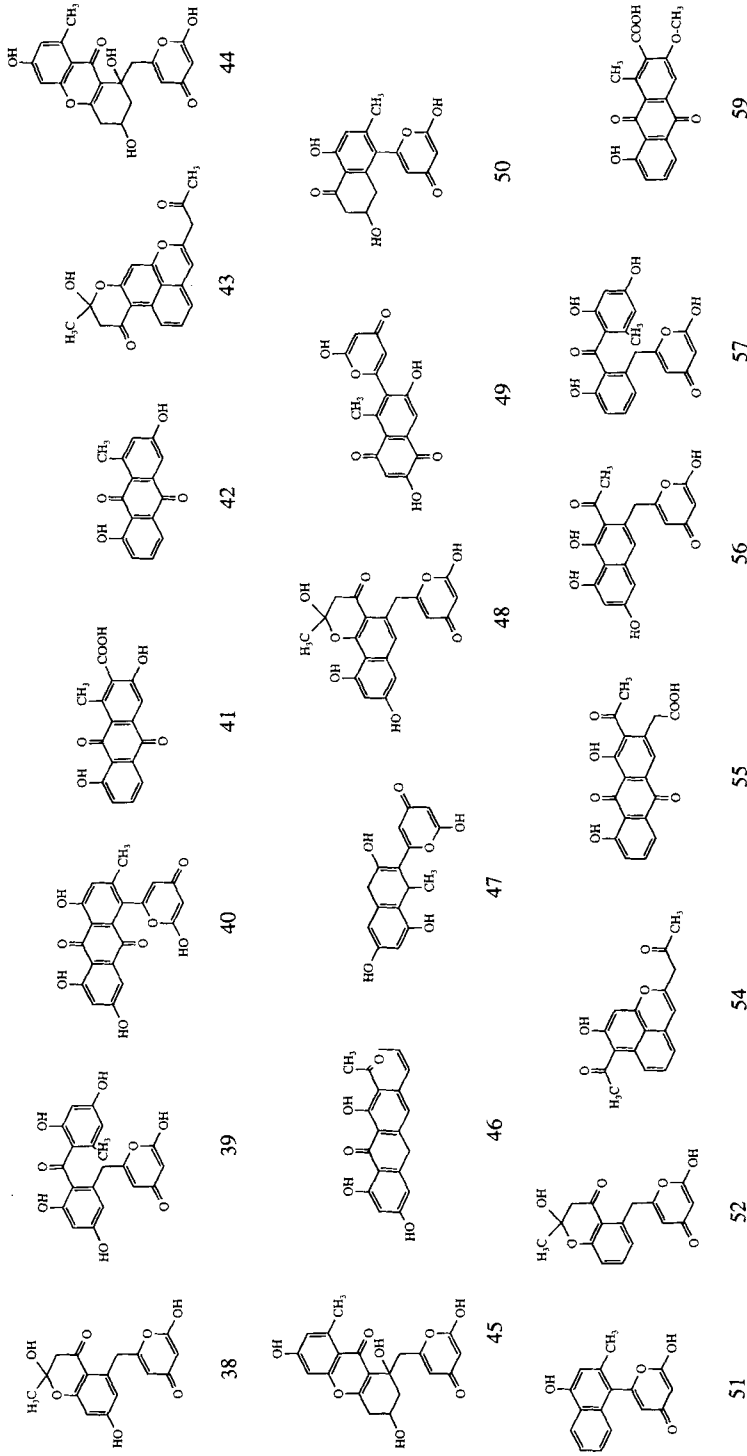
### 12.3.5.1 Polyketide Synthases

Excellent reviews have been written about polyketides and their biosynthesis [14, 29–34], and therefore we will only briefly describe the nature and characteristics of this important family of compounds. Polyketides are molecules that are synthesized as secondary metabolites following the onset of the stationary phase in the life cycle of different organisms. Many polyketides are important therapeutic agents, including antibiotics, anti-cancer drugs, immunosuppressants and antifungal products. Polyketides are synthesized by the action of polyketide synthases (PKSs) in a similar manner to fatty acid biosynthesis. In contrast to fatty acid biosynthesis, the natural pool of polyketide starter units includes acetate, propionate, malonate, butyrate, isobutyrate, and others. Two starter units derived from their corresponding CoA thioester are loaded onto the phosphopantetheine thiol of an acyl carrier protein (ACP). A decarboxylative condensation reaction between two acyl groups of the starter units results in the generation of an intermediate possessing a reactive  $\beta$ -carbonyl group. This group may either be left unreduced or converted into a hydroxyl, olefin or methylene group. The growing chain is transferred from the ACP to the reactive cysteine of the ketosynthase (KS) in preparation of another cycle of chain elongation. The PKS system determines the final length, the level of ketoreduction, the stereochemistry, and the cyclization pattern of the polyketide chain. PKSs have been classified into iterative and modular PKSs. Bacterial aromatic PKSs, which are composed of several monofunctional proteins, fungal or bacterial multifunctional PKSs and plant PKSs belong to the group of iterative PKSs while PKSs catalyzing the formation of macrolides, avermectins or rapamycin are modular PKSs. As dictated by the convention for fatty acid synthases, microbial PKSs have also been classified into two types, type I (modular PKSs) and type II (iterative PKSs). New drugs have mainly been produced using genes of the type II PKSs. This might be the case because information derived from DNA sequencing about type II PKSs predated the discovery of type I systems. The following section outlines recent advances in type II and I PKS research, and shows how PKS-encoding genes can be used to make novel natural products.

#### 12.3.5.1.1 Iterative Polyketide Synthases

The first two biosynthetic gene clusters characterized were those for actinorhodin (**1**) from *S. coelicolor* [4] and tetracenomycin (**2**) from *S. glaucescens* [5]. It was shown that the polyketide backbone of these two drugs was synthesized by an iterative PKS [35, 36]. Studies on the PKS gene clusters of these two pathways led to the discovery that the different PKSs of actinomycete polyketide producers were very similar [37]. This facilitated the isolation of many additional clusters from further strains (see 1). The iterative PKS of actinorhodin and tetracenomycin (and those of other iterative polyketide systems) consist of several genes coding for ketosynthase/acyltransferases (KS/AT), chain length factor proteins (CLF), ACP, cyclases (CYC), and aromatases (ARO). Three proteins, KS/AT, CLF and ACP are essential for polyketide synthesis and comprise the ‘minimal PKS’ [38]. Together with ARO, CYC and ketoreductases (KR), these proteins control the chain length, reductive state and the folding pattern of the polyketide backbone. Additional





**Figure 12.11.** (a) Recombinant sets carrying iterative polyketide synthase genes which were expressed in *S. coelicolor* CH999. KS(AT): ketosynthase/acyltransferase, CLF: chain length factor protein. ACP: acyl-carrier-protein. KR: ketoreduktase. Aro: aromatase. Cyc: cyclase. Met: methyltransferase; *act*: gene from the actinorhodin biosynthetic gene cluster. *cm*: gene from the tetracenomyacin biosynthetic gene cluster. *freni*: gene from the frenolin biosynthetic gene cluster. *gris*: gene from the griseusin biosynthetic gene cluster. (b) New compounds generated by recombinant assembly of enzymatic subunits. (38): SEK4, (39): SEK15, (40): SEK15b, (41): 3,8-DMAC, (42): albesaparin, (43): RM20, (44): RM20b, (45): RM20c, (46): tetracenomycin F2, (47): UWMI, (48): RM80, (49): RMT7, (50): mutacin, (51): dehydromutacin, (52): SEK34, (54): RM18, (55): SEK26, (56): PK8, (57): SEK43, (59): Met-DMAC.



genes encoding oxygenases, acyltransferases and glycosyltransferases act on the products of the PKSs and are known as 'tailoring enzymes'.

As mentioned above, iterative PKSs have been studied in detail, focusing especially on the biosynthesis of actinorhodin and tetracenomycin. Later, further genes and gene clusters have become available. In this section we describe how knowledge of the function of these enzymatic systems can be used to generate new molecules. The most important results are summarized in Figure 12.11a and b.

### Replacing ACPs and KSs with Heterologous Proteins

Heterologous KSs and ACPs appear to function nearly identically, as was shown by mutant complementation and targeted gene replacements [39, 40]. Mutations in the *actI-orf3* gene coding for the actinorhodin (**1**) ACP were complemented with the tetracenomycin (**2**), granaticin (**29**) or oxytetracycline (**5**) ACP genes. Similar results were obtained by complementing mutations in the *actI-orf1* gene coding by the equivalent KS genes of the granaticin and oxytetracycline pathways. These experiments indicate that ACPs and KSs are less valuable targets for the production of new compounds by combinatorial biosynthesis than other genes.

### Variation of the Chain Length

One of the most defining characteristics of a polyketide is the number of condensation events necessary for its biosynthesis. In contrast to the complementation experiments described for ACPs and KSs mutations, the *actI-orf2* gene (encoding the CLF component of the actinorhodin cluster) was not complemented by most heterologous genes of other clusters. To investigate the function of ActI-orf2 (and other CLFs), Hopwood, Khosla and co-workers developed a host-vector system for the regulated expression of actinorhodin genes in combination with other iterative PKS genes [38]. As a host, a mutant of *S. coelicolor* was generated in which all the actinorhodin cluster was deleted; this mutant is known as *S. coelicolor* CH999. As vector, the plasmid pRM5 was created containing the *actII-orf4* gene which codes for a positive regulator that induces transcription from the actinorhodin promoters during the transition from growth phase to stationary phase, a transcription terminator from phage fd and the divergent *actII/actIII* promoter pair. To facilitate replacement of PKS genes, the actinorhodin PKS genes were cloned in pRM5 as individual cassettes with flanking restriction sites and their own ribosome binding site. After transformation of this construct in *S. coelicolor* CH999, different metabolites were produced. Expression of the minimal PKS of the actinorhodin cluster (*actI-orfs 1, 2 and 3*) resulted in the accumulation of a compounds designated as SEK4 (**38**) [41], whereas after expression of the tetracenomycin minimal PKS (*tcmKLM*), SEK15 (**39**) and SEK15B (**40**) were produced [42]. In these cases the formation of six-membered aromatic rings was due to spontaneous aldol reaction, as no cyclase was present in the expression cassettes. In addition, KS, CLF and ACP dictate the regioselectivity of the first cyclization in the absence of a specific cyclase. The co-expression of *actI-orf1*, *actI-orf2*, *actI-orf3* with

*actIII*, *actIV* and *actVII* (the last three genes are required for the synthesis of an early intermediate in the actinorhodin biosynthesis) was leading to the production of 3,8-dihydroxymethylantraquinone carboxylic acid (3,8-DMAC) (**41**) and aloesaponarin II (**42**), shunt products of the actinorhodin pathway, while after expression of a cassette containing *tcmK*, *tcmL* and *tcmM* with *actIII*, *actIV* and *actVII*, the accumulation of compounds RM20 (**43**), RM20b (**44**) and RM20c (**45**) was observed [38]. While 3,8-DMAC contains 16 carbons, RM20b and RM20c contain 20 carbons, indicating the influence of *actI-ORF2* and *tcmL* on the chain length. Using a similar vector host system (null mutants of *S. glaucescens* were used as hosts and pWHM3 as vector), Hutchinson and co-workers expressed the *tcmJKLMN* (*tcmJ* and *tcmN* code for cyclases). This was leading to the accumulation of tetracenomyacin F2 (**46**) (a 20-carbon compound) [43, 44]. The same compound was produced after expression of *actI-orf1* with *tcmL*, *tcmM*, *tcmN* and *tcmJ*. Replacement of *tcmL* with *actI-orf2* (co-expression of *tcmL*, *actI-orf2*, *tcmM*, *tcmN* and *tcmJ*) resulted in the formation of tetracenomyacin F2 and UWM1 (**47**), a new 16-carbon compound. This compound predominated when both *tcmK* and *tcmL* were replaced by *actI-orf1* and *actI-orf2*, indicating a joint activity of the KSs and CLFs [31].

These and other experiments proposed the designation of the KS as chain length factor (CLF) determinant [38]. It must be mentioned that genes coding for CLFs can not be interchanged in all cases as some of the KS-CLF combinations are not functional.

#### Variation of the Folding Pattern – Influence of Aromatases and Cyclases

As mentioned above, the co-expression of *tcmK*, *tcmL*, and *tcmM* or *actI-orf1*, *actI-orf2* and *actI-orf3* with *tcmN* and *tcmJ* was leading to different cyclic compounds. Also, different products were produced when *tcmK*, *tcmL* and *tcmM* (product RM80 (**48**)) or *actI-orf1*, *actI-orf2* and *actI-orf3* (product RM77 (**49**)) were co-expressed only with *tcmN* [45]. From these results, the function of *tcmN* can be assigned as a second-ring aromatase.

The gene *actVII* is coding for an aromatase catalyzing the formation of ring A of actinorhodin. ActVII mutants of *S. coelicolor* accumulate the shunt products mutacin (**50**) and dehydromutacin (**51**). The same products were accumulated when *actI-ORF1*, *actI-ORF2*, *actI-ORF3* and *actIII* were expressed. In contrast the co-expression of *actI-ORF1*, *actI-ORF2*, *actI-ORF3*, *actIII* and *actVII* resulted in the formation of SEK34 (**52**) [46].

#### Influence of Ketoreduction on the Bioengineered Product

The influence of the ActIII KR on the final product has been studied in some detail. ActIII interacts with the bacterial PKSs and reduces the carbonyl that is nine carbons from the carboxyl end of the poly- $\beta$ -ketone intermediates; this results in the absence of a hydroxyl at the corresponding position of the polyketide. The reduction influences the cyclization pattern. When *actIII* was expressed together with *tcmK*, *tcmL* and *tcmM*, RM20b (**44**) was produced. In the absence of *actIII*, SEK15 (**39**) was the main accumulated product (see above).

## Designing New Molecules using Genes of Different Gene Clusters

The potential use of biosynthetic genes to create new molecules excited scientists to clone further biosynthetic gene clusters. Some of these genes have been used extensively to generate new compounds because of their relaxed substrate specificity. A few examples are given to describe the potential use of these genes in generating new polyketides. Several studies had revealed that the frenolicin (**53**) PKS is capable of generating 16- and 18-carbon-membered polyketides in an approximately 1:2 ratio. First, it was speculated that the frenolicin minimal PKS showed relaxed chain length specificity. Later, it was discussed that KRs and CYCs can modulate the intrinsic specificity of the minimal PKS with respect to both the folding pattern and the chain length of the final product. This was demonstrated by expressing the minimal PKS of the frenolicin biosynthetic gene cluster together with a number of different genes of the actinorhodin- and tetracenomycin- biosynthetic gene cluster in *S. coelicolor* CH999. These experiments were leading to the production of new compounds, namely RM18 (**54**), SEK26 (**55**) and PK8 (**56**). The production of SEK26 is especially worthwhile in mentioning because its biosynthesis is the result of the rationally designed combination of four components, the minimal PKS and the aromatase from the frenolicin cluster and a KR and a cyclase from the actinorhodin cluster. A further rationally designed polyketide is SEK43 (**57**); this is produced after expressing several genes of the tetracenomycin- (*tcmK*, *tcmL*, *tcmM*) and actinorhodin-cluster (*actIII*) with an aromatase of the griseusin (**58**) cluster. This aromatase shows broad substrate specificity as it recognizes chains of 16, 18 and 20 carbons [47].

In further experiments, *tcmO* – a methyltransferase of the tetracenomycin cluster that methylates the C8-hydroxyl of tetracenomycin B3, an intermediate in the biosynthesis of tetracenomycin – was used to generate the new polyketide 3-O-Methyl-DMAC (**59**). Again, this experiment demonstrated that the variable substrate specificity of an enzyme increases its potential use in creating new molecules.

### 12.3.5.1.2 Modular Polyketide Synthases

Modular PKSs are large multifunctional enzymes. Active sites (domains) within these enzymes (ketosynthases, KS; acyltransferases, AT; dehydratases, DH; enoyl reductases, ER; ketoreductases, KR; acyl carrier proteins, ACP; and thioesterases, TE) are organized into modules such that each module catalyzes the stereospecific addition of a new monomer onto a growing polyketide chain, and also sets the reduction level of the carbon atoms of the resulting intermediate [33, 34]. Among the type I systems, the 6-deoxyerythronolide B synthase (DEBS) from *Saccharopolyspora erythraea* has been studied in more detail. It catalyzes the biosynthesis of 6-deoxyerythronolide B (6-dEB) (**60**), the polyketide moiety of the macrolide antibiotic erythromycin (**4**). DEBS consists of three large enzymes (EryAI, EryAII, EryAIII), each with a molecular weight >300 kDa. Each enzyme contains two modules which comprise a different number of active sites. The expression of the complete DEBS in *S. coelicolor* CH999 was achieved in 1994. The recombinant strain produced 6-dEB in relatively large quantities (>40 mg/l) [48] (Fig. 10.12).

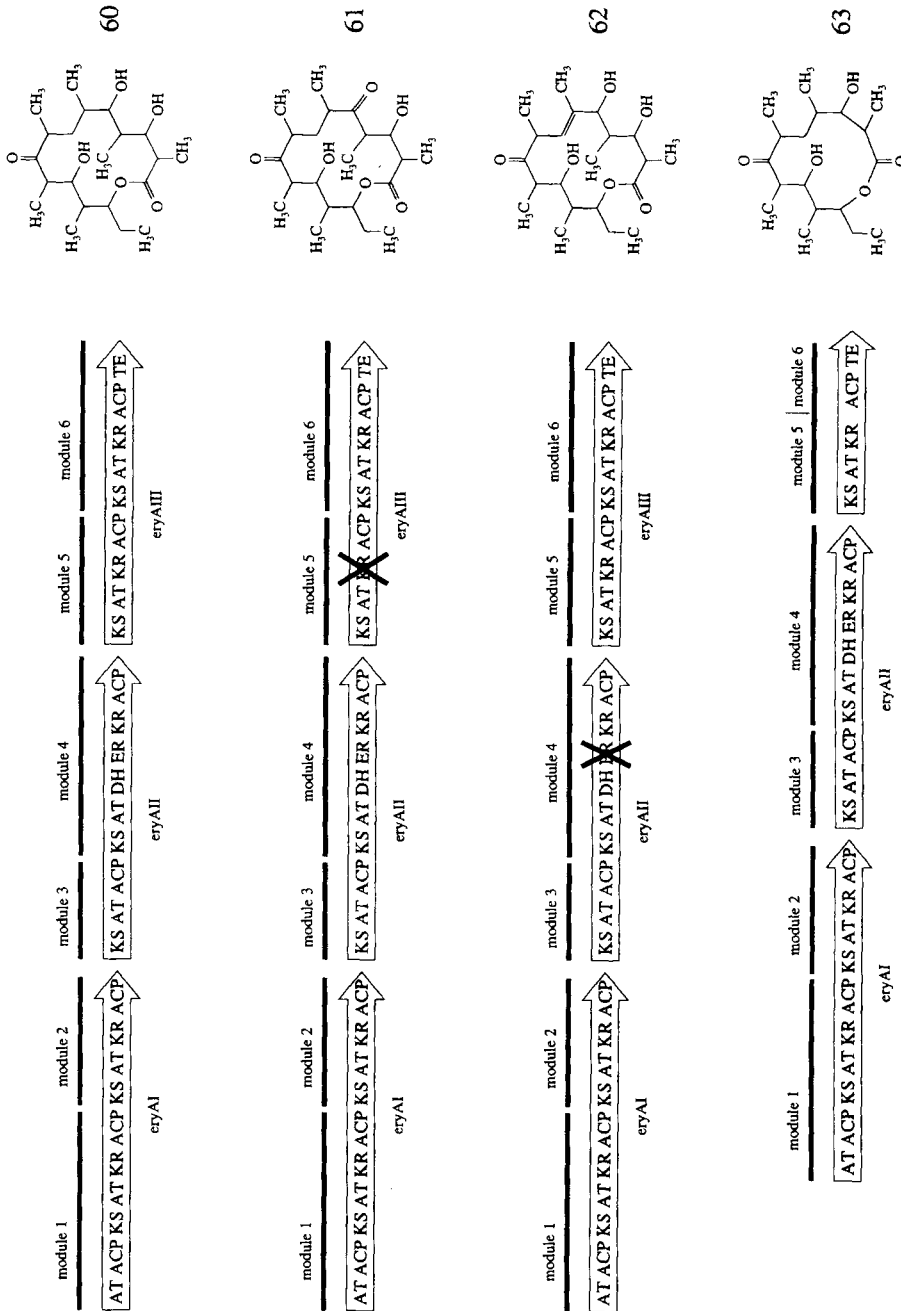


Figure 12.12. New compounds produced after deletion of domains of DEBS from *Saccharopolyspora erythraea*.

The production of modified full-length derivatives of erythromycin has been achieved first by gene disruption experiments (see Section 12.3.1). Recent success in engineering new compounds via deletion of modules and domains, expression of single enzymes (e.g., EryAI) and domain replacement have been described.

#### Deletion of Modules and Domains

Mutants of *Saccharopolyspora erythraea* were generated carrying *in frame* deletions in the KR domain of module 5 or in the ER domain of module 4. These mutants were producing a predicted keto-derivative (**61**) (KR mutant) and a predicted anhydro-derivative (**62**) (ER mutant) of 6-dEB [49, 50]. A deletion mutant of DEBS was constructed containing EryAI, EryAII and a chimeric fifth module with KS, AT and KR of module 5 of EryAIII and ACP and TE of module 6. This modified DEBS was transformed into *S. coelicolor* CH999, leading to the production of the new macrolactone 8R, 9S-8,9-dihydro-8-methyl-9-hydroxy-10-deoxymethynolide (**63**) [48]. The formation of this 12-membered macrolactone instead of a 14-membered lactone as 6-dEB indicates that the deletion of domains or modules can lead to the production of novel macrolides with changes in the ring system that increase the chance to produce novel lead structures by combinatorial biosynthesis (Fig. 12.12).

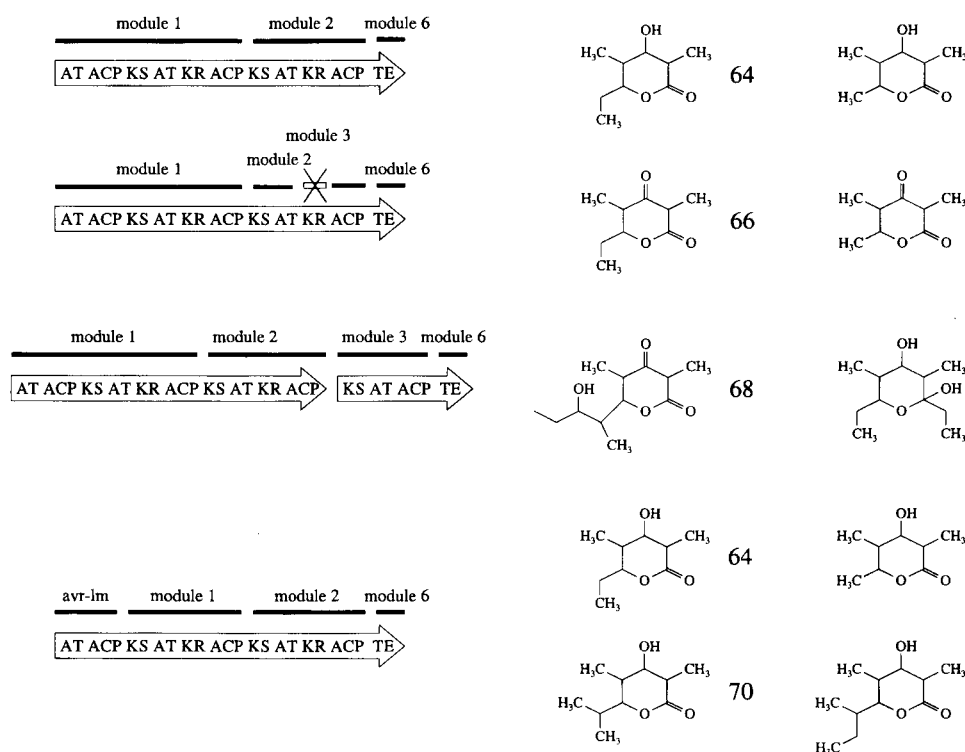
#### Expression of Single Enzymes

The expression of a modified version of EryAI, which also contains the TE domain of EryAIII, in a heterologous host was leading to the production of two triketide lactones (**64**, **65**) [51]. After replacement of the functional KR domain in module 2 with a non functional KR from module 3 and expression of this gene in the same host, the production of two different triketide ketolactones (**66**, **67**) was observed [52]. Tetra-ketide lactones (CK13a (**68**) and CK13b (**69**)) were obtained after expression of a protein consisting of EryAI and module 3 of EryAII containing the TE domain of EryAIII.

After expressing a hybrid gene consisting of the loading domain (AT and ACP) of the avermectin (**73**) producer, modules 1 and 2 of EryAI and TE of EryAIII different trice-tide lactones (**66**, **67**, **70**, **71**) were produced. The relaxed substrate specificity of the loading domain from the avermectin producing PKS of *S. avermitilis*, which accepts branched carboxylic acids as starter units, was influencing the formation of the new compounds [53] (Fig. 12.13).

#### Domain Replacement

The loading domain of the erythromycin PKS (AT and ACP) is not essential for erythromycin biosynthesis; however, mutants lacking the domain produce erythromycin at much lower levels than the wild-type strain [54].



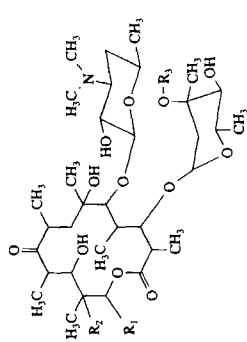
**Figure 12.13.** New compounds produced after expression of modified versions of *eryAI* from *Saccharopolyspora erythraea* in a heterologous strain.

The accumulation of novel erythromycin-derivatives has been reported after replacing the 'loading domains' (AT and ACP) of module 1 of DEBS with the 'loading domains' from the avermectin producer. As expected, and again due to the broad substrate specificity of AT and ACP from *S. avermitilis*, the hybrid enzyme was producing a number of new antibiotics (**72-80**) [53] (Fig. 12.14).

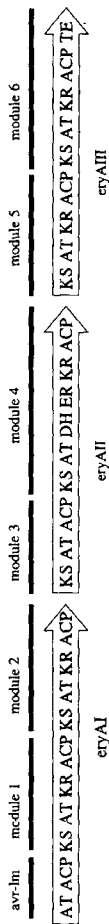
In a similar approach, the 'loading domains' AT and ACP of the spiramycin (**6**) PKS were replaced by those of the tylosin (**7**) PKS. As the new loading domains accept propionyl-CoA instead of acetyl-CoA as substrate, a new hybrid polyketide (**81**) was produced [55].

### 12.3.5.2 Peptide Synthetases

Peptide biosynthesis may occur through two different systems. Most of cellular peptides and proteins are produced by the ribosomal machinery connecting 20 proteinogenic amino acids to the desired products. However, most of the bioactive peptides are produced non-ribosomally by large peptide synthetases. These peptides are used as antibiotics, enzyme inhibitors, toxins, and other medically useful drugs. The biosynthesis of



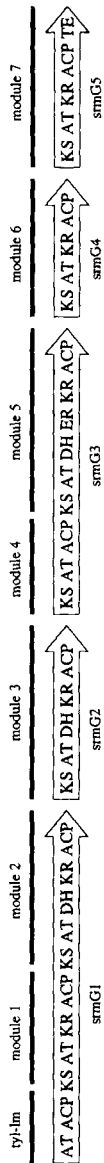
R <sub>1</sub> :	R <sub>2</sub> :	R <sub>3</sub> :	compound
H	OH	CH <sub>3</sub>	72
-C-CH <sub>3</sub>	H	CH <sub>3</sub>	73
H	H	H	74
H H	OH	CH <sub>3</sub>	75
-C-C-CH <sub>3</sub>	H	CH <sub>3</sub>	76
H H	H	H	77
H H	OH	CH <sub>3</sub>	78
-C-C-CH <sub>3</sub>	H	CH <sub>3</sub>	79
H <sub>3</sub> C H	H	H	80



eryAI

eryAII

eryAIII



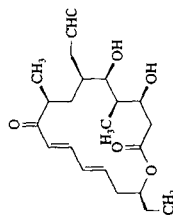
srmG1

srmG2

srmG3

srmG4

srmG5



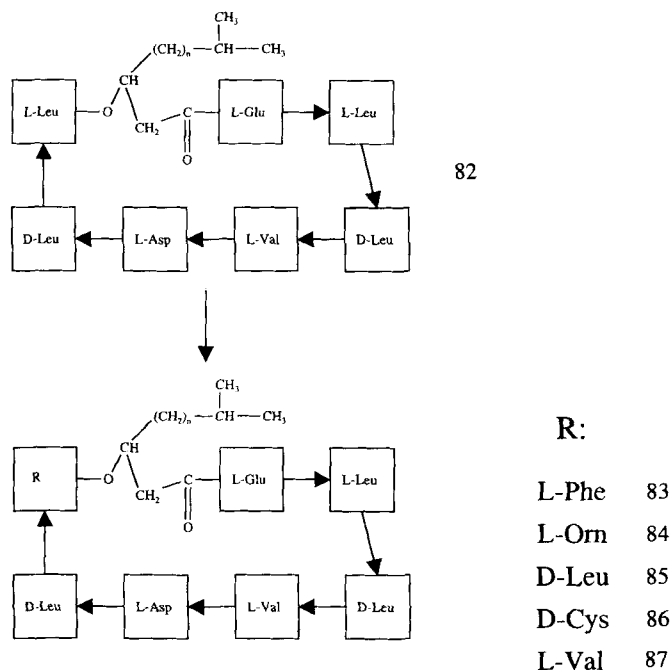
81

**Figure 12.14.** Production of new macrolides after replacing the loading domains of the erythromycin and spiramycin polyketide synthases with domains of the avermectin and tylosin polyketide synthases.

non ribosomal peptides begins in general with the enzyme-catalyzed activation of amino acids. An ATP-dependent amino-acyladenylation step is followed by a thioesterification of the amino acid on a prosthetic 4'-phosphopantethein group, that is bound to the enzyme. One peptide synthetase consists of several domains each necessary for the binding of one amino acid. Each domain catalyzes not only the binding of the amino acid but also modifications (racemization, N-methylation) as well as the formation of peptide bonds as each domain contains a peptidyl acceptor and a peptidyl donor site. These sites facilitate the ordered shift of the amino acids between the domains. To terminate the process, the thioester bond connecting the peptide product and the enzyme is cleaved by hydrolysis and cyclization or transfer to a functional group. The specification of a peptide sequence is determined by the number and order of domains in one polypeptide synthetase [56, 57].

### Novel Peptides by Gene Replacement

*Bacillus subtilis* produces the cyclic peptide antibiotic surfactin (82). This is an acyl-peptidolactone composed of a  $\beta$ -hydroxy fatty acid side chain and seven amino acid residues, including a leucine at position 7. The biosynthesis of surfactin is catalyzed by a peptide synthetase consisting of three modules. The first and second modules are carrying three domains, the last module just one domain. The last domain catalyzes the activation and



**Figure 12.15.** Production of new surfactin antibiotics after replacing a L-leucine domain in *Bacillus subtilis* with domains from other organisms.



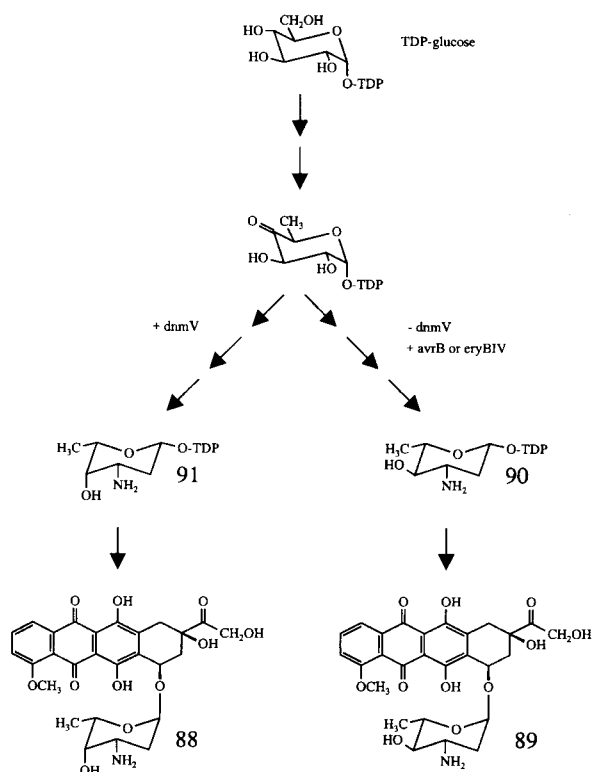
connection of D-leucine into the molecule. The targeted replacement of this domain has been carried out successfully for several bacterial and fungal domains. It could be replaced by a L-phenylalanine, a L-ornithine and a D-leucine domain of the gramicidin S peptide synthetase from *Bacillus brevis* and a D-cysteine and a L-valine domain of the ((L-aminoadipyl)-cysteinyl)-D-valine synthetase from *Penicillium chrysogenum*, creating a number of new peptide antibiotics (**83-87**) [58] (Fig. 12.15).

### 12.3.5.3 Proteins Involved in Deoxysugar Biosynthesis

Carbohydrates are the most abundant group of natural compounds, and the role of sugars and deoxysugars as energy resources (e.g., glycolysis) and key structure elements in the formation of biological backbones (e.g., 2-deoxyribose for DNA) is generally known [59, 60]. However, they are also important constituents of natural compounds. It has been shown for some antibiotics and antitumor compounds that the sugar moiety is contributing as recognition element to the mechanism of action of the respective drug [61–63]. A large number of pharmaceutical useful compounds contain deoxysugars, those carbohydrates in which one or more of the (normally) occurring oxygen atoms has been replaced by hydrogen or any other heteroatom or heteroatomic group. The requirement of a number of genes from different biosynthetic gene clusters for deoxysugar formation has been shown. The possible function of many genes has been assigned by comparing the deduced protein sequence with sequences of known proteins in the data base. In addition a few genes have been expressed in heterologous hosts and in these cases proteins were purified and chemically characterized. These have led to hypothetical assignments of some steps in deoxysugar biosynthesis [64, 65].

### Novel Saccharide Antibiotics by Gene Replacement Experiments

Recently, Hutchinson and co-workers were able to demonstrate that a valuable hybrid antibiotic can be made by combinatorial biosynthesis with deoxysugar biosynthesis genes. *S. peuceitius* produces the anticancer drug, doxorubicin (**88**), which can be converted to the clinically more useful antitumor drug epirubicin (**89**) by a costly seven-step chemical method. Epirubicin contains the deoxysugar 4-epidaunosamine (**90**) instead of daunosamine (**91**) (Fig. 12.16). The gene cluster for doxorubicin biosynthesis has been cloned, and the function of several genes assigned. It had been speculated that a 4-ketoreductase (DnmV) is involved in the biosynthesis of TDP-daunosamine, catalyzing the stereoselective conversion of a hypothetical TDP-4-keto-6-deoxy sugar intermediate to the final deoxysugar. Reduction of a C-4 keto intermediate is also required for mycarose and oleandrose biosynthesis, but with stereoselectivity opposite. EryBIV and AvrE are involved in the biosynthesis of these deoxysugars as components of erythromycin and avermectin respectively. *EryBIV* and *avrE* genes were introduced into a *dnmV* mutant of *S. peuceitius*. Transformed strains produced epirubicin (4'-epidoxorubicin) instead of doxorubicin (Fig. 12.16), indicating the potential use of deoxysugar biosynthetic genes in the formation of hybrid antibiotics [66].



**Figure 12.16.** Production of epirubicin instead of doxorubicin after replacing the gene *dnmV* from *S. peucetius* with *avrB* from *S. avermitilis* or *eryBIV* from *Saccharopolyspora erythraea*.

## 12.4 Future Perspectives

### 12.4.1 Use of Genes Involved in the Biosynthesis of Oligosaccharide Antibiotics

So far, glycosylated hybrid antibiotics have been produced more or less by introducing one sugar moiety to a given compound. Many antibiotics consist of more than one sugar and, as outlined in Section 12.3.5.3, these sugars are contributing to the mechanism of action of these compounds. In 1994, the authors of this chapter initiated investigations on the genetic analysis of avilamycin (**92**), mitramycin (**34**), landomycin (**93**) and urdamycin (**18**) biosynthesis. All these antibiotics consist of polyketide- and sugar-moieties. More than 80 biosynthetic genes have been cloned and sequenced in the biosynthetic gene clusters, and several aspects about the function of genes have been determined [67–69] (see also Sections 12.3.2 and 12.3.3). We are now focusing on the elucidation of the function of glycosyltransferase genes in the clusters, and are testing the potential of these genes to create new molecules.

### 12.4.2 Use of Genes from Other Sources – New Compounds from Earth

Since less than 1 % of the micro-organisms present in soil are cultivatable by standard methods, the discovery of novel microbial secondary metabolites is restricted to this small portion. The isolation of total DNA from soil, and use of this DNA in combinatorial biosynthesis, is likely to be a valuable approach in overcoming this limitation. Recently, Prof. C.R. Hutchinson, Prof. J. Davies and co-workers amplified chain length factor protein genes from soil using PCR primers derived from conserved regions of known ketosynthase and ACP genes. Expression of these genes, together with genes involved in polyketide biosynthesis of tetracenomycin and actinorhodin, led to the production of different octaketides and decaketides [70].

### 12.4.3 Changing the Substrate Specificity of an Enzyme

The function of a protein can be changed by multiple rounds of mutagenesis and functional screening – a process named DNA shuffling [71]. This promising strategy to create enzymes with altered functions was first described to improve the activity of a protein [72], but was used recently also to change the specificity of an enzyme [73]. Although this approach was used initially to change the function of biosynthetic genes, it must be expected that using genes modified by gene shuffling will, in the near future, lead to the production of new antibiotics.

## 12.5 References

- [1] Feitelson, J.S. and D.A. Hopwood (1983). Cloning of a streptomycetes gene for an O-methyltransferase involved in antibiotic biosynthesis. *Mol. Gen. Genet.* 190: 394–398.
- [2] Gil, J.A. and D.A. Hopwood (1983). Cloning and expression of a p-aminobenzoic acid synthetase gene of the candidicin-producing *Streptomyces griseus*. *Gene* 25: 119–132.
- [3] Jones, G. and D.A. Hopwood (1984). Molecular cloning and expression of the phenoxazinone synthase gene from *Streptomyces antibioticus*. *J. Biol. Chem.* 259: 2703–2714.
- [4] Malpartida, F. and D.A. Hopwood (1984). Molecular cloning of the whole biosynthetic pathway of a *Streptomyces antibioticus* and its expression in a heterologous host. *Nature* 309: 462–464.
- [5] Motamedi, H. and C.R. Hutchinson (1987). Cloning and heterologous expression of a gene cluster for the biosynthesis of tetracenomycin C, the anthracycline antitumor antibiotic of *Streptomyces glaucescens*. *Proc Natl Acad Sci U S A* 84: 4445–4449.
- [6] Rohr, J. 1995. Kombinatorische Biosynthese – eine Zukunftstrategie. *Angew Chem* 107: 963–966.
- [7] Khosla, C. and R.J.X. Zawada (1996) Generation of polyketide libraries via combinatorial biosynthesis. *Trends in Biotechnology* 14: 335–341.
- [8] Ikeda, H., S. Omura. (1997). Avermectin biosynthesis. *Chem. Rev.* 97: 2591–2609.
- [9] Thompson, C.J., J.M. Ward, D.A. Hopwood. (1980). DNA cloning in *Streptomyces*: Resistance genes from antibiotic-producing species. *Nature* 286: 525–527.
- [10] Staunton, J., B. Wilkinson (1997). Biosynthesis of erythromycin and rapamycin. *Chem. Rev.* 97: 2611–2629.
- [11] Kim, E.S., D.A. Hopwood, D.H. Sherman (1994). Sequences of the oxytetracycline polyketide synthase-encoding *otc* genes from *Streptomyces rimosus*. *Gene* 141: 141–142.
- [12] Kuhstoss, S., M. Huber, J. R. Turner, J. W. Paschal, R.N. Rao (1996). A new shuttle cosmid vector, pKC505, for streptomycetes: its use in the cloning of three different spiramycin-resistance genes from a *Streptomyces ambofaciens* library. *Gene* 183: 231–241.

- [13] Fishman, S.E., K. Cox, J. L. Larson, P. A. Reynolds, E. T. Seno, W. K. Yeh, R. van Frank, C.L. Hershberger (1987). Cloning genes for the biosynthesis of a macrolide antibiotic. *Proc. Natl. Acad. Sci. USA* 84: 8248–8252.
- [14] Hopwood, D.A. (1997). Genetic contribution to understanding polyketide synthases. *Chem. Rev.* 97: 2465–2497.
- [15] Weber, J.M., J.O. Leung, S.J. Swanson, K.B. Idler and J.B. McAlpine (1991). An erythromycin derivative produced by targeted gene disruption in *Saccharopolyspora erythraea*. *Science* 252: 114–117.
- [16] Motamedi, H., A. Shafiee, S.-J. Cai, S.L. Streicher, B.H. Arison and R.R. Miller (1996). Characterization of methyltransferase and hydroxylase genes involved in the biosynthesis of the immunosuppressants FK506 and FK520. *J. Bacteriol.* 178: 5243–5248.
- [17] Epp, J.K., M.L.B. Huber, J.R. Turner, T. Goodson and B.E. Schoner (1989). Production of a hybrid macrolide antibiotic in *Streptomyces ambofaciens* and *Streptomyces lividans* by introduction of a cloned carbomycin biosynthetic gene from *Streptomyces thermotolerans*. *Gene* 85: 293–301.
- [18] Hara, O., and C.R. Hutchinson (1990). A macrolide 3-O-acyltransferase gene from the midecamycin-producing *Streptomyces mycarofaciens*. *J. Bacteriol.* 174: 5141–5144.
- [19] Decker, H. and S. Haag (1995). Cloning and characterization of a polyketide synthase gene from *Streptomyces fradiae* Tü2717, which carries the genes for biosynthesis of the angucycline antibiotic urdamycin A and a gene probably involved in its oxygenation. *J. Bacteriol.* 177: 6126–6136.
- [20] Solenberg, P.J., P. Matsushima, D.R. Stack, S.C. Wilkie, R.C. Thompson and R. H. Baltz (1997). Production of hybrid glycopeptide antibiotics in-vitro and in *Streptomyces toyocaensis*. *Chem & Biol.* 4: 195–202.
- [21] Sezonov, G., V. Blanc, N. Bamas-Jacques, A. Friedmann, J.L. Pernodet and M. Guerineau (1997). Complete conversion of antibiotic precursor to pristinamycin IIA by overexpression of *Streptomyces pristinaespiralis* biosynthetic genes. *Nature Biotechnol.* 15: 349–353.
- [22] Hong, S.-T., J.R. Carney and S.J. Gould (1997). Cloning and heterologous expression of the entire gene clusters for PD 116740 from *Streptomyces* strain WP4669 and tetrangulol and tetrangomycin from *Streptomyces rimosus* NRRL3016. *J. Bacteriol.* 179: 470–476.
- [23] Ichinose, K., D.J. Bedford, D. Tornus, A. Bechthold, M.J. Bibb, W. P. Reville, H. G. Floss, D. A. Hopwood (1998). The granaticin biosynthetic gene cluster of *Streptomyces violaceoruber* Tü22: sequence analysis and expression in a heterologous host. Manuscript in preparation.
- [24] Decker, H., J. Rohr, H. Motamedi, H. Zähler and C.R. Hutchinson (1995). Identification of *Streptomyces olivaceus* Tü2353 genes involved in the production of the polyketide elloramycin. *Gene* 166: 121–126.
- [25] Wohler S.-E., Blanco G., Lombo F., Udvarnoki G., Mendez C., Decker H., Salas J.A., Rohr J. (1998) Novel hybrid tetracenomycins through combinatorial biosynthesis using a glycosyltransferase encoded by the *erm*-Genes in Cosmid 16F4 which shows a very broad sugar substrate specificity. Manuscript in preparation.
- [26] Oki, T., Matsuzawa, Y., Kiyuoshima, K. and A. Yoshimoto (1981) New anthracyclines, feudomycins, produced by the mutant from *Streptomyces coeruleorubiscus* ME130-A4. *J. Antibiot.* 34: 783–790.
- [27] Fu, H., S. Ebert-Khosla, D.A. Hopwood and C. Khosla (1994b). Relaxed specificity of the tetracycline polyketide synthase for an acetate primer in the absence of a malonamyl primer: *J. Am. Chem. Soc.* 116: 6443–6444.
- [28] Dutton, C.J., S.P. Gibson, A.C. Goudie, K.S. Holdom, M.S. Pacey, J.C. Ruddock, J.C. J.D. Bullock, and M.K. Richards (1991). Novel avermectins produced by mutational biosynthesis. *J. Antibiot.* 44: 357–365.
- [29] Hopwood, D.A. and D.H. Sherman (1993). Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu. Rev. Genet.* 24: 37–66.
- [30] Katz, L. and S. Donadio (1993). Polyketide synthesis: prospects for hybrid antibiotics. *Annu. Rev. Microbiol.* 47: 875–892.
- [31] Hutchinson, C.R. and I. Fujii (1995). Polyketide synthase gene manipulation: a structure-function approach in engineering novel antibiotics. *Annu. Rev. Microbiol.* 49: 201–238.
- [32] Hutchinson, C.R. (1997). Biosynthetic studies of daunorubicin and tetracenomycin C. *Chem. Rev.* 97: 2525–2535.
- [33] Katz, L. (1997). Manipulation of modular polyketide synthases. *Chem. Rev.* 97: 2557–2575.
- [34] Khosla, C. (1997). Harnessing the biosynthetic potential of modular polyketide synthases. *Chem. Rev.* 97: 2577–2590.

- [35] Bibb, M.J., S. Biro, H. Motamedi, J.F. Collins and C.R. Hutchinson (1989). Analysis of the deduced nucleotide sequence of the *Streptomyces glaucescens tcmI* genes provide key information about the enzymology of polyketide antibiotic biosynthesis. *EMBO J.* 8: 2727–2736.
- [36] Fernández-Moreno, M.A., E. Martínez, L. Boto, D.A. Hopwood and F. Malpartida (1992). Nucleotide sequence and deduced functions of a set of cotranscribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin. *J. Biol. Chem.* 267: 19278–19290.
- [37] Malpartida, F., S.E. Hallam, H.M. Kieser, H. Motamedi, C.R. Hutchinson, M.J. Butler, D.A. Sudgen, M. Warren, C. McKillop, C.R. Bailey, G.O. Humphrey and D.A. Hopwood (1987). Homology between *Streptomyces* genes coding for synthesis of different polyketides used to clone antibiotic biosynthetic genes. *Nature* 325: 818–821.
- [38] McDaniel, R., S. Ebert-Khosla, D.A. Hopwood and C. Khosla (1993). Engineered biosynthesis of novel polyketides. *Science* 262: 1546–1550.
- [39] Sherman, D., E.S. Kim, M.J. Bibb and D.A. Hopwood (1992). Functional replacement of genes for individual polyketide synthase components in *Streptomyces coelicolor* A3(2) by heterologous genes from a different polyketide pathway. *J. Bacteriol.* 174: 6184–6190.
- [40] Khosla, C. (1992) Targeted gene replacements in a *Streptomyces* polyketide synthase gene cluster: role for the acyl carrier protein. *Mol. Microbiol.* 21: 3237–3249.
- [41] McDaniel, R., S. Ebert-Khosla, H. Fu, D.A. Hopwood and C. Khosla (1994a). Engineered biosynthesis of novel polyketides: influence of downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. *Proc. Natl. Acad. Sci. USA* 91: 11 542–11 546.
- [42] Fu, H., S. Ebert-Khosla, D.A. Hopwood and C. Khosla (1994). Engineered biosynthesis of novel polyketides: dissection of the catalytic specificity of the act ketoreductase. *J. Am. Chem. Soc.* 116: 4166–4170.
- [43] Summers, R.G., E. Wendt-Pienkowsky, H. Motamedi, and C.R. Hutchinson. CR. 1992. Nucleotide sequence of the *tcmII-tcmIV* region of the tetracenomycin biosynthetic gene cluster of *Streptomyces glaucescens* and evidence that the *tcmN* gene encodes a multifunctional cyclase-dehydratase-O-methyltransferase. *J. Bacteriol.* 174: 1810–1820.
- [44] Summers, R.G., E. Wendt-Pienkowsky, H. Motamedi, C.R. Hutchinson, CR. 1993. The *tcmVI* region of the tetracenomycin C biosynthetic gene cluster of *Streptomyces glaucescens* encodes the tetracenomycin F1 monooxygenase, tetracenomycin F2 cyclase, and most likely, a second cyclase. *J. Bacteriol.* 175: 7571–7580.
- [45] McDaniel, R., C.R. Hutchinson, and C. Khosla (1995) Engineered biosynthesis of novel polyketides: analysis of *tcmN* function in tetracenomycin biosynthesis. *J. Am. Chem. Soc.* 117: 6805–6810.
- [46] McDaniel, R., S. Ebert-Khosla, D.A. Hopwood, and C. Khosla (1994b). Engineered biosynthesis of novel polyketides: *actVII* and *actIV* genes encode aromatase and cyclase enzymes, respectively. *J. Am. Chem. Soc.* 116: 10855–10859.
- [47] McDaniel, R., S. Ebert-Khosla, D.A. Hopwood, and C. Khosla (1995b). Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. *Nature* 375: 549–554.
- [48] Kao, C.M., G. Luo, L. Katz, D. E. Cane, and C. Khosla (1995). Manipulation of macrolide ring size by directed mutagenesis of a modular polyketide synthase. *J. Am. Chem. Soc.* 117: 9105–9106.
- [49] Donadio, S., M. J. Staver, J. B. McAlpine, S. J. Swanson, and L. Katz (1991). Modular organization of genes required for complex polyketide biosynthesis. *Science* 252: 675–679.
- [50] Donadio, S., J.B. McAlpine, P.J. Sheldon, M. Jackson, and L. Katz (1993). An erythromycin analog produced by reprogramming of polyketide synthesis. *Proc. Natl. Acad. Sci. USA* 90: 7119–7123.
- [51] Cortes, J., K. E. H. Wiesmann, G. A. Roberts, M. J. B. Brown, J. Staunton, and P. F. Leadlay (1995). Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage. *Science* 268: 1487–1489.
- [52] Bedford, D., J. R. Jacobsen, G. Luo, D. E. Cane, and C. Khosla. (1996) A functional chimeric modular polyketide synthase generated via domain replacement. *Chem. Biol.* 3: 827–831.
- [53] Andrew, F., A. Marsden, B. Wilkinson, J. Cortes, N.J. Dunster, S. Staunton, and P.F. Leadlay (1998). Engineering broader specificity into an antibiotic-producing polyketide synthase. *Science* 279: 199–201.
- [54] Pereda, A., R. G. Summers, D. L. Stassi, X. Ruan, and L. Katz (1998). The loading domain of the erythromycin polyketide synthase is not essential for erythromycin biosynthesis in *Saccharopolyspora erythraea*. *Microbiol.* 144: 543–553.
- [55] Kubstoss, S., M. Huber, J. R. Turner, J.W. Paschal, and R.N. Rao (1996). Production of a novel polyketide through the construction of a hybrid polyketide synthase. *Gene* 183: 231–236.

- [56] Marahiel, M.A., T. Stachelhaus, and H.D. Mootz (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. 1997. Chem. Rev. 97: 2651–2673.
- [57] Von Döhren, H., U. Keller, J. Vater, and R. Zocher (1997). Multifunctional peptide synthetases. Chem. Rev. 97: 2675–2705.
- [58] Stachelhaus, T., A. Schneider, M.A. Marahiel (1995). Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. Science 269: 69–72.
- [59] Streyer, L. (1988). Biochemistry. Freeman and Company, New York.
- [60] Metzler, D.E. (1977). Biochemistry – The chemical reactions of living cells. Academic Press, New York.
- [61] Moore, M.H., W.N. Hunter, B.L. d'Estaintot, and O. Kennard (1989). DNA-drug interactions. The crystal structure of d(CGATCG) complex with daunomycin. J. Mol. Biol. 206: 693–705.
- [62] Sun, D., M. Hansen, J.J. Clement, and L.H. Hurley (1993). Structure of the altromycin B (N7-guanine)-DNA adduct. A proposed prototypic DNA adduct structure for the pluramycin antitumor antibiotics. Biochemistry 32: 8068–8074.
- [63] Liu, C., B.M. Smith, K. Ajito, H. Komatsu, L. Gomezpaloma, T.H. Li, E.A. Theodorakis, K.C. Nicolaou, and P.K. Vogt (1996). Sequence selective carbohydrate DNA interaction: dimeric and monomeric forms of the calicheamicin oligosaccharide interfere with transcription factor function. Proc. Natl. Acad. Sci. USA 93: 940–944.
- [64] Kirschning, A., A. Bechthold, and J. Rohr (1997). Chemical and biochemical aspects of deoxysugars and deoxysugar oligosaccharides. Topics Current Chemistry 188: 1–84.
- [65] Liu, H.-W., and J.S. Thorson (1994). Pathways and mechanisms in the biogenesis of novel desoxy-sugars by bacteria. Annu. Rev. Microbiol. 48: 223–256.
- [66] Madduri, K., J. Kennedy, G. Rivola, A. Inventi-Solari, S. Filippini, G. Zanuso, A.L. Colombo, K. M. Gewain, J. L. Occi, D. J. MacNeil, and C.R. Hutchinson (1998). Production of the antitumor drug epirubicin (4'-epidoxorubicin) and its precursor by a genetically engineered strain of *Streptomyces peuceiius*. Nature Biotechnology 16: 69–74.
- [67] Gaisser, S., A. Trefzer, S. Stockert, A. Kirschning, A., and Bechthold (1997). Cloning of an avilamycin biosynthetic gene cluster from *Streptomyces viridochromogenes* Tü57. J. Bacteriol. 179: 6271–6278.
- [68] Lombo F, K. Siems, A.F. Brana, C. Mendez, K. Bindseil, J. A. Salas (1997). Cloning and insertional inactivation of *Streptomyces argillaceus* genes involved in the earliest steps of biosynthesis of the sugar moieties of the antitumor polyketide mithramycin. J. Bacteriol. 179: 3354–3357.
- [69] Lombó, F., G. Blanco, E. Fernández, C. Méndez, J. A. Salas (1996). Characterization of *Streptomyces argillaceus* genes encoding a polyketide synthase involved in the biosynthesis of the antitumor mithramycin. Gene 173: 89–91.
- [70] Seow, K-T., G. Meurer, M. Gerlitz, E. Wendt-Pienkowski, C.R. Hutchinson and J. Davies (1997). A study of iterative type II polyketide synthases, using bacterial genes cloned from soil DNA: a means to access and use genes from uncultured microorganisms. J. Bacteriol. 179: 7360–7368.
- [71] Stemmer, W.P.C (1994). Rapid evolution of a protein in vitro by DNA shuffling. Nature 370: 387–388.
- [72] Crameri, A., E.A. Whitehorn, E. Tate, W.P.C. Stemmer (1996). Improved green fluorescent protein by molecular evolution using DNA shuffling. Nature Biotechnology 14: 315–319.
- [73] Zhang, Ji-Hu, G. Dawes, W.P.C. Stemmer (1997). Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. Proc. Natl. Acad. Sci. USA 94: 4504–4509.

This Page Intentionally Left Blank

# 13 Design and Diversity Analysis of Compound Libraries for Lead Discovery

*Hans Matter and Matthias Rarey*

## 13.1 Introduction

Today, novel technologies with potential for shortening the drug discovery process are essential for the pharmaceutical industry to rapidly develop new drug entities, as patent protection for many of their products will soon disappear. Combinatorial chemistry [1–4] and high-throughput screening (HTS) are perceived as such innovative technologies, and both are finding wide application as integrated methods for the identification and optimization of new drug candidates.

For lead finding, the use of chemical compound collections and combinatorial libraries is indispensable to every pharmaceutical company for a timely and efficient discovery of novel bioactive entities [5]. Until recently, most attention was paid only to internal historic compound collections, maintained over several years of medicinal research [6]. However, those libraries typically contain a limited number of structural classes as the result of past research projects which focused on ‘biological activity islands’. Today there is a growing interest in techniques that expand structural and – even more important – biological diversity, this perhaps being accomplished by the direct acquisition of an external compound collection, or by the use of internal or external synthesis capacities [7].

The aim of this chapter is to present motivations, concepts and practical considerations for the design, assessment and comparison of combinatorial and other compound libraries [8]. In contrast to traditional chemical synthesis, the goal of combinatorial chemistry is to manufacture populations of diverse small-molecule entities [9] for biological screening by systematic, repetitive connection of building blocks with varying chemical topology. The possibility of creating huge libraries in different formats is unprecedented in medicinal chemistry [10]. The technology to produce and extract diverse, yet representative, subsets [11] has captured the imagination of drug-discovery teams. While combinatorial chemistry has started with the production of large libraries of mixtures followed by sophisticated deconvolution strategies, the automated parallel or combinatorial synthesis of designed, smaller libraries consisting of single, small-molecule chemical entities is routinely used by the majority of research groups today [12]. Thus, the sections of this chapter will refer to this library concept.

The evaluation of chemical databases and the rational design of combinatorial libraries and subsets is important to optimize resources for a successful discovery strategy. Although today’s synthesis and assay technology allows the processing of large libraries, their production and testing must be carried out in an efficient and cost-effective way



[13]. Here, the identification of ‘redundant’ compounds based on molecular diversity considerations is a key requirement, as any reduction in the number of compounds to be synthesized and tested, by reducing redundancy within a database yet not introducing voids, should impact on research efficiency and costs [14]. Even with new miniaturization technologies, smaller compound subsets are essential if more assays are to be handled in a given time frame. Designing subsets using an adequate method enables project teams to find the appropriate balance between the number of compounds for HTS and the project’s lifetime.

## 13.2 Concepts and Issues for Combinatorial Library Design

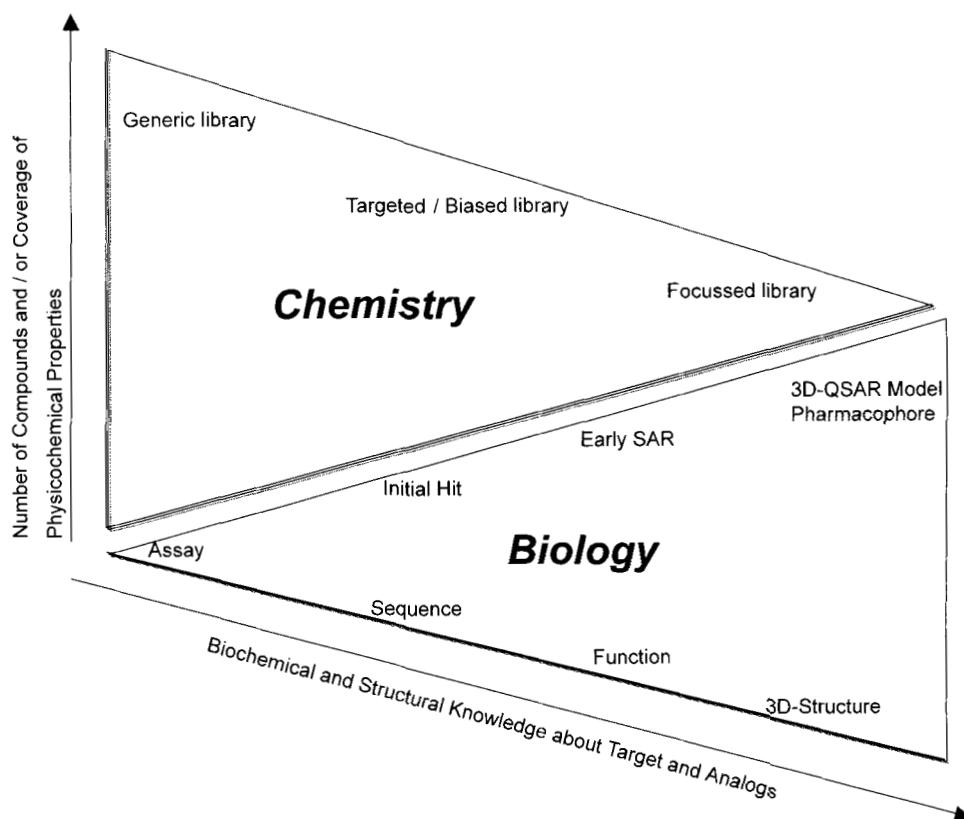
General design concepts in combinatorial chemistry were identified earlier by Martin et al. [11]. Efficiently designed libraries for general screening should minimize redundancy among its members by employing a basic set of compounds with very diverse structures. Conversely, highly focused screening for lead optimization should employ monomers with similar features to those in analogue leads.

Hence, the process of computer-assisted library design (CALD) as a tool for drug discovery is dependent on *a priori* knowledge about target or analogues (Fig. 13.1). If only the assay has been established for an otherwise uncharacterized biological target, a generic screening library should encompass a larger number of compounds with a broad coverage of structural motifs and physico-chemical properties. While one approach is to accept that this is purely a ‘numbers game’ and requires huge libraries with millions of randomly generated molecules, more attention has been drawn to rational design for enhancing the structural variations in those libraries, while maintaining moderate sizes [15].

Representative compound subsets should span the entire chemical property space of the parent library with a smaller number of (preferably) drug-like compounds. Applying such a design strategy requires additional screening iterations to synthesize and test similar compounds (analogue libraries) to identified hits, in order to derive structure–activity relationships (SARs). The ultimate goal of any rational design is to ensure that all new compounds add new information to the existing HTS compound collection. For such databases, an optimally *diverse* subset [14, 15] is designed, which includes as few compounds as possible, but still is representative of the entire collection. However, the design procedures must ensure that each new compound is of biological and chemical relevance [16].

If more biochemical and structural information on the target is accumulated, any rational design should bias the final combinatorial library towards structural features from the target, initial hits or analogues. As at this level, no detailed SAR can be derived, this bias being introduced by global similarity considerations during the selection from the underlying virtual library. The terms ‘biased’ or ‘targeted’ libraries are typically used for those ensembles.

If additional information is available, such as 3D-QSAR models [17], pharmacophore hypotheses [18], or 3D target protein structures [19], the library design should be focused towards those motifs or to regions shown to be essential in explaining the fundamental



**Figure 13.1.** Library design depending on knowledge about biological targets and analogues.

SAR. The global similarity is then replaced by local similarity considerations – some regions or functional groups are strongly correlated with biological activity, while other regions are not. This knowledge typically results from QSAR studies or analysis of protein–ligand 3D structures, while the small, highly focused libraries are used for lead optimization. Interesting strategies towards the design of biased or focused libraries have recently been reported [20–26].

All those design alternatives require a ‘virtual library’ as a fundamental basis, which refers to an electronic description of all possible compounds, which might be synthesized by a short sequence of specific transformations with particular reagents and one or multiple scaffolds. Appropriate fast and relevant searching techniques are available today, which can be used for selection and design of actual synthesis subsets in accordance with the level of biological knowledge [27, 28] from such a parent library.

Subsequently, for every validated primary screening hit from rationally designed libraries, the nearest neighbors must be identified using similarity searching [29, 30] based on the same descriptors as were used for design. If not already available, this subset of nearest neighbors must be synthesized and tested in a second iteration to establish the valid-

ity of every hit, and to develop an initial SAR. If a nearest neighbor was identified in a virtual library, there is a high probability that the selected compound actually can be made using the same, validated reaction protocol.

### 13.3 The Similarity Principle

The approaches today employed in library design and diversity analysis [31–33] have their origin in computational methods like substructure and similarity searching, clustering and QSAR. Their underlying concept of molecular diversity [34] is based on the ‘Similarity Principle’ formulated by Johnson and Maggiora in 1990 [35]: structurally similar molecules are assumed to have similar physico-chemical and biological properties. This principle leads to the design and evaluation of compound libraries spanning a wide range of chemical and biological properties and to the prediction of target properties for new molecules using known values for similar compounds.

Hence, the use of very similar molecules for primary screening does not enhance the probability to find different types of biological activities, while using dissimilar molecules should enhance the probability for finding interesting leads on different targets. Such a diverse chemical library should cover a wide range of diverse, non redundant compounds, whatever definition of diversity is appropriate for a particular case. Not only should the risk of missing a lead compound be low, but so also should the total number of compounds.

There is, however, no objective definition of the term ‘molecular diversity’ [12], and it has been shown previously that closely related molecules can exhibit significant differences in biological activity [36]. Thus, a fundamental requirement for any valid, useful molecular diversity descriptor is to minimize the number of exceptions from the similarity principle by considering only those physico-chemical properties which really reflect biological diversity. Several validation strategies and results will be presented below. The safest strategy certainly is to test all possible compounds, but a compromise between library size and representivity is required for reasons of efficiency. On the other hand, medicinal chemistry has operated for many decades according to this principle by introducing minor changes into a lead compound in order to slightly alter biological activity. A deeper understanding of the SAR then helps to focus on alterations, with a positive effect on activity.

While a typical historical database is characterized by well-separated clusters of similar compounds, an optimal distribution of molecules will avoid this degree of redundancy and the respective loss of information. Such a database can be designed taking structural properties into account to determine the smallest acceptable distance between two molecules (similarity radius). Taking a large variety of compounds within an initial virtual library, an optimal procedure would select only dissimilar compounds outside this similarity radius, leading to a more diverse subset, which increases the probability of finding lead structures. Moreover, such a distribution may help to identify structurally different molecules being active on the same target, supposing that the interest lies in the differentiation of active versus inactive compounds, regardless of any quantification of activity. Subtle activity differences on the other hand are subject of a lead optimization program based

on targeted libraries. Here, libraries should be designed which are enriched by compounds within the similarity radius around a given hit.

However, the premise to apply this library design strategy successfully is to uncover the relationship between structural similarity encoded in several molecular descriptors and biological similarity, *i.e.*, activities in the same assay. Important details of this concept require further investigation: Which similarity descriptor is useful for diverse selections? Which chemical diversity metric reflects biological diversity? How large is the similarity radius for useful descriptors based on activity data in a single biological assay? Some studies to address those issues will be discussed in the following sections.

## 13.4 Molecular Descriptors and Selection Techniques

### 13.4.1 Two-Dimensional Fingerprints

Interestingly, many two-dimensional (2D) descriptors based on fragment counts from 2D substructure searching systems and associated physico-chemical properties find novel applications to characterize, design and analyze compound libraries. It is well accepted by organic chemists that a 2D structural diagram encodes physico-chemical properties and reactivity, though it also was suggested that some 2D descriptors might contain 3D information, as they can predict biological activity [37]. Recently, it was uncovered that 2D substructure descriptors also contain information about relevant terms for protein–ligand interactions, like molecular shape, hydrophobicity, size, flexibility and hydrogen-bonding potential [38]. Those binary descriptors such as 2D fingerprints [39–41] encoding the absence or presence of specific chemical fragments were originally combined with non hierarchical clustering algorithms for similarity-based selections [42]. A critical assessment of 2D fingerprint-based similarity considerations is given in [43].

The following description of UNITY 2D fingerprints [39] should give a representative example for such a descriptor: A list of all fragments of a particular size is generated for each molecule and converted into a bitstring. Due to the large number of possible fragments in a database, it is necessary to assign multiple fragments to each bit of the bitstring. This is achieved by a two-step procedure: First, each fragment's line notation [44] is mapped to a unique integer using a cyclic redundancy check algorithm [45]. Each integer is then projected to this size-limited bitstring by a procedure known as 'hashing', setting one or multiple bits to 1. Specific bits are reserved for particular fragment sizes. For each fragment, multiple occurrences lead to more neighboring bits set to 1.

Alternative 2D fingerprint descriptors are Daylight 2D fingerprints [40, 46, 47] and MACCS substructure keys [41]. While Daylight 2D fingerprints also encode paths of a given fragment length as features, they do not have specific bit ranges reserved for particular fragment lengths after hashing. Both hashed fingerprint methods are designed to describe all structures with a similar quality independent from a predefined fragment dictionary.

MACCS substructure keys on the other hand encode the presence of a predefined set of relevant 2D fragments, originally designed for speeding up database substructure searching [48, 49] by eliminating those compounds from detailed consideration that can-

not possibly match the query. Here, every bit is mapped unambiguously to a single 2D fragment.

Encoding molecular structures in the form of bitstrings allows quantification of the molecular similarity using different functions, like the Tanimoto [37, 50] or Cosine [51] coefficients. Both are counting the number of bits in common set to 1 in slightly differing ways. The Tanimoto coefficient is widely used in database analysis, as it has certain properties making the work with larger datasets very efficient. A similarity coefficient of 0 means that both structures have no '1' bits in common; there is no intersection between both sets of fragments. In contrast, a value of 1 indicates that the fingerprints are identical.

### 13.4.2 Pharmacophore Definition Triplets

The major deficiency of those fingerprints is that they are '2D' in nature, carrying only information about atoms, bonds, and fragments. A molecular 3D structure cannot be derived, although the very existence of successful 3D-QSAR and protein–ligand docking studies suggests that biological activity and specificity are highly correlated with the ligand's 3D structure.

Reports by Mason et al. [52–54], Martin et al. [55, 58] and Davies [56] have highlighted interesting properties of novel 3D descriptors for selecting representative subsets, leading to commercial software products [57, 58]. The term 'pharmacophore definition triplets' refers to a set of three pharmacophoric points within a molecule, where multiple points can have a similar type (*e.g.*, acceptor–acceptor–hydrophobic). Each possible triplet, together with its point-to-point distance disregarding the order of specification, can be encoded in a fingerprint-like manner. Individual bits refer to different triangles, which can be formed between potential pharmacophoric points (*e.g.*, acceptor-atoms, acceptor-sites, donor-atoms, donor-sites and hydrophobic centers [18a, 58]). Different approaches have in common that individual bits set to 1 encode a particular triangle geometry in pharmacophoric space. Triplet fingerprints are computed for conformational ensembles to account for flexibility. The pharmacophoric definitions are reflecting biologically relevant conditions and accommodate tautomeric potential. Typically, 27 distance bins (from 2.5 to 15 Å) and five point types are used, resulting in 307 020 bits encoding triangle geometries. Typically a single composite fingerprint is accumulated as union of all molecules in the database. Successful applications of enhanced four-point pharmacophoric descriptors have also recently been reported [59].

### 13.4.3 Other Descriptors

The following descriptors have also been applied for diversity analysis (for recent reviews, see [60, 61]).

- **Atom pair fingerprints** count the shortest path of bonds between two atom types [62, 63], each bit corresponds to a specific pair separated by a predefined number of bonds. Recently, physico-chemical properties rather than atom typing were successfully implemented [64].

- **Topological indices** encode the molecular connectivity pattern [65, 66] and quantify features like atom types, numbers of atoms and bonds, and the extent and position of branchings and rings. Furthermore, electrotopological state values [67], molecular shape indices [68] and topological symmetry indices are used. Often, the large number of correlated indices is condensed by principal component analysis (PCA) [69].
- **2D and 3D autocorrelation vectors** [70] represent intramolecular 2D topologies or 3D distances within particular molecules. An autocorrelation coefficient is a sum over all atom pair properties separated by a predefined number of bonds (2D) or distance (3D), while the entire vector represents a series of coefficients for all topological or Cartesian distances. Atomic properties involve hydrophobicity [71], partial atomic charges, hydrogen bonding potential and others. Again, a PCA is often used to reduce the number of variables. 3D autocorrelation vectors of properties based on distances calculated from 3D molecular surfaces [72] have also been applied to visually assess the diversity of different libraries [73].
- **Flexible 3D fingerprints** [74] encode interatomic distances of flexible molecules, in particular the ability to adopt a conformation with a specific atomic distance. Distances between predefined features are stored in a bitstring.
- **Steric fields** derived from the CoMFA method [75] were used as 3D shape descriptors, while their major drawback is that a reliable superposition rule for databases is required. Thus, topomeric fields as significant enhancement for comparing the diversity of a set of variable groups attached to a common scaffold have been developed and validated [76]. Those fields rely on canonical conformations of aligned compounds from combinatorial libraries using the reaction core. They use a novel way to encode flexibility depending on the shortest bond distance between any atom to the reaction core.
- **Weighted Holistic Invariant Molecular (WHIM) indices** [77] represent a different 3D approach to overcome the molecular 3D alignment problem, as they are invariant to molecular rotations and translations. These indices encapsulate information about the molecular 3D structure in terms of size, shape, symmetry and atom distribution, solely derived from Cartesian coordinates.
- **Physico-chemical properties** are also used to quantify the diversity of molecules [11, 78], although today's algorithms still fail on a remarkable number of candidate molecules, which are then rejected from subsequent considerations.

Several other interesting 2D [11, 79] and 3D [80] descriptors have been successfully applied in the recent literature.

#### 13.4.4 Compound Selection Techniques

Different techniques for selecting compounds from larger virtual or existing databases have been proposed in the literature, ranging from computationally very expensive clustering [81, 82] to rapid maximum dissimilarity [83–86] or grid-based partitioning methods [87, 88]. A comparison of grid-based partitioning applications to similarity selections is given in [27]. In maximum dissimilarity methods, new compounds are successively selected such that they are maximally dissimilar from those previously selected. This process

can be terminated either when a preset maximum number of compounds is selected, or when no other molecules can be selected without being too similar to already selected molecules. There are several variations to this technique, which have recently been compared in a detailed study [89]. To result in a deterministic procedure, the first three compounds are rejected after the fourth selection, but are allowed for later picking [58]. A more flexible implementation of this method, the OptiSim algorithm, has recently been described [90].

Cluster analysis [81, 82] as an alternative method offers more specific control by assigning every structure to a group which exhibits a high degree of both intracluster similarity and intercluster dissimilarity. A variety of different clustering techniques have been critically evaluated [29, 50a, 91]. There are no *a priori* guidelines for which technique would be most appropriate for a particular dataset, though some methods perform better than others for grouping similar compounds [46, 47].

The idea of hierarchical clustering is to start with singleton clusters and sequentially to merge those two clusters with minimal intercluster distance. This approach does not require any assumption to be made about the number of clusters to generate, but computationally is the most expensive technique. Different methods can be used to compute distances between clusters: the distance between the closest pair of data points in both clusters (single linkage clustering), the distance between the most distant pair of data points in both clusters (complete linkage clustering), the average of all pairwise data points between two clusters, and the distance between two cluster centroids. For compound selection, typically one or more structures closest to the cluster centers are chosen to form a representative set.

Furthermore, genetic algorithms [92–95], neural networks [73] or simulated annealing [96] have also been applied for selecting appropriate sublibraries.

### 13.5 Selected Examples of New Approaches to Molecular Similarity

Despite the wide range of existing methods for describing molecular similarity, new techniques are under continuous development. Their major goal is an improved description of molecular similarity, while maintaining computation speed.

As a quantitative performance measure of a similarity descriptor, the number of molecules from a database belonging to the same class as a reference molecule can be used. The enrichment factor is such a measure, and this takes the ratio of hits to the entire database size into account. It is defined as number of hits in the first percent of the database sorted according to the similarity measure, divided by the expected number of hits retrieved with a random selection. Intuitively, the enrichment factor describes how much is gained by using a similarity measure for selecting molecules compared to a random selection. For the design of diverse libraries, the detection of molecular classes is a necessary feature of the similarity measure.

Besides the quantitative measure of the number of retrieved hits, their quality in the sense of usefulness for lead discovery is also of interest. Molecules bearing a common scaffold are typically less useful, as they do not reveal a chemically new motif. Therefore,

approaches describing similarity not on the 2D structure level may be of great value, even if they show a lower performance in a quantitative evaluation.

The main difficulty for commonly used 3D descriptors results from the treatment of conformational flexibility. Considering that only a single conformation is inadequate for most classes of pharmaceutically relevant molecules, averaging over conformational space provides only a very rough view as to what is intended to be described, namely the possible arrangement of pharmacophoric groups in Cartesian space. In this section a more detailed view on three subjectively selected approaches to describe molecular similarity will be presented. The approaches have in common the fact that similarity is quantified to a less dependable degree from only the 2D structure or specific 3D conformations.

### 13.5.1 Affinity Fingerprints

In 1995, researchers at Terrapin proposed a completely different way to describe molecular similarity [97]. Rather than look at potential drug molecules alone, they took a set of reference proteins (the reference panel) into account. For each molecule, the binding affinity to each protein in the panel was determined experimentally. The list of binding affinities with respect to the reference panel is named the 'affinity fingerprint' and is used as descriptor. The underlying idea is that while some similarity exists between a target protein and proteins in the reference panel, then binding affinities should correlate due to similar sub-pockets or pharmacophores which are common between the proteins.

This concept of correlated affinity data was validated [98] by demonstrating that binding data for human serum albumin exhibit a good correlation with three molecules from their reference panel. Indeed, it becomes increasingly difficult to add new proteins to the reference panel without introducing redundancy.

Affinity fingerprints can be used like other conventional descriptors in similarity searching and diversity calculations, in that they need additional experimental effort as well as physically existing compounds for profiling. Dixon and Villar [98] used affinity fingerprints with a Euclidean distance measure and a simple optimization algorithm to design diverse libraries from existing compound collections. They also showed that a wide range of structurally diverse active molecules could be detected using the affinity data of a first screening experiment.

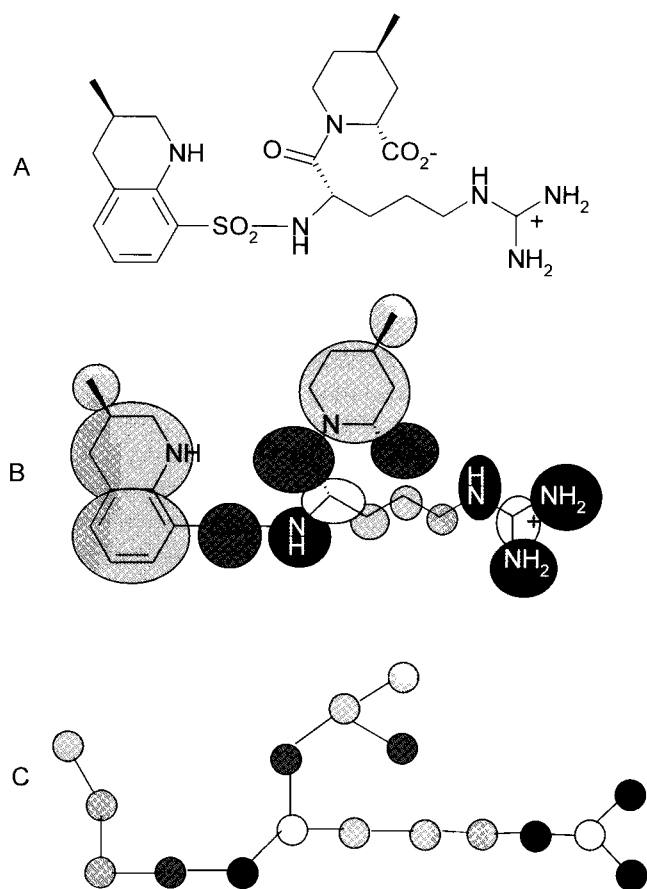
In order to avoid the large experimental effort of the Terrapin approach, Briem and Kuntz [99] used the program DOCK [100] for generating *in-silicio* affinity fingerprints by estimating binding affinities for a series of molecules. For this approach, the 3D structure of the proteins in the reference panel must be known. Their work showed that – although performing more poorly than 2D techniques – an enrichment could be achieved, and this led to hits being structurally diverse on the 2D level. Currently, this approach is enhanced by using a docking program which explicitly considers conformational flexibility [101] and by extending the reference panel of proteins. The selection of proteins for the panel is also addressed. Here, optimization schemes are developed in order to increase predictivity of the *in-silicio* affinity fingerprints [102].



## 13.5.2 Feature Trees

Most descriptors referred to above represent molecular features in a linear fashion, a bitstring or a vector, regardless of whether the features themselves are macromolecular, two-, or three-dimensional. The advantage of such a linear representation is that for a comparison of two molecules only a simple mapping between their vectors is required. In most cases, the mapping function is the identity function such that bit  $i$  of descriptor  $a$  is compared with bit  $i$  of descriptor  $b$ . While comparisons can be performed rapidly with simple algorithms, the disadvantage is that information about the arrangement of features in a molecule is only partially represented, or even lost.

In order to achieve a better approximation of a molecular structure, a new class of descriptors was recently described [103] which bridged classical descriptors and methods for



**Figure 13.2.** The feature tree descriptor: A molecule (A) is divided into its major building blocks highlighted by gray spheres (B). Each building block is represented by a node in the feature tree (C). The shade of gray indicates major interactions, which a building block is able to form with a potential receptor: hydrogen bond donor (black), hydrogen bond acceptor (gray), hydrophobic (light gray).

structural alignment in 3D-QSAR. Here, a molecule is described by a tree structure, called a 'feature tree'. The nodes of this feature tree represent individual molecular fragments, while edges are connecting nodes in correspondence with fragments joined in the 2D chemical structure. Thus, a feature tree is a rough approximation to the structural formula of a particular molecule (Fig. 13.2).

Each node of this tree contains a set of features derived from the underlying substructure fragment. Those features describe substructure properties at the particular feature tree nodes, and typically are subdivided into steric and chemical features. However, there is no limitation on the physico-chemical nature of features, as long as the following functions can be computed:

- a generator function calculating the feature value from a molecular fragment;
- a comparison function calculating a similarity value from two feature values; and
- a combine function calculating a feature value for a set of nodes.

The steric feature is an approximated van der Waals volume to represent the size of a molecular fragment. The chemical feature is an interaction profile, represented as array where the  $i$ -th entry gives the number of possibilities to form an interaction of type  $i$ .

If two feature trees – each representing one molecule – should be compared, a mapping between parts of the trees, so-called subtrees, must first be computed. Two recursive algorithms were developed to address this matching problem: the 'split-search' and the 'match-search' algorithms. The basic idea of the split-search algorithm is that on each recursion level, both feature trees ( $a;b$ ) are cut into two subtrees ( $a1,a2;b1,b2$ ), called a split, such that ( $a1;b1$ ) and ( $a2;b2$ ) form the input of the next recursion level. The resulting matching considers the topology of the trees, but the matches may be disconnected, *i.e.*, nodes in the middle of tree  $a$  or  $b$  are unmatched. The 'match search' algorithm starts with an initial split and then adds matches to both sides of the split subsequently. In contrast to the first algorithm, the match search algorithm produce a connected set of matches.

This feature tree descriptor introduces several new concepts into the descriptor technology. As mentioned above, due to the tree structure of the descriptor, the overall molecular structure is represented. The tree structure is a trade-off between exactness (a better representation would be a graph structure allowing to handle complex macrocycles) and computational speed. Tree structures can be mapped quite efficiently allowing average comparison times in the order of 50 milliseconds. In contrast to methods based on 2D molecular fragments, a comparison of two feature tree nodes detects similarity if the same kind of interactions can be performed from both represented molecular fragments, even if the 2D structures are dissimilar. In addition, the developed matching algorithms are very flexible.

This feature tree approach shows similar – and in some cases even improved – performance compared to a 2D fingerprint approach [103]. Based on 3D protein–ligand complex structures, it was also shown that the mapping as basis for the similarity computation represents the relative binding orientation in the active site of a protein. In 61 % of all test cases, the average distance between mapped fragments is less than 4 Å.

### 13.5.3 Automated Structural Superposition of Fragments

Another interesting approach is based on rigid-body structural superpositions of molecular fragments to reference molecules [104], which can be extended to a descriptor technology in a straightforward manner. In order to search 3D databases for compounds that are similar to a particular query, a set of fragments is generated from that starting molecule. Subsequently, each fragment is superposed onto each database entry using a rapid rigid-body superposition algorithm (RigFit) [105]. The resulting similarity score from each superposition allows the retrieval of similar database entries. A molecular descriptor that is independent of a query molecule can be generated by comparing the entire database to a reference set of predefined fragments. As fragment-based descriptors have been proven to be very effective on the 2D level (see Section 13.6), this superposition approach is a natural extension to 3D, being less dependent than solely 2D structure similarity.

When comparing this approach to a 2D fingerprint descriptor on a dataset containing about 1000 molecules from different biological classes, an increased enrichment for three out of six classes was observed. As this approach requires approximately 1 second of CPU time per fragment-to-molecule comparison, it is applicable to database searching and profiling. If the method is used to generate query-independent descriptors, this superposition step must be performed only once for each fragment–molecule pair in a database generation phase.

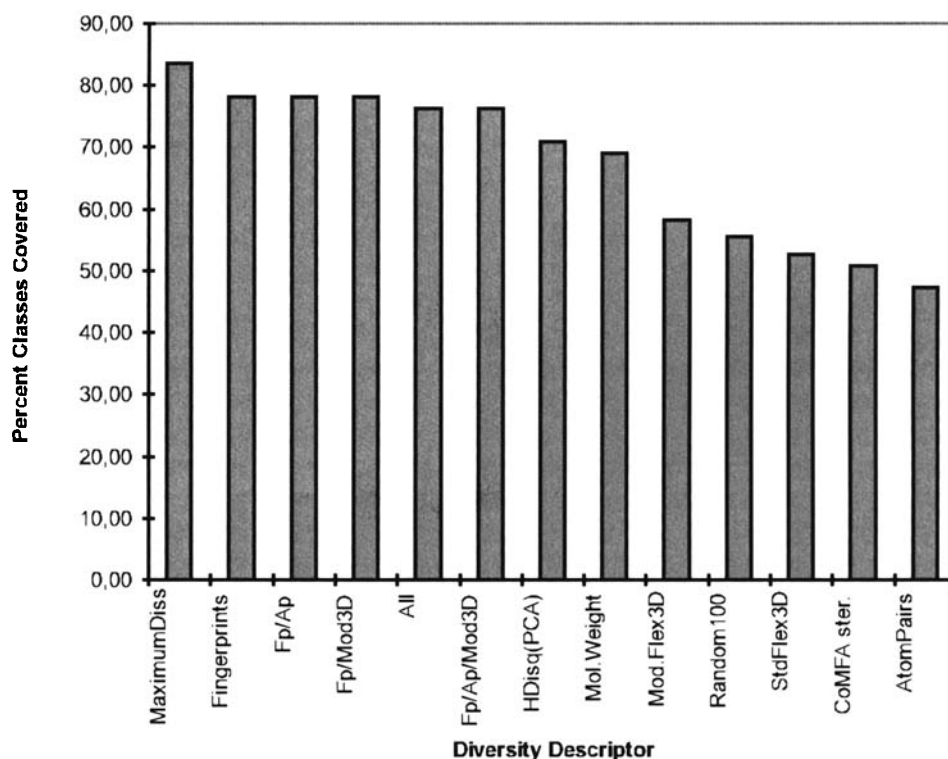
## 13.6 Descriptor Validation Studies

### 13.6.1 2D versus 3D Descriptors for Global Diversity

Given these numerous descriptors and selection techniques, the question arises, whether chemical diversity can be correlated to biological diversity. This section summarizes own and many published validation studies for selected diversity descriptors. Several physico-chemical descriptors were investigated in different studies to uncover the relationship between 2D/3D similarity and biological activity.

Compound selections to design diverse subsets were reported [15, 106, 107] on a part of the IC93 database [108] with 1283 active compounds and 55 biological classes derived according to the biological activity strings extracted from that database [106]. It was assessed whether subsets represent all biological properties of the parent database. As a second example, 334 compounds from 11 QSAR series were analyzed (Bayer) [106], with the advantage that biological data were collected under identical conditions.

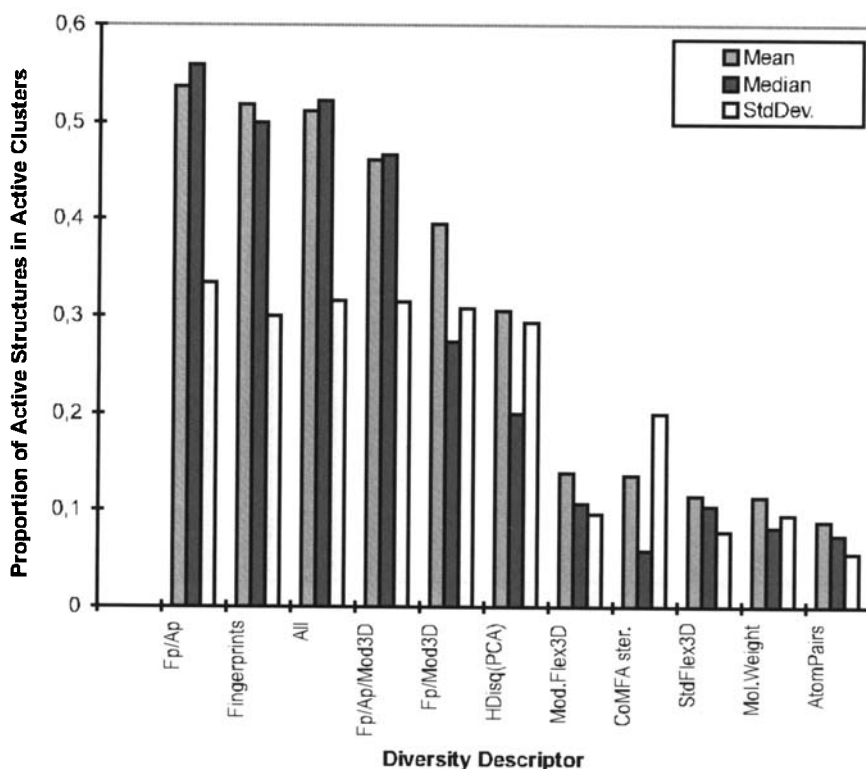
First, the coverage rate of 120 selections after hierarchical clustering or dissimilarity-based selections for all classes was monitored (Fig. 13.3). 2D fingerprints show the best performance using hierarchical cluster analysis (78.2 % classes represented) or maximum dissimilarity methods (83.2 %), while selection on clusters generated using atom-pair descriptors or standard flexible 3D descriptors perform unsatisfactory (cf. Fig. 13.3). On the other hand, combined descriptors containing 2D fingerprints and one or more of the above-mentioned metrics generally show similar performances compared to standard 2D-fingerprints.



**Figure 13.3.** Sampling of 55 biological classes from the IC93 database. Each descriptor was used to select 120 structures using hierarchical clustering except for the first one. The percentage of covered biological classes is indicated on the y-axis. The following descriptors were used: 1. UNITY 2D Fingerprints and maximum dissimilarity; 2. 2D Fingerprints and hierarchical clustering; 3. Fp/Ap: 2D and atom pair fingerprints; 4. Fp/Mod3D: 2D and pharmacophoric flexible 3D fingerprints; 5. All: 2D fingerprints, atom pair fingerprints, topological indices and pharmacophoric flexible 3D fingerprints; 6. Fp/Ap/Mod3D: 2D, atom pair and pharmacophoric flexible 3D fingerprints; 7. HDisq(PCA): topological indices condensed into eight principal properties (HDisq); 8. Molecular weight; 9. Mod.Flex3D: pharmacophoric flexible 3D fingerprints; 10. Average over 100 random selections; 11. StdFlex3D: standard flexible 3D fingerprints; 12. CoMFA ster: molecular steric fields; 13. Atom pair fingerprints.

Second, the ability of hierarchical cluster analysis to separate bioactive compounds from inactive structures for a single target was assessed [107] according to Brown and Martin [47] in their study to evaluate several 2D and 3D descriptors along with clustering techniques. An active cluster contains at least one active compound, and the active cluster subset refers to the total number of structures in all active clusters. Then, the proportion of active structures in a subset is computed and compared to the proportion of active structure in all clusters. This proportion  $p$  ranges from 0 to 1. The averaged proportion over all biological classes is shown in Figure 13.4. Any increase indicate the separation of active from inactive structures for a particular descriptor.

Again, 2D fingerprints alone or in combination with secondary descriptors perform better than other metrics, and  $p$  values ranged from 0.54 to 0.51. Replacing standard by

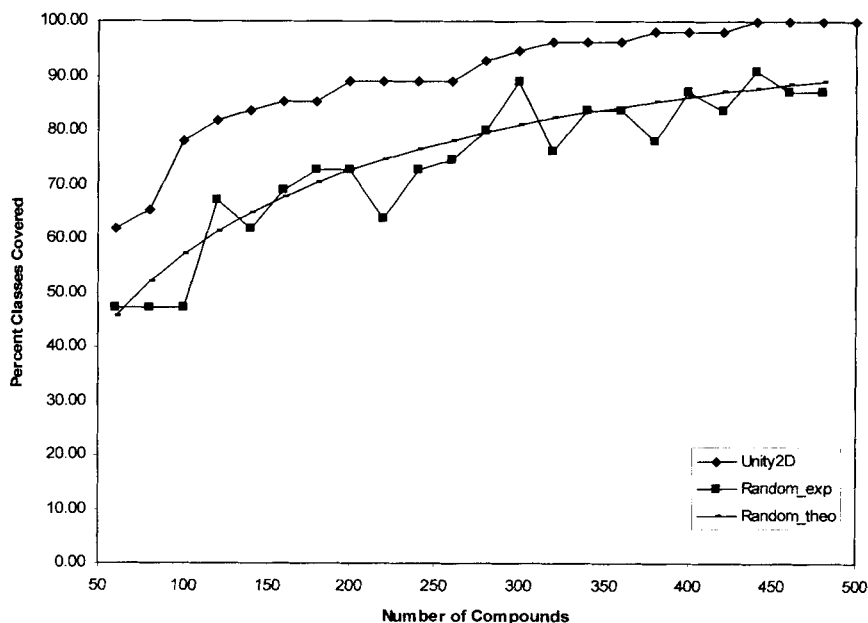


**Figure 13.4.** Correspondence between biological activity and cluster formation. Using hierarchical clustering and various descriptors, the averaged proportions of active structures in active clusters are computed.

modified (pharmacophoric) 3D fingerprints slightly increases the cluster separation tendency, but all 3D descriptors do not sufficiently separate active from inactive structures.

### 13.6.2 Random versus Rational Design for Global Diversity

In additional studies [106], the efficiency of rational design to maximize the structural diversity of databases was compared to random selections using 2D fingerprints and 3D PDT fingerprints [52–54, 58, 109]. Maximum dissimilarity methods and a direct dissimilarity-based selection were used to select subsets with sizes ranging from 60 to 500 members from the IC93 database. The sampling of biological classes in each subset was then evaluated (Fig. 13.5). Probability calculations and 100 random selections were used to compare rational design to a random approach. It is possible to compute the probability  $p$  to find  $nI$  hits by  $n$  selections in a database with a total of  $N$  compounds and  $NI$  hits for a particular target. Thus, one can evaluate how many target classes are covered by a pure random selection of  $n$  compounds.

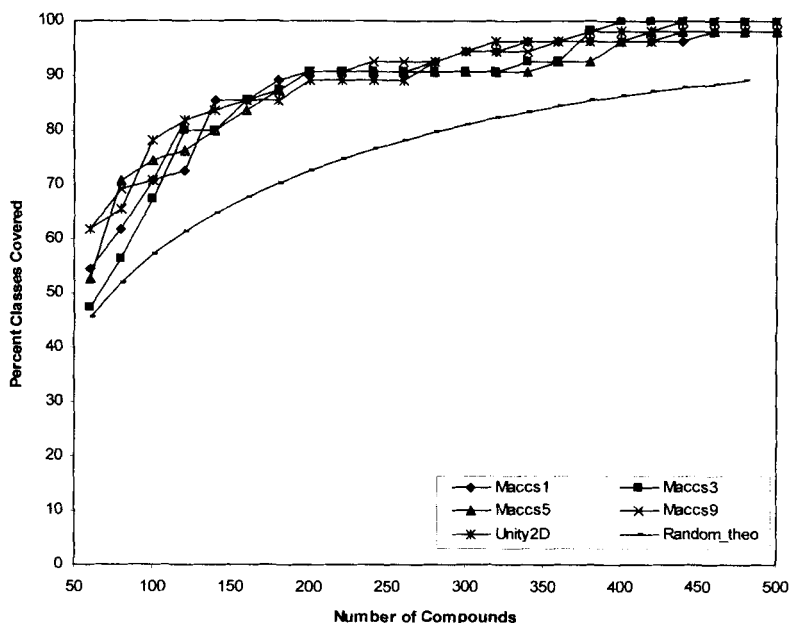


**Figure 13.5.** Selection of subsets from the IC93 database using random picking (theoretical expectation and representative experimental result) and maximum dissimilarity selection based on UNITY 2D fingerprints. The percentage of biological classes sampled from the IC93 database is plotted versus the subset size.

With increased sizes for random or rational subsets, the mean and maximum Tanimoto coefficient is also increased, while more diverse subsets are obtained using rational methods. Interestingly, those subsets also sample more biological classes than the corresponding randomly selected subsets (Fig. 13.5). A subset of more than 440 rationally obtained members contain at least one member per biological class, while many classes are not represented in the corresponding random subsets. The mean Tanimoto coefficient of about 0.85 for the smallest subset sampling all biological classes is in agreement with the similarity radius for 2D fingerprints derived using other methods [46, 47, 107, 110].

In order to represent 90% of all biological classes, 3.47 times more compounds must be randomly selected compared to the 2D fingerprint-based maximum dissimilarity method. This biological sampling rate is the most efficient selection when comparing random to rational design. Such coverage rate shows the highest efficiency of design techniques compared to a random approach. Filling the gap between 90% and 100% biological representivity requires many more compounds. For example, more than 440 compounds of the IC93 database sample 100% biological targets instead of 280 compounds for 90%. Remarkably, lower numbers of selected molecules led to almost similar coverage of biological classes for the rational and random approach. All results are similar for the Bayer database.

As it was previously shown that MACCS substructure keys outperform UNITY and Daylight 2D fingerprints [46], the IC93 database was investigated using an implementa-

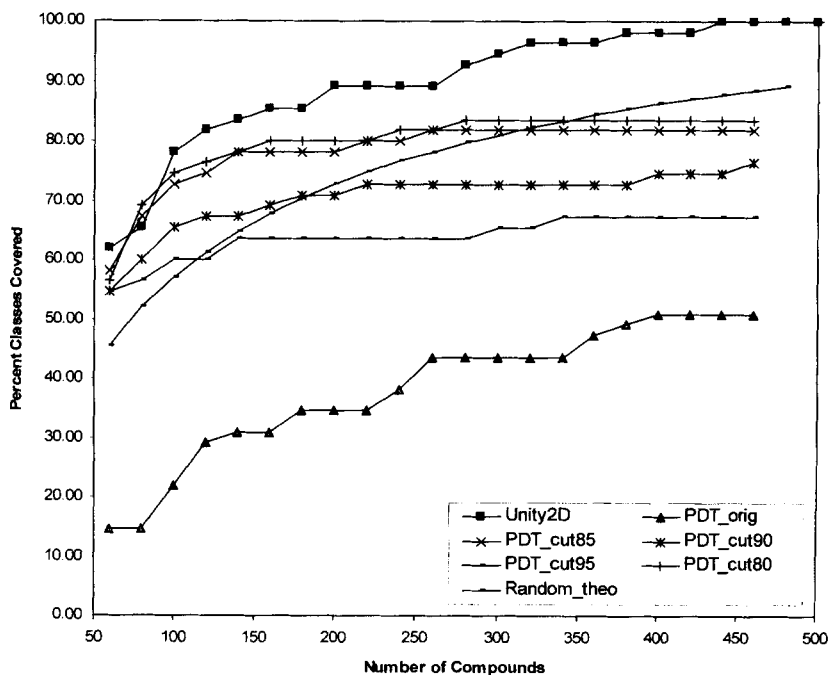


**Figure 13.6.** Percent biological classes covered from the IC93 database versus subset sizes for maximum dissimilarity selections using selected MACCS substructure keys counting up to 1, 3, 5 or 9 occurrences of a particular fragment key, UNITY 2D fingerprints (*Unity2D*), and theoretical random selections (*Random\_theo*).

tion of MACCS substructure keys in UNITY 4.0. To count multiple occurrences, up to 12 neighboring bits were reserved in different fingerprint files. Thus, each time a fragment is identified, one bit is set, until the maximum number of bits reserved for this fragment is reached. Compound subsets were generated using maximum dissimilarity methods based on the Tanimoto coefficient. All fingerprint files counting from 1 to 12 occurrences perform significantly better than a random approach. In Figure 13.6, the differences of MACCS key results to the UNITY 2D fingerprint coverage rates of biological classes are plotted versus subset sizes. From this figure it is clear that MACCS substructure keys outperform UNITY 2D fingerprints. This is especially true for the *Maccs\_3* (3 bits per fragment) fingerprint definition for larger populations and *Maccs\_5*/*Maccs\_9* for smaller-sized subsets.

### 13.6.3 3D Pharmacophore Definition Triplets versus 2D Fingerprints

Using the MaxDiss algorithm, 2D fingerprints were shown to perform more efficiently than a random selection. In contrast, the performance of PDTs is strongly dependent on the initial ordering of compounds in the parent database [109]. To select compounds based on a composite database fingerprint for PDTs, all structures were evaluated in a se-



**Figure 13.7.** Percent biological classes covered from the IC93 database versus subset sizes for selection using 2D fingerprints and maximum dissimilarity selection (*Unity2D*), theoretical random selections (*Random\_theo*) and different implementations of the PDT selection method (*PDT\_orig*, *PDT\_cut80*, etc.).

quential order. A new structure is selected, if its individual PDT fingerprint is more diverse than a given Tanimoto cut-off to the previous fingerprint of the selection. Even after database presorting according to 2D fingerprint diversity using the MaxDiss algorithm and subsequent selecting using PDTs, such a selection strategy based on PDTs is not as efficient as a selection with 2D fingerprints alone (Fig. 13.7).

Only 15% of all IC93 classes were covered using a subset with 60 compounds, while a subset with 460 members only represents 49% of all biological classes. In contrast, the 2D fingerprint subset represents all classes, while even the random selection covers 87.8%. It was unexpected to find that 3D PDT selections always show lower coverage rates than a random selection.

Setting up a combined descriptor strategy by ordering the dataset first according to the most diverse compounds using 2D fingerprints with maximum dissimilarity sorting, and using this ordered set for PDT-based selections, significantly improves the coverage rates. This is in agreement with the ChemDiverse presorting strategy [57]. Coverage results for small populations are very similar to 2D fingerprint-based selection, as they are dominated by the effect of the 2D fingerprint-based presorting and not the 3D pharmacophoric key diversity. No improvement of the initial 2D fingerprint-based selections by such an advanced molecular description could be observed. On the other hand, a selection of more than 440 structures represents only 78% of all classes (100% for 2D fingerprints



only). Again, it is remarkable that a random selection for the IC93 database performs better for larger subsets than this combination of 2D fingerprint presorting and PDT selection (Fig. 13.7). Another slight improvement could be achieved, when using a cut-off Tanimoto coefficient of 0.85 to generate the cumulative database PDT key. The obtained subset performs significantly better than random for small populations, while an increased subset size led to a better performance of the random selection, showing the inappropriateness of PDT fingerprints using this particular selection technique alone or in combination with 2D fingerprints for selecting representative larger subsets.

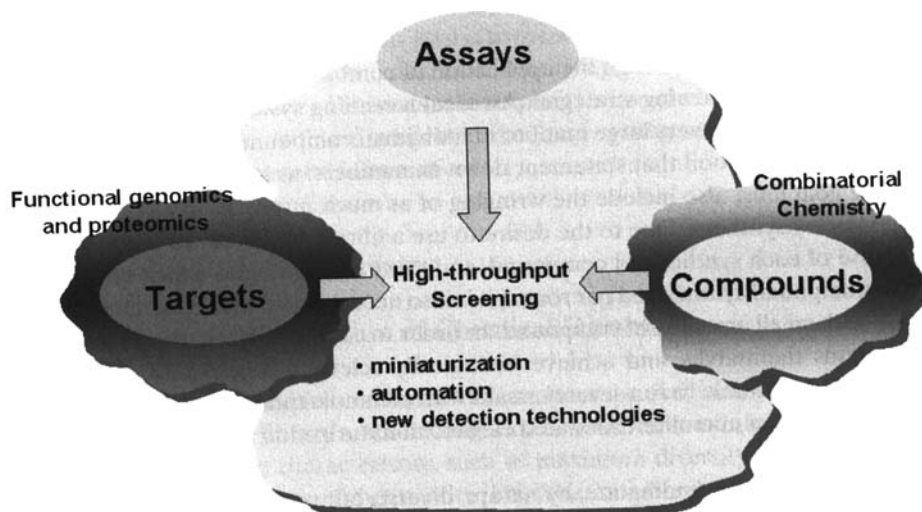
### 13.6.4 Local Similarity – The Radius of Similarity

The match between chemical and biological similarity was also studied using structurally similar molecules acting on different biological targets. In a pioneering work, Patterson et al. evaluated the ability of 11 different descriptors to group compounds from 20 representative QSAR datasets according to expectations from the similarity principle [110]. The ‘neighborhood plot’ was introduced as a useful graphical representation (Fig. 13.8). A similarity radius for 2D fingerprints and molecular steric fields was estimated, and compounds within this radius of another molecule were shown to have comparable biological properties, as expected from this principle.

In order to be useful for library design, any molecular descriptor must exhibit a relationship to biological activity. Within a similarity radius of a bioactive compound there should be a high probability that any other compound belongs to the same activity class – otherwise the design is similar in result to random screening. Hence, a descriptor can be characterized using this similarity radius allowing the selection of minimally redundant compounds.

An additional investigation was undertaken to confirm the relationship between local structural similarity and biological activity for 10 biologically diverse QSAR datasets with 729 molecules [107] and a biological activity range larger than 2.5 orders of magnitude. For each of the  $n(n-1)/2$  pairs of  $n$  molecules per assay, the pairwise differences for the physico-chemical descriptors and the biological activity differences were plotted in a neighborhood plot.

Example neighborhood plots for 138 structurally diverse ACE inhibitors [111] are shown in Figure 13.8. Such a plot of pairwise descriptor differences versus biological distances should reveal a characteristic shape for any valid metrics, allowing the derivation of a maximum change for the biological activity per change in this descriptor. The indicated arrows for steric fields and 2D fingerprints highlight a linear gradient with some discontinuities and a slope of 0.04 and 0.10 for steric fields (Fig. 13.8(a)) and 2D fingerprints (Fig. 13.8(b)). In contrast, molecular weight (Fig. 13.8(c)) or random numbers do not allow the derivation of such a similarity relationship. From the averaged gradients, a similarity radius of 0.85 for 2D fingerprints and 0.88 for steric fields was estimated, taking a tolerance of three orders of magnitude for biologically active molecules into account. Thus, two molecules with a Tanimoto coefficient larger than 0.85 should exhibit related biological activities. If a molecule A falls within the similarity radius of molecule B, it is sufficiently represented by B without loss of biological information. This result is in good



**Figure 13.8.** Comparison of pairwise biological distances versus descriptor differences for 138 angiotensin-converting enzyme (ACE) inhibitors (upper panel) as example for neighborhood plots: (a) CoMFA steric fields; (b) 2D fingerprints; (c) molecular weight.

agreement with previously published data [55, 85, 106, 110]. It does not imply, however, any kind of linear correlation, and does not hold for larger biological differences.

Thus, it was shown that designed subsets are superior in the sense of sampling more biological targets compared to randomly picked ensembles [106]. 2D fingerprints as molecular descriptors were shown to be appropriate for designing subsets representing biological properties of parent databases. They perform better, when comparing the sampling properties of other descriptors carrying 2D or 3D molecular information. All studies reveal that 2D fingerprints alone or in combination with other metrics as primary descriptor allow handling of global diversity. Based on these findings, an optimally diverse subset encompassing only 38% of the IC93 database can be designed using 2D fingerprints, with no structure being more similar than 0.85 to any other compounds, but all biological classes being represented.

## 13.7 Designing Combinatorial Libraries

### 13.7.1 Strategy

This section summarizes strategy and practical considerations for designing and analyzing diverse chemical structures in combinatorial or parallel synthesis library format based on concepts and results introduced above.

The initial step in library design is an idea of a new synthetic route or an innovative scaffold. Such a new scaffold might be synthetically accessible through a novel, elegant route, or it is of medicinal chemistry relevance, such as protein secondary structure motifs, turn mimetics, benzodiazepines, pyrimidines and many other examples. It must be clearly evaluated, which building blocks are compatible with a particular synthetic scheme. This reaction validation phase is the key step for a successful library design and realization, as computer-assisted design is encompassing the rules derived from these systematic studies.

Any library design procedure needs to make sure that every new compound is unique and adds useful information to the entire library. As shown, there is a higher probability of finding a given biological activity in clusters of similar molecules where activity has been already detected than in dissimilar ones. This finding suggests that structurally similar compounds behave similarly in a biological environment for some physico-chemical descriptors. To design a generic library for lead finding, the consideration of two very similar compounds does not enhance the ability to find different types of biological activity; indeed, it would be better to replace one of these compounds with a more dissimilar one. This concept is the key motivating factor for the design of optimally diverse compound libraries [14, 15]. A major component is the definition of an exclusion region around each compound (similarity radius), where it is undesirable to have overlapping candidates. As suggested by validation studies, we typically choose a Tanimoto coefficient of 0.85 for 2D fingerprints to represent the boundary between similarity/dissimilarity in terms of biological activity for product structures.

The library itself should exhibit a wide coverage of the physico-chemical property space. New compounds should be true drug-like compounds with biological relevance. There have been some very interesting approaches to estimate the 'drug-likeness' of organic compounds, either by an empirical set of rules [16], a detailed substructure analysis using a genetic algorithm approach [112], their match to frequently occurring molecular frameworks [113] or the use of scoring schemes to discriminate between drugs and non drugs [114, 115]. This latter promising method is based on simple structural parameters like extended atom types [71] or short selected fragments, which are subjected to data-analysis using neural networks to account for non linearity. These approaches are effective in providing a reasonable ranking of compounds in a virtual or existing library, so that synthetic and screening resources can be focused on compound subsets of general biological interest. We are typically using a combined method based on: (a) acceptable ranges for computed physico-chemical and structural properties derived from databases containing drug-like molecules; (b) rejection of molecules with undesirable substructures; and (c) scoring using an internal implementation of the drug-like classification scheme [116].

Another point is to evaluate which of many library proposals is more diverse and/or better suited for a particular screening application. Here, validated quantitative measures of diversity are routinely applied to assess diversity, representativity plus complementarity to any corporate compound collection. It must be decided whether it is more effective to manufacture a sublibrary variation of an existing library, or to establish a new synthesis scheme with a novel scaffold.

Another important consideration in the library planning stage is whether to use diversity-based selection on reactants or on the entire virtual product matrix [92, 96]. It has been shown previously that the most diverse library will be obtained by a pure diversity selection on the full product matrix. However, this type of selection is sometimes not very practical, as the combinatorial scheme is violated – not all building blocks for attachment point A will combine with all reactants B. In contrast, parallel synthesis schemes are compatible with this pure diversity selection, when it is possible to solve the logistic problems [14, 117]. Gillet et al. [92] demonstrated that a product-based design of true combinatorial libraries using a genetic algorithm results in more diverse libraries than a simple reactant-based design, while the maximum diversity can only be achieved by product-based pure diversity selection, disregarding the combinatorial scheme. Different solutions for generic library design under combinatorial constraints were obtained by other groups [118].

Thus, diversity-based selections are utilized to select an optimal generic library from the underlying virtual library, in some cases under combinatorial constraints. In contrast, different global similarity searching strategies around earlier hits are applied to design biased or focused libraries. The most detailed level of information can be incorporated by scoring a virtual library utilizing 2D or 3D-QSAR models, pharmacophore models, or direct protein–ligand interaction scoring functions [119] after flexible docking. However, scoring of virtual libraries based on detailed 3D structure information is time-consuming, and thus an effective filtering strategy to reduce the size of the underlying library based on simple parameters is essential [120].

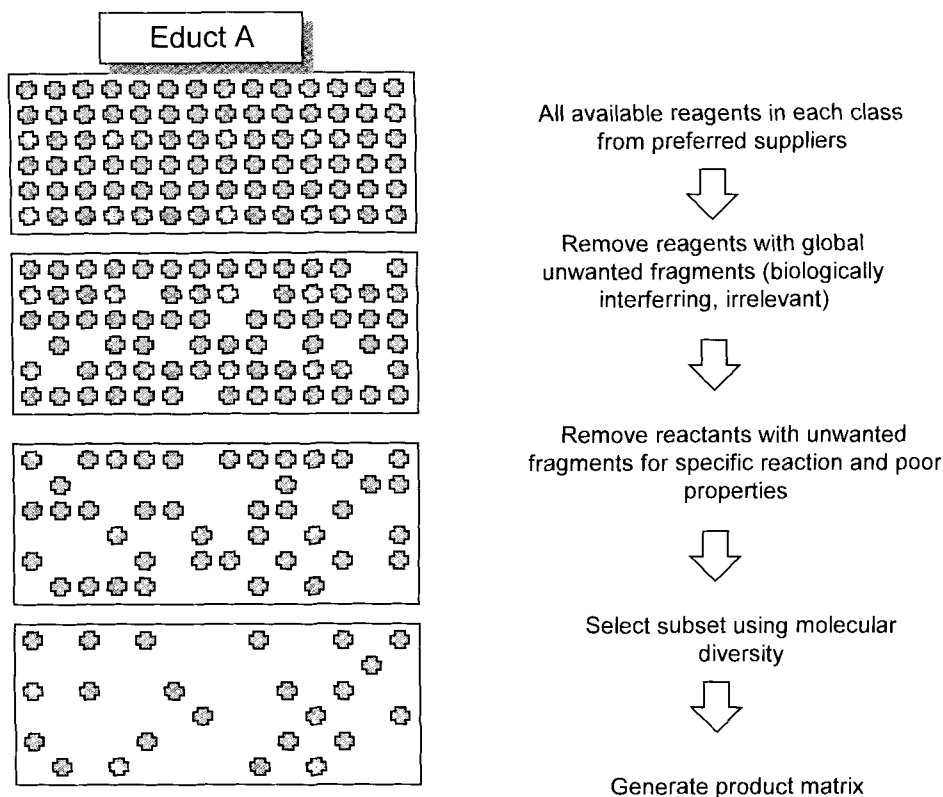
### 13.7.2 Practical Issues

Any research group involved in parallel synthesis or combinatorial chemistry will use different individual steps during the phases to design compound libraries according to their personal set-up, validation studies, experiences with descriptors and techniques, and their focus on particular library types. Clearly, there is no single accepted strategy in the literature, although a better understanding of the term ‘molecular diversity’ emerged from many studies in various groups. Thus, any comments on individual steps given below relate to own experience in library design. The general procedure for computer-assisted small-molecule library design, subdivided into two separate parts outlined below, assumes that for a new library reaction validation studies have been carried out:

#### **Scheme 1:** Individual Steps in Computer-Assisted Library Design:

Reagents:

- Search for reagents in internal and external databases like ACD [121]
- Clean up reactant lists (duplicates, counterions, suppliers, price)



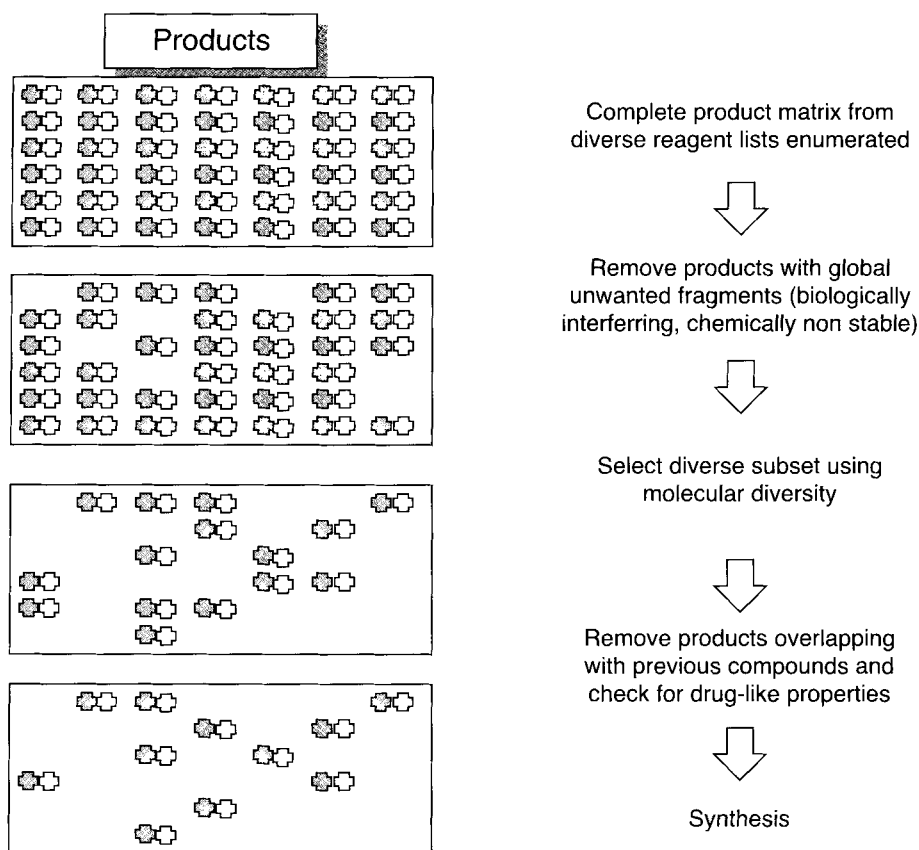
**Figure 13.9.** Schematic flow to illustrate individual steps related to educt selection in library design.

- Remove reagents with unwanted substructures
- Remove reagents with non drug-like properties
- Select reagents for a specific reaction after constraints from reaction validation
- Compute diversity descriptors and select diverse or similar reactants according to type of library

**Products:**

- Combine reactants to virtual product matrix
- Select products according to drug-like definitions
- Compute diversity descriptors and select products based on diversity or similarity considerations according to type of library
- Reject undesirable products by physico-chemical properties and unwanted substructures
- Deliver designed library to chemistry including plate-layout depending on synthesis scheme

Many parts of this procedure shown in Figures 13.9 and 13.10 have been automated to allow for a seamless flow of information. The undesirability of certain reagents is defined by the presence of substructures that are known to have toxic side effects and/or be metabolically hazardous [122], as well as by the presence of groups that could interfere with or prevent the reaction. Many companies are using lists compiled from analysis of earlier internal screening data. Furthermore, critical information from the chemistry (reaction



**Figure 13.10.** Schematic flow on product-based filter and selection steps prior to chemical synthesis.

validation) has to be incorporated into this procedure in order to exclude reaction-specific unwanted substructures based on the likelihood of there being significant proportions of by-products, the inertness of reacting groups, the reactivity of non reacting functional groups, the stability of products, or solubility problems connected with extraction. Rejected and accepted compounds are stored in different files with physico-chemical descriptions, such that a later selection is possible at every stage. This is especially useful if, at a later stage, it transpires that a particular educt is no longer commercially available, or delivery times are too long. Furthermore, this set-up allows additional reactivity information to be incorporated at a later stage. Such a refined assortment of small, reactive building blocks in accord with a reaction protocol is now the essential starting point for diversity or similarity analyses, depending on the type and motivation for a planned combinatorial or parallel synthesis library. As pointed out earlier, for targeted libraries, different descriptors known to be relevant in a particular biological system are chosen, while for the design of generic libraries, 2D fingerprints in combination with 3D molecular fields are used for selecting a diverse array of potential reagents. The final acquisition and main-

tenance of a building block collection is a resource-intensive activity. Hence, such a collection will be used more than once by integrating them into the design of different libraries around alternative scaffolds with alternative reaction schemes to end up with a diverse portfolio of complementary libraries.

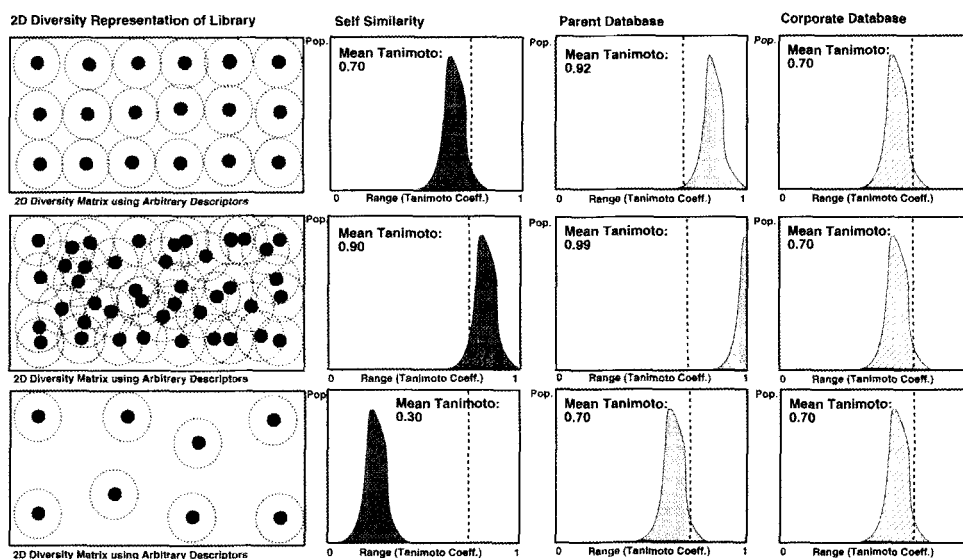
Important points to consider when compiling building block lists are whether to use commercially available reagents only, or whether to incorporate building blocks from other sources such as internal or external custom synthesis. Such a synthesis of building blocks itself can be subject of an automated parallel synthesis program. Proprietary building blocks offer certain advantages, but it requires a huge effort to establish such a tailored collection. Furthermore, many researchers suggest the incorporation of building blocks derived from historical medicinal chemistry knowledge (privileged structures) [123, 124]. In fact, there is no general answer to the question of which type of reagents should be used for a particular library synthesis project. This answer depends on the library type and motivation, the information accumulated regarding the biological target, and the logistics and availability of particular educts.

All acceptable reagents are then combined to form a virtual combinatorial library. The theoretical number of different individual compounds contained in such a library prepared by an ideal synthesis scheme is given by two factors: (i) the number of building blocks for each individual step; and (ii) the total number of steps. Efficient searching first rejects compounds using unwanted substructure queries and simple property ranges such as molecular weight, computed log P, number of hydrogen bond donors, acceptors and others. A scoring system might then be applied to classify drug-like and non drug-like compounds according to simple substructure parameters and a neural network approach [114, 115], before diversity considerations are taken into account, again depending on the type of library to be generated (Fig. 13.10).

### 13.7.3 Analysis of Diverse Libraries

Ideally, any diversity-based design using a similarity radius-based concept should result in a diverse, but representative child library. Not too few compounds should be selected to maintain information from the entire virtual library. Database self-similarity and comparison plots are useful tools in this context for analysis to monitor diversity, representivity and complementarity to a corporate collection, as shown schematically in Figure 13.11.

For this application the Tanimoto coefficient (or any other appropriate metric) for every database structure to its nearest neighbor is computed and used to generate a histogram. The maximum Tanimoto coefficient represents the closest pair of any two compounds in the dataset. For an idealized optimal diverse database selection using a Tanimoto coefficient of, *e.g.*, 0.85 defining the boundary between similar and diverse (Fig. 13.11 upper panel), a self-similarity plot (central left column) with a peak on the left-most side is expected, while the comparison with the parent virtual library should not reveal any voids, leading to a histogram (central right column) with a peak and a cut-off at on the right side of this similarity radius (dotted vertical line). For a very redundant database on the other hand (central panel), both plots show maxima on the right side of the similarity



**Figure 13.11.** Database comparison histograms to illustrate an optimal diverse database selection (upper panel), a highly redundant database selection (middle panel), and a database selection with loss of information (lower panel). The left column show simplified representations of databases as distribution of molecules (filled circles) in an arbitrary 2D molecular property space. In the middle left column, idealized self-similar histograms are given, while the plots in the middle right column show plots obtained by comparing the database subset to the entire database. The right column refers to plots obtained by comparison to a corporate database. Dotted vertical lines indicate the similarity radius for a particular descriptor.

radius. While such a non redundant database selection (lower panel) is very diverse (central plot, maximum on the left side), too few selected compounds would lead to a loss of information (right plot, maxima or cutoff on the left side of the similarity radius). These analysis tools can also be used to monitor complementarity to corporate or other third-party databases. Here, the parent database is replaced by this additional database and a good fit is indicated by a peak of this distribution, preferably on the right side of the dotted vertical line (right-most column). Many other visual and numerical analysis tools have also been reported to be very useful to monitor design success for diverse libraries in the literature (see above).

## 13.8 Conclusions

The rational design of combinatorial libraries and assessment of compound databases is an important issue to optimize resources in today's discovery strategies. There is a growing interest in augmenting the structural and biological diversity in databases, and how to relate chemical with biological diversity. Many approaches have been reported towards the design of combinatorial libraries, molecular diversity considerations, and vali-



dation on commonly used descriptors and selection techniques. Quantifying diversity has been shown to be crucial for the design and selection of representative subsets from larger collections. Those strategies are also required to compare libraries with regard to diversity, representivity and complementarity.

This chapter has summarized some approaches and applications in virtual library design towards the generation of screening libraries and libraries directed at optimizing a particular biological activity. The view presented here is a subjective one, and our own results are set in context to those of published studies. Several researchers have correspondingly reported that 2D substructure information is indeed important for valid descriptors, while it also was shown that design techniques perform significantly better than a pure random selection. There is a need for objective descriptor validation and comparison methods. In this chapter, three of these were highlighted: the sampling of biological classes for diverse library design; enrichment factors; and neighborhood plots.

The question of 2D versus 3D descriptors remains highly controversial, though it may be expected that more effort in the development of powerful descriptors will enhance success rates for generic library design. In general, the question arises whether public datasets for assessment are biased towards 2D similarity, as they are mainly compiled according to biologically successful chemistry projects. Chemically feasible variations around a scaffold for a particular target often translates to common 2D structural motifs for biological activity.

Furthermore, there is a gap between modern QSAR methods to predict binding affinities for smaller datasets and molecular descriptors for database searching and library design. Some approaches highlighted in this chapter are aiming to close this gap by the use of efficient algorithms to analyze larger datasets. In addition, the combinatorial nature of those libraries introduces some structure into the computational search space of compound collections. Because of a common building scheme, together with common building blocks, information relating to a molecule can be gained from previously considered molecules in a search procedure, which certainly will lead to more sophisticated algorithms for the design and analysis of combinatorial libraries.

Not only diversity, but also physico-chemical properties, when correlated with pharmacokinetic parameters [125] or a more general description of 'drug-likeness' of a candidate molecule, will become increasingly important in any design or selection strategy to ensure the generation of relevant compounds, thus supplementing current library design strategies. Practicability and automation of the design process are other key issues for the successful integration of combinatorial chemistry into the drug design process.

Thus, there is a broad variation of approaches for addressing molecular diversity, and these have been applied successfully in many projects. The emerging detailed picture leads to an understanding of the correlation between chemical and biological diversity. The design of suitable libraries for different purposes becomes possible due to the integration of innovative approaches, along with molecular modeling techniques on one side, but with broad medicinal chemistry experience on the other. The result is the creation of successful libraries for screening which will have great impact on research efficiency. As synthetic and computational chemists have gained experience with library design and production, it is now becoming possible to design and produce tailored libraries for lead identification, and optimization directed towards many biological targets.

## 13.9 Acknowledgements

The authors especially thank W. Guba, K.H. Baringhaus, C. Giegerich, S. Henke, C. Turpin, R. Vaz, E. Chan (Hoechst Marion Roussel), T. Pötter (Bayer AG), H. Briem (Boehringer Ingelheim KG), B. Wendt, R.D. Cramer, D.E. Patterson, R.D. Clark, D. Lassen (Tripos) and J.S. Dixon (Metaphorics Inc.) for helpful discussions and collaborations on various issues of molecular similarity and diversity, library design and descriptor validation.

## 13.10 References

- [1] Combinatorial Peptide and Nonpeptide Libraries, G. Jung, Ed. VCH, Weinheim 1996.
- [2] Combinatorial Chemistry and Molecular Diversity in Drug Discovery, Gordon, E.M.; Kerwin, J.F. Jr (Eds.) Wiley, New York 1998.
- [3] a) Gallop, M.A.; Barrett, R.W.; Dower, W.J.; Fodor, S.P.A.; Gordon, E.M. *J. Med. Chem.* **1994**, *37*, 1233–1251.  
b) Gordon, E.M.; Barrett, R.W.; Dower, W.J.; Fodor, S.P.A.; Gallop, M.A. *J. Med. Chem.* **1994**, *37*, 1385–1399.
- [4] a) Madden, D.; Krchnak, V.; Lebl, M. *Persp. Drug Discovery and Design* **1995**, *2*, 269–285.  
b) Ellman, J.A. *Acc. Chem. Res.* **1996**, *29*, 132–143.  
c) Gordon, E.M.; Gallop, M.A.; Patel, D.V. *Acc. Chem. Res.* **1996**, *29*, 144–154.  
d) Fruchtel, J.S.; Jung, G. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 17–42.  
e) Thompson, L.A.; Ellman, J.A. *Chem. Rev.* **1996**, *96*, 555–600.  
f) Choong, I.C.; Ellman, J.A. *Annu. Rep. Med. Chem.* **1996**, *31*, 309–318.  
g) Balkenhohl, F.; von dem Bussche-Hünnefeld, C.; Lansky, A.; Zechel, C. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2288–2337.  
h) Nefzi, A.; Ostresh, J.M.; Houghten, R.A. *Chem. Rev.* **1997**, *97*, 449–472.  
i) Bristol, J.A. (Ed.) *Tetrahedron* **1997**, *53*, 6573–6706.
- [5] Ash, J.E.; Warr, W.A.; Willett, P. *Chemical Information Systems*; Ellis Horwood: Chichester 1991.
- [6] Shemetulskis, N.E.; Dunbar, J.B.; Dunbar, B.W.; Moreland, D.W.; Humblet, C. *J. Comp. Aided Mol. Design* **1995**, *9*, 407–416.
- [7] Warr, W.A. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 134–140.
- [8] Martin, E.J.; Critchlow, R.E.; Spellmeyer, D.C.; Rosenberg, S.; Spcar, K.L.; Blaney, J.M. in *Trends in Drug Research II*, van der Goot, H. (Ed.) 1998, Elsevier, pp 133–146.
- [9] Pavia, M.R.; Sawyer, T.K.; Moos, W.H. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 387–396.
- [i0] a) Baum, R.M. *Chem. & Eng. News*, Febr. 7. **1994**, 20–26.  
b) Alper, J. *Science* **1994**, 264, 1399–1401.
- [11] Martin, E.J.; Blaney, J.M.; Siani, M.A.; Spellmeyer, D.C.; Wong; A.K. Moos, W.H. *J. Med. Chem.* **1995**, *38*, 1431–1436.
- [12] Kubinyi, H. *Curr. Opin. Drug. Disc. Dev.* **1998**, *1*, 16–27.
- [13] Spilker, B.; Cuatrecasas, P. *Inside the Drug Industry*, Prous Science Publ., Barcelona, 1990.
- [14] Ferguson, A.M.; Patterson, D.E.; Garr, C.; Underiner, T. *J. Biomol. Screen.* **1996**, *1*, 65–73.
- [15] Matter, H.; Lassen, D. *Chim. Oggi* **1996**, *14*, 6, 9–15.
- [16] Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. *Adv. Drug Delivery Rev.* **1997**, *23*, 3–25.
- [17] a) *3D-QSAR in Drug Design. Theory, Methods and Applications*, Kubinyi, H., Ed. ESCOM, Leiden (NL), 1993.  
b) Kubinyi, H.; Folkers, G.; Martin, Y.C. (Eds.): *Drug Design*, Vol 2, Dordrecht, Kluwer/ ESCOM; 1998.  
c) Hansch, C.; Leo, A. *Exploring QSAR. Fundamental and Applications in Chemistry and Biology*. Washington, DC, American Chemical Society; 1995.
- [18] a) Martin, Y.C.; Bures, M.G.; Danaher, E.A.; DeLazzer, J.; Lico, I.; Pavlik, P.A. *J. Comp.-Aided Mol. Design* **1993**, *7*, 83–102.  
b) Jones, G.; Willett, P.; Glen, R.C. *J. Comp.-Aided Mol. Des.* **1995**, *9*, 532–549.  
c) Greene, J.; Kahn, S.; Savoj, H.; Sprague, P.; Teig, S. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1297–1308.

- [19] PDB Database (<http://www.pdb.bnl.gov>): Bernstein, F.C.; Koetzle, T.F.; Williams, G.J.B.; Meyer, E.F.; Brice, M.D.; Rodgers, J.R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535–542.
- [20] a) Zheng, W.; Cho, S.J.; Tropsha, A. *Int. J. Quant. Chem.* **1998**, *69*, 65–75.  
b) Zheng, W.; Cho, S.J.; Tropsha, A. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 251–258.  
c) Cho, S.J.; Zheng, W.; Tropsha, A. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 259–268.
- [21] Li, J.; Murray, C.W.; Waszkowycz, B.; Young, S.C. *Drug Disc. Today* **1998**, *3*, 105–112.
- [22] a) Kick, E.K.; Roe, D.C.; Skillman, A.G.; Lin, G.; Ewing, T.J.A.; Kuntz, I.D. *Chem. Biol.* **1997**, *4*, 297–307.  
b) Roe, D.C.; Kuntz, I.D. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 269–282.
- [23] Rockwell, A.; Melden, M.; Copeland, R.A.; Hardman, K.; Decicco, C.P.; DeGrado, W.F. *J. Am. Chem. Soc.* **1996**, *118*, 10337–10338.
- [24] Saleme, F.R.; Spurlino, J.; Bone, R. *Curr. Biol.* **1997**, *5*, 319–324.
- [25] Liu, D.X.; Jiang, H.; Chen, K.; Ji, R. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 233–242.
- [26] Norman, T.C.; Gray, N.S.; Koh, J.T.; Schultz, P.G. *J. Am. Chem. Soc.* **1996**, *118*, 7430–7431.
- [27] Van Drie, J.H.; Lajiness, M.S. *Drug Disc. Today* **1998**, *3*, 274–283.
- [28] Cramer, R.D.; Patterson, D.E.; Clark, R.D.; Soltanshahi, F.; Lawless, M.S. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1010–1023.
- [29] Willett, P.; Barnard, J.M.; Downs, G.M. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 983–996 and many references in this review.
- [30] Wild, D.J.; Willett, P. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 159–167.
- [31] Ashton, M.J.; Jaye, M.C.; Mason, J.S. *Drug Disc. Today* **1996**, *1*, 71–78.
- [32] Bures, M.G.; Martin, Y.C. *Curr. Opin. Chem. Biol.* **1998**, *2*, 376–380.
- [33] Willett, P. (Ed.), *Computational Methods for the Analysis of Molecular Diversity*, Vol. 8, Dordrecht, Kluwer/ESCOM; 1997.
- [34] Dean P.M. (Ed.), *Molecular Similarity in Drug Design*, Chapman and Hall, London 1995.
- [35] Maggiora, G.M.; Johnson, M.A. *Concepts and Applications of Molecular Similarity*, Wiley, New York, 1990 pp. 99–117.
- [36] Kubinyi, H. In *3D-QSAR in Drug Design*, vol 2. Kubinyi, H.; Folkers, G.; Martin, Y.C. (Eds.), Kluwer, Dordrecht 1998, pp 225–252.
- [37] Willett, P. *Similarity and Clustering in Chemical Information Systems*. (1987), Letchworth: Research Studies Press.
- [38] Brown, R.D.; Martin, Y.C. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 1–9.
- [39] UNITY 4.0, Chemical Information Software, Tripos Inc., 1699 S. Hanley Road, St. Louis, MO 63144, USA.
- [40] Daylight Chemical Information Systems, Inc.; 3951 Claremont Street, Irvine, CA 92714.
- [41] ISIS/Base 2.1.3.; Molecular Design Ltd.; 14600 Catalina Street, San Leandro, CA 94577.
- [42] a) Willett, P.; Winterman, V.; Bawden, D. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 109–118.  
b) Bawden, D. in *Chemical Structures 2* (Warr, W.A., Ed.) Springer, 1990, pp. 173–176.
- [43] Flower, D.R. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 379–386.
- [44] Ash, S.; Cline, M.A.; Homer, W.; Hurst, T.; Smith, G.B. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 71–79.
- [45] Knuth, D.E. *Sorting and searching*; Addison-Wesley: Reading, MA.
- [46] Brown, R. D.; Bures, M. G.; Martin, Y. C. American Chemical Society Meeting, Anaheim, CA, 1995; COMP3.
- [47] Brown, R. D.; Martin, Y. C. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 572–584.
- [48] a) Adamson, G.W.; Cowell, J.; Lynch, M.F.; McLure, A.H.W.; Town, W.G.; Yapp, A.M. *J. Chem. Doc.* **1973**, *13*, 153–157.  
b) Adamson, G.W.; Bush, J.A.; McLure, A.H.W.; Lynch, M.F. *J. Chem. Doc.* **1974**, *14*, 44–48.  
c) Hodes, L.J. *J. Chem. Inf. Comput. Sci.* **1976**, *16*, 88–93.  
d) Hodes, L.J.; Feldman, A. *J. Chem. Inf. Comput. Sci.* **1978**, *18*, 96–100.
- [49] McGregor, M.J.; Pallai, P.V. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 443–448.
- [50] Willett, P.; Winterman, V. *Quant. Struct. Activ. Relat.* **1986**, *5*, 18–25.
- [51] Holliday, J.D.; Ranade, S.S.; Willett, P. *QSAR*, **1996**, *14*, 501–506.
- [52] Ashton, M.J.; Jaye, M.C.; Mason, J.S. *Drug Disc. Today* **1996**, *2*, 71–78.
- [53] Pickett, S.D.; Mason, J.S.; McLay, I.M. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1214–1223.
- [54] Mason, J.S.; Pickett, S.D. *Perspect. Drug Des. Discovery* **1997**, *7*, 1–29.
- [55] Brown, R.D.; Martin, Y.C. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 572–584.

- [56] Davies, K. In *Molecular Diversity and Combinatorial Chemistry: Libraries and Drug Discovery*; Chaiken, I.M., Janda, K.D., Eds.; American Chemical Society: Washington D.C., 1996; pp 309–316.
- [57] ChemDiverse from Chemical Design Ltd., Roundway House, Cromwell Park, Chipping Norton, Oxon, OX7, 5SR, UK.
- [58] Selector from TRIPOS Inc., 1699 South Hanley Road, St. Louis MO 63144, U.S.A.
- [59] Mason, J.S. *J. Mol. Graphics Mod.* **1998**, *16*, 51.
- [60] Brown, R.D. *Perspect. Drug Disc. Des.* **1997**, *7/8*, 31–49.
- [61] Martin, Y.C.; Brown, R.D.; Bures, M.G. In *Combinatorial Chemistry and Molecular Diversity in Drug Discovery*, Gordon, E.M.; Kerwin, J.F. Jr. (Eds.) Wiley, New York 1998, pp 369–385.
- [62] Carhart, R.E.; Smith, D.H.; Venkataraghavan, R. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 64–73.
- [63] Sheridan, R.P.; Nachbar, R.B.; Bush, B.L. *J. Comp.-Aided Mol. Design*, **1994**, *8*, 323–340.
- [64] a) Sheridan, R.P.; Miller, M.D.; Underwood, D.J.; Kearsley, S.K. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 128–136.  
b) Kearsley, S.K.; Sallamack, S.; Fluder, E.M.; Androse, J.D.; Mosley, R.T.; Sheridan, R.P. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 118–127.
- [65] Kier, L.B.; Hall, L.H. *Molecular Connectivity and Drug Research*, Academic Press, NY, USA, 1976.
- [66] Hall, L.H.; Kier, L.B. In *Reviews in Computational Chemistry*, vol. 2; Lipkowitz, K.B.; Boyd, D.B., Eds.; New York, VCH, pp 367–422.
- [67] a) Hall, L.H.; Mohny, B.; Kier L.B. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 76–82.  
b) Gombar, V.K.; Enslein, K. *Quant. Struct.-Act. Relat.* **1990**, *9*, 321–325.
- [68] a) Gombar, V.K.; Jain, D.V.S. *Ind. J. Chem.* **1987**, *26A*, 554–555.  
b) Kier, L.B. *Medicinal Research Reviews* **1987**, *7*, 417–440.
- [69] a) Stahle, L.; Wold, S. in *Progress in Medicinal Chemistry*, Eds. Ellis, G.P., West, G.B., Elsevier 1988, pp 292–338.  
b) Wold, S.; Albano, C.; Dunn, W.J.; III, Edlund, U.; Esbensen, K.; Geladi, P.; Hellberg, S.; Johansson, E.; Lindberg, W.; Sjöström, M. in: *Chemometrics: Mathematics and Statistics in Chemistry*, Kowalski, B.R. (Ed.) NATO, ISI Series C 138, D. Reidel Publ. Co., Dordrecht, Holland 1984, pp 17–96.  
c) Rännar, S.; Lindgren, F.; Geladi, P.; Wold, S. *J. Chemometrics* **1994**, *8*, 111–125.
- [70] a) Moreau, G.; Broto, P. *Nouv. J. Chim.* **1980**, *4*, 359–360.  
b) Moreau, G.; Broto, P. *Nouv. J. Chim.* **1980**, *4*, 757–764.  
c) Broto, P.; Moreau, G.; Vandycke, C. *Eur. J. Med. Chem. Chim. Ther.* **1984**, *19*, 66–70.  
d) Moreau, G.; Turpin, C. *Analysis* **1996**, *24*, M17–M22.
- [71] a) Ghose, A.; Crippen, G. *J. Comp. Chem.* **1986**, *7*, 565–577.  
b) Viswanadhan, V.N.; Ghose, A.K.; Revankar, G.R.; Robins, R.K. *J. Chem. Inf. Comput. Sci.*, **1989**, *29*, 163–172.
- [72] Wagener, M.; Sadowski, J.; Gasteiger J. *J. Am. Chem. Soc.* **1995**, *117*, 7769–7775.
- [73] Sadowski, J.; Wagener, M.; Gasteiger, J. *Angew. Chem., Int. Ed. Engl.* **1996**, *34*, 23–24.
- [74] Only available in UNITY version 2.5, Chemical Information Software, Tripos Inc., 1699 S. Hanley Road, St. Louis, MO 63144, USA.
- [75] Cramer, R.D.; Patterson, D.E.; Bunce, J.E. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
- [76] Cramer, R.D.; Clark, R.D.; Patterson, D.E.; Ferguson, A.M. *J. Med. Chem.* **1996**, *39*, 3060–3069.
- [77] a) Todeschini, R.; Lasagni M.; Marengo, E. *J. Chemometrics*, **1994**, *8*, 263–272.  
b) Todeschini, R.; Gramatica, P.; Provenzani, R.; Marengo, E. *Chemometr. Intelligent Lab. Systems* **1995**, *27*, 221–229.  
c) Todeschini, R.; Bettiol, C.; Giurin, G.; Gramatica, P.; Miana, P.; Argese, E. *Chemosphere* **1996**, *33*, 71–79.
- [78] Lewis, R.A.; McLay, I.M.; Mason, J.S. *Chem. Des. Autom. News* **1995**, *10* (4), 37–38.
- [79] a) Shemetulskis, N.E.; Weininger, D.; Blankley, C.J.; Yang, J.J.; Humblet, C. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 862–871.  
b) Nilakantan, R.; Bauman, N.; Haraki, K.S. *J. Comp.-Aided Mol. Des.* **1997**, *11*, 447–452.
- [80] Chapman, D. *J. Comp.-Aided Mol. Des.* **1996**, *10*, 501–512.
- [81] Barnard, J.M.; Downs, G.M. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 644–649.
- [82] Everitt, B.S. *Cluster Analysis*; 3<sup>rd</sup> ed.; Arnold, London 1993.
- [83] Lajiness, M.; Johnson, M.A.; Maggiora, G.M. In: *QSAR: Quantitative Structure-Activity Relationships in Drug Design*; Fauchere, J.L., Ed.; Alan R. Liss Inc.: New York, USA, 1989; pp 173–176.

- [84] Lajiness, M.S. *Perspect. Drug Disc. Des.* **1997**, 7/8, 65–84.
- [85] Taylor, R. *J. Chem. Inf. Comput. Sci.* **1995**, 35, 59–67.
- [86] Holliday, J.D.; Ranade, S.S.; Willett, P. *Quant. Struct.-Act. Relat.* **1995**, 14, 501–506.
- [87] a) Pearlman, R.S. *Network Sci.* **1996**, June (<http://www.netsci.org>).  
b) Pearlman, R.S.; Smith, K.M. *Perspect. Drug Disc. Des.* **1998**, 9, 339–353.
- [88] Lewis, R.A.; Mason, J.S.; McLay, I.M. *J. Chem. Inf. Comput. Sci.* **1997**, 37, 599–614.
- [89] Snarey, M.; Terrett, N.K.; Willett, P.; Wilton, D.J. *J. Mol. Graphics* **1997**, 15, 372–385.
- [90] Clark, R.D. *J. Chem. Inf. Comput. Sci.* **1997**, 37, 1181–1188.
- [91] Clark, R.D.; Langton, W.J. *J. Chem. Inf. Comput. Sci.* **1998**, 38, 1079–1086.
- [92] Gillet, V.J.; Willett, P.; Bradshaw, J. *J. Chem. Inf. Comput. Sci.* **1997**, 37, 731–740.
- [93] Brown, R.D.; Martin, Y.C. *J. Med. Chem.* **1997**, 40, 2304–2313.
- [94] Weber, L.; Wallbaum, S.; Broger, C.; Gubernator, K. *Angew. Chem. Int. Ed. Engl.* **1995**, 34, 2281–2282.
- [95] Weber, L. *Drug Disc. Today* **1998**, 3, 379–385.
- [96] Good, A.C.; Lewis, R.A. *J. Med. Chem.* **1997**, 40, 3926–3936.
- [97] Kauvar, L.M.; Higgins, D.L.; Villar, H.O.; Sportsman, J.R.; Engvist-Goldstein, A.; Bukar, R.; Bauer, K.E.; Dille, H.; Rocke, D.M. *Chemistry & Biology* **1995**, 2, 107–118.
- [98] Dixon, S.L.; Villar, H.O. *J. Chem. Inf. Comput. Sci.* **1998**, 38, 1192–1203.
- [99] Briem, H.; Kuntz, I.D. *J. Med. Chem.* **1996**, 39, 3401–3408.
- [100] a) Kuntz, I.D.; Blaney, J.M.; Oatley, S.J.; Langridge, R.L.; Ferrin, T.E. *J. Mol. Biol.* **1982**, 161, 269–288.  
b) Shoichet, B.K.; Bodian, D.L.; Kuntz, I.D. *J. Comp. Chem.* **1992**, 13, 380–397.  
c) Meng, E.C.; Shoichet, B.K.; Kuntz, I.D. *J. Comp. Chem.* **1992**, 13, 505–524.
- [101] a) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996**, 261, 470–489.  
b) Rarey, M.; Kramer, B.; Lengauer, T. *J. Comp.-Aided Mol. Des.* **1997**, 11, 369–384.  
c) Makino, S.; Kuntz, I.D. *J. Comp. Chem.* **1997**, 18, 1812–1825.
- [102] Briem, H. personal communication.
- [103] Rarey, M.; Dixon, J.S. *J. Comp.-Aided Mol. Des.* **1998**, 12, 471–490.
- [104] Lemmen, C.; Lengauer, T. In *Molecular Modelling and Prediction of Bioactivity, Proceedings of the 12th European Symposium on Quantitative Structure-Activity Relationships (QSAR'98)* Gunder-tofte, K.; Jorgensen, F.S. (Eds.), Plenum Press, New York 1999, in press.
- [105] Lemmen, C.; Hiller, C.; Lengauer, T. *J. Comp.-Aided Mol. Des.* **1998**, 12, 491–502.
- [106] Pötter, T.; Matter, H. *J. Med. Chem.* **1998**, 41, 478–488.
- [107] Matter, H. *J. Med. Chem.* **1997**, 40, 1219–1229.
- [108] Index Chemical Database - Subset from 1993, Institute for Scientific Information, Inc. (ISI), 3501 Market Street, Philadelphia, PA, USA.
- [109] Matter, H.; Pötter, T. in preparation.
- [110] Patterson, D.E.; Cramer, R.D.; Ferguson, A. M.; Clark, R.D.; Weinberger, L.E. *J. Med. Chem.* **1996**, 39, 3049–3059.
- [111] DePriest, S.A.; Mayer, D.; Naylor, C.B.; Marshall, G.R. *J. Am. Chem. Soc.* **1993**, 115, 5372–5384.
- [112] Gillet, V.J.; Willett, P.; Bradshaw, J. *J. Chem. Inf. Comput. Sci.* **1998**, 38, 165–179.
- [113] Murcko, M.A.; Bemis, G.A. *J. Med. Chem.* **1996**, 39, 2887–2893.
- [114] Sadowski, J.; Kubinyi, H. *J. Med. Chem.* **1998**, 41, 3325–3329.
- [115] Ajay, Walters, W.P.; Murcko, M.A. *J. Med. Chem.* **1998**, 18, 3314–3324.
- [116] Matter, H.; Giegerich, C. unpublished results.
- [117] Garr, C.D.; Peterson, J.R.; Schultz, L.; Oliver, A.R.; Underiner, T.L.; Cramer, R.D.; Ferguson, A.M.; Lawless, M.S.; Patterson, D.E. *J. Biomol. Screen.* **1996**, 1, 179–186.
- [118] Pötter, T. personal communication.
- [119] a) Hirst, J.D. *Curr. Opin. Drug Disc. Dev.* **1998**, 1, 28–33.  
b) Böhm, H.J. *J. Comp.-Aided Mol. Des.* **1998**, 12, 309–322.  
c) Böhm, H.J. *J. Comp.-Aided Mol. Des.* **1994**, 8, 243–256.  
d) Eldridge, M.D.; Murray, C.W.; Auton, T.R.; Paolini, G.V.; Mee, R.P. *Comp.-Aided Mol. Des.* **1997**, 11, 425–445.  
e) Jain, A.N. *J. Comp.-Aided Mol. Des.* **1996**, 10, 427–440.  
f) Oprea, T.I.; Marshall, G.R. In *3D-QSAR in Drug Design*, vol 2. Kubinyi, H.; Folkers, G.; Martin, Y.C. (Eds.), Kluwer, Dordrecht 1998, pp 35–61.  
g) Head, R.D.; Smythe, M.L.; Oprea, T.I.; Waller, C.L.; Green, S.M.; Marshall, G.R. *J. Am. Chem. Soc.* **1996**, 118, 3959–3969.

- [120] Walters, W.P.; Stahl, M.T.; Murcko, M.A. *Drug Disc. Today* **1998**, 3, 160–178.
- [121] ACD: Available Chemical Directory, from Molecular Design Ltd.; 14600 Catalina Street, San Leandro, CA 94577.
- [122] Rishton, G.M. *Drug. Disc. Today* **1997**, 2, 382–384.
- [123] Gordon, E.M. In *Combinatorial Chemistry and Molecular Diversity in Drug Discovery*, Gordon, E.M.; Kerwin, J.F. Jr. (Eds.) Wiley, New York 1998, pp 17–38.
- [124] Lewell, X.Q.; Judd, D.B.; Watson, S.P.; Hann, M.M. *J. Chem. Inf. Comput. Sci.* **1998**, 38, 511–522.
- [125] Guba, W.; Cruciani, G. In *Molecular Modelling and Prediction of Bioactivity, Proceedings of the 12th European Symposium on Quantitative Structure-Activity Relationships (QSAR'98)* Gundertofte, K.; Jorgensen, F.S. (Eds.), Plenum Press, New York 1999, in press.

This Page Intentionally Left Blank

# 14 How to Scale-up while Scaling down: EVOscreen™, a Miniaturized Ultra-High-Throughput Screening System

*Rodney Turner, Sylvia Sterrer, Karl-Heinz Wiesmüller  
and Franz-Josef Meyer-Almes*

## 14.1 Introduction to Drug Discovery and Screening

### 14.1.1 A Rationale for New Methods in Drug Discovery

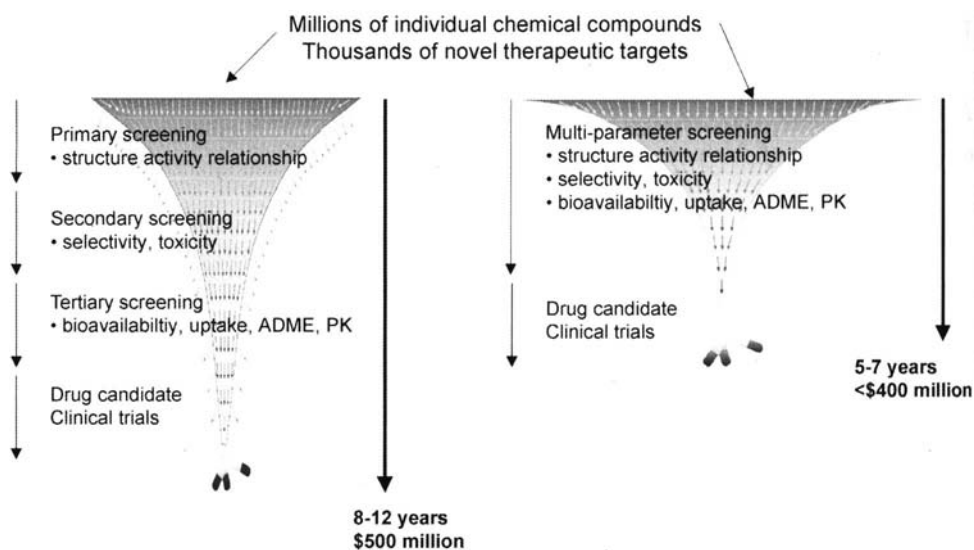
Combinatorial chemistry has been one of the most important ‘buzz-words’ in the drug discovery field for the past seven years [1–3]. Together with genomics and proteomics [4, 5] – the other driving force of the current biotechnology boom – combinatorial chemistry offers the promise of new and better medications. These two new driving forces promise to lead to an explosive increase in both the ability to identify new therapeutic targets (largely thanks to genomics), and the ability to provide tailor-made molecules which can rapidly enter the drug development pipeline (the promise of combinatorial chemistry). The simultaneous, dramatic increase in both the number of targets and potential hit compounds is the reason for what has been called a change in paradigm in drug discovery [6]. To capitalize fully on the promise of both genomics and combinatorial chemistry, the drug discovery process must change in a fundamental way.

Rather than screen 50 chemical compounds against a particular target, combinatorial approaches often require that thousands be screened. However, because of the preponderance of targets, that screening must take place in days rather than months. Out of necessity, high-throughput screening (HTS) will be a key factor in realizing the potential of new drug discovery strategies by building a bridge between the increased number of targets and the vast number of compounds to be screened.

Not only does screening throughput need to be increased, but novel assay techniques and detection technologies will be employed in order to provide more information about the suitability of specific compounds earlier in the screening process. This combination of quality and quantity must be achieved in order to take advantage fully of the increased potential for target identification offered by genomics and proteomics and the potential for direct, rapid access to novel chemical compounds offered by combinatorial chemistry and combinatorial biology [7]. Figure 14.1 illustrates the time and cost savings to be achieved by using novel technologies and miniaturization.

By employing new methods for the synthesis of discrete molecules, modern combinatorial chemistry has surpassed the capacity of traditional screening strategies in the search for new active substances. Whether screening libraries of high diversity, such as combinatorial peptide and organic compound libraries, or large libraries of individual natural or rationally designed compounds, modern screening strategies dictate that





**Figure 14.1.** A comparison of today's and tomorrow's drug discovery paradigm.

$10^5$ – $10^6$  compounds be screened with a single assay system. Using 96-well microtiter plate automated screening technology, a month would be required to screen a diversity of  $10^5$  at a rate of 50 plates per day. This rate must be cut by roughly a factor of ten for a research laboratory to remain competitive in today's pharmaceutical research environment.

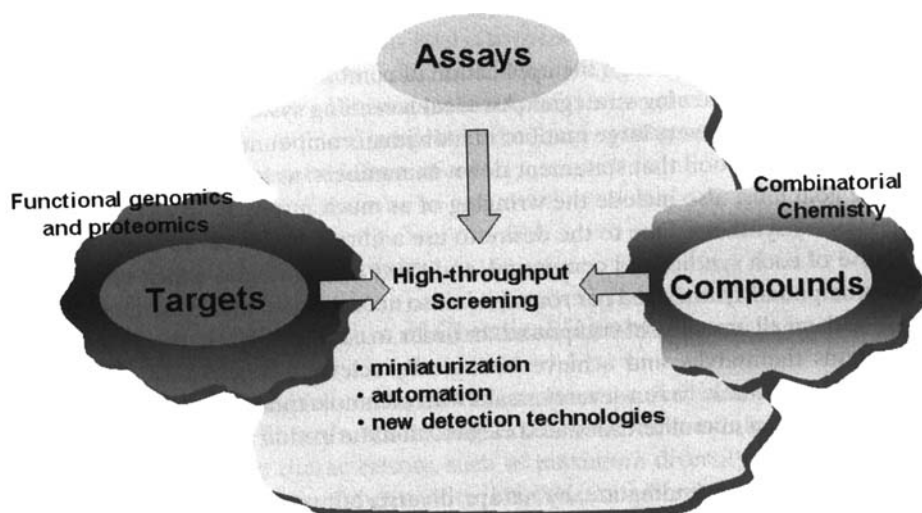
### 14.1.2 Chemistry, Biology and Technology

The necessary increase in screening throughput in light of combinatorial strategies brings several related aspects to the fore. At the screening rates enabled by high- and ultra-high-throughput applications, reagent consumption is a critical factor. By reducing the volume of reagent required for synthesis and the amount of compound required for analysis, miniaturization during both synthesis and screening phases can lead to significant savings over current methods [8]. Additionally, given the number of compounds to be screened, often of similar chemistry, novel assay techniques and detection technologies must be employed in order to provide better characterization of specific compounds, and this earlier in the screening process. A collection of synthetic compounds is likely to be used in several different assays – another reason to keep reagent consumption to a minimum. Multiplexing strategies, whereby multiple parameters can be measured by employing multiple assays and/or read-outs in a single sample, will also improve screening efficiency. This combination of quality and quantity must be achieved in order to take full

advantage of the increased potential for target identification offered by genomics and proteomics and the potential for direct, rapid access to novel chemical compounds offered by combinatorial chemistry and combinatorial biology [8]. High- and ultra-high-throughput screening (uHTS) will not only allow realization of the potential of these new technologies, it will also enable better usage of natural compound libraries – long a primary backbone of drug discovery. Screening, therefore, is the link between biology (targets) and chemistry (drug-like molecules). And in the age of genomics and combinatorial chemistry, that link must achieve high throughput. The unification of targets, compounds and assays in screening applications for drug discovery is illustrated in Figure 14.2.

The effective use of HTS for combinatorial applications depends on effective synergy between the traditional disciplines of drug discovery, biology, biochemistry and chemistry, the technologies associated with combinatorial chemistry [9], and such diverse disciplines as information technology, robotics, physics and fluidics. Screening was once the exclusive province of giants in the pharmaceutical industry. Today, screening is increasingly characterized by alliances between biotech ventures devoted to developing screening strategies, and research laboratories, whether they be large pharmaceutical firms or other biotech ventures, who provide a specific screening target. As these alliances are forged, the appropriate union between target, assay method and screening technology is of primary importance.

Miniaturization is an inherent characteristic of most modern HTS systems, but the detection technology is often largely unchanged from many of the read-out technologies



**Figure 14.2.** The integration of the three primary components of drug discovery screening: compounds, targets, and assays.

used in conventional assays. To provide the most effective synergy between screening technology and assay development, screening systems must provide detection technology which is matched to the demands of miniaturization, on the one hand, and offers significant flexibility of read-out methods, on the other. Fluorescence-based detection technologies are by far the choice of most modern HTS systems.

### 14.1.3 Sources of Compounds for HTS

Natural product libraries, shelf compounds, libraries from chemically modified natural products and combinatorial compound collections, including peptides, peptidomimetics and small organic molecules, are well established as sources for new lead structures. Solution-phase or solid-phase synthesis with a variety of chemistries and scaffolds allows the experienced chemist to produce libraries directed at almost any class of targets. Combinatorial chemistry is used in structure-based drug design, for lead optimization, and to generate highly diverse compound collections for random screening. Methods for their reproducible, effective and fully automated synthesis and analysis have been established and are dependent on the heterogeneity of the mixtures [10]. As the number of products accessible from a given set of components increases exponentially, synthesis strategies had to be developed that allow the highly parallel and simultaneous production of compounds which were generated and tested as mixtures or as individuals [11]. These strategies are supported by new methodologies for diversity measurement and compound selection [12]. Combinatorial and high-throughput synthesis has boosted the generation of compounds in many pharmaceutical companies 10- to 50-fold.

### 14.1.4 Demands on HTS

The diversity accessible through the application of combinatorial chemistry places significant demands on screening strategies. An ideal screening system must provide sufficient flexibility to allow a very large number of individual compounds to be tested against a variety of targets. To boil that statement down to numbers: at least 50 000 compounds per day. The goal must also include the wringing of as much information as possible out of a single library synthesis. Due to the desire to use a library a number of times to maximize the use of each synthesized compound, and synthesis strategies which restrict the amount of compound synthesized per round, it is also necessary that HTS systems be able to operate with small amounts of compound. In order to conserve synthetic reagents and the compounds themselves, and achieve biologically relevant concentrations, modern HTS systems must be able to run several assays with picomole quantities of compound in volumes of less than a microliter. As stated earlier, miniaturization is a prerequisite of an efficient HTS system.

Libraries used for lead finding are, by nature, diverse, but screening places the added demand that this diversity be describable: we need to be able to register and store a variety of characteristics associated with each individual compound, as well as collate chemical and biological information. Not only is it necessary to keep track of each compound, the

results of screening must also be tracked for a variety of assays. In the future, efficient integrated synthesis and screening systems will be able to take the stream of compounds directly from synthesis to screening without the need for interim storage. Hits will be resynthesized in scaled-up quantities for downstream studies based on screening and synthesis information alone.

The assays themselves present a significant demand on HTS systems. Established biochemical assays must be readily adapted to miniaturized format. Additionally, assay throughput can be increased by multiplexing strategies: the rapid sequential or simultaneous detection of multiple parameters. This multiplies the amount of information which can be obtained from a single assay – and thereby the information that can contribute to the characterization of a single compound. This also means that the lines between primary and secondary screening will become increasingly blurred. A variety of read-out modes must be available in any HTS system, not only to enable assay multiplexing and thereby the rapid and efficient characterization of compounds, but also to allow for flexibility in assay design. This flexibility also contributes to the ease with which specific assays can be translated from standard manual formats to miniaturized automated formats. This translation of assay protocols is a challenge not to be underestimated: a fact to which many of those who are already involved in screening will most likely testify grudgingly. For this reason, those same scientists would also be glad to have a variety of read-out modes with which to test different assay strategies.

To sum up this list of demands: a screening system must be capable of achieving throughput of 50 000 compounds/day, in volumes of less than a microliter, possess multiple read-out modes, and master multiplexed assays with the results being tracked by a database capable of correlating synthesis information with assay results. It must also be cost-effective. So, you may ask: Is miniaturized uHTS worth all the trouble? The answer is yes, and many companies are active in the field of HTS. Most, however, have chosen to concentrate either on specific types of assays, such as cell-based assays, or to rely on a single detection technology. EVOTEC's goal is to meet all the requirements of HTS through the use of EVOscreen to achieve a degree of sensitivity and specificity through high spatial resolution in miniaturized formats, a multiplicity of read-out modes all operating in high-throughput mode, with an integrated database system capable of correlating the data required to obtain useful information.

## 14.2 Combinatorial Chemistry

### 14.2.1 Design and Production of Chemical Diversity

The initial design phase for combinatorial compound collections focuses on computer-generation of a virtual library and selecting a subset of compounds for chemical synthesis on the basis of specific characteristics, such as maximum diversity, desired lipophilicity, and lack of toxic and reactive functionality [13]. When combinatorial compounds are used for lead optimization, collections can be designed on the basis of a reference structure with the collection chosen to represent the desired degree of diversity. Software for the design of such combinatorial reactions is commercially available. An alternative ap-

proach to generate a diverse compound collection is selection from an existing database of compounds [12]. Similarity-based selection of test compounds from an existing compound collection increases the effectiveness of lead finding and lead optimization [14].

The construction of a fully automated laboratory where all manual steps – from chemical synthesis to product characterization – are managed by robots is no longer a futuristic dream. Currently, about 20 companies are specialized in the design of chemistry robots of different sizes and levels of automation [15, 16; Chapt. 19]. The robotized laboratory for chemical synthesis has to be connected with databases for building blocks, for reaction information, and for library information. The tools for data analysis should be accessible to the operator to design libraries and lead optimization in cooperation with specialists for screening, bioinformatics, database management, computational and medicinal chemistry, and for compound characterization.

### 14.2.2 Microreactor Systems

Today, most combinatorial and high-throughput synthesis strategies are calculated for 10 mg, 10–20 micromolar, yield. In the first generation of microreactors, single compound beads or spots on functionalized surfaces have been introduced for the synthesis of pico- to nanomolar amounts of product [17, 18]. As they come into wide use, microreactors will bring about a fundamental change in combinatorial synthesis and screening strategies.

Microreactors will be directly connected to microscale purification units and highly sensitive analytical devices [19]. Nanotiter plates are already being used for direct collection from microfluidic separation units, transferring samples directly to biological assay units. Such integrated labs-on-a-chip – combining microreactors, microscale purification and separation units, and highly sensitive analytical devices – are likely to make compound storage systems obsolete. Test compounds will move directly from synthesis to screening. Furthermore, high-throughput miniaturized synthesis will be designed for parallel synthesis, as well as for the initial scale-up process. To fulfil all of these requirements, all components of integrated microreactors, mixers, valves, and process-control sensors, must be stable to aggressive, highly reactive chemicals.

The benefits of microreactors are many: more ideal temperature exchange for gas phase reactions; they enable the reaction process to take place in an electric field; and the performance of complex catalytic reactions can be simplified [18]. These benefits will also provide access to novel structures. Additionally, the bottleneck caused by limited availability of unusual building blocks and scaffolds will be eased, or altogether eliminated, by microreactor operations and miniaturized screening formats. The advent of microreactor systems for chemical synthesis not only opens the door for direct integration of synthesis and screening, but also provides better access to novel compounds.

### 14.2.3 Procedures for the Synthesis of Combinatorial Compounds

Compound collections and libraries generated by the combination of different chemical reactions and building blocks are valuable tools in the search for novel ligands. Such

combinatorial synthesis strategies result in residues at either defined or degenerate positions on defined chemical scaffolds. Compound mixtures can be designed for screening either in solution [22], or with compounds immobilized on solid supports, such as polypropylene pins [23] or resin beads [24]. Pharmaceutical companies have started to establish a database for published and proprietary procedures for solid-phase organic synthesis which can be downloaded to their automated laboratory for chemical synthesis.

Pioneering publications by Leznoff [20] and Furka et al. [21] provided impetus for the field of solid-phase combinatorial synthesis. The 'split-and-pool' method was developed to generate resin beads that are structurally homogeneous, and is characterized by coupling different building blocks in separate vessels, batch deprotection with pooling of all resins, with redistribution prior to the next coupling. Polymers offering a hydrophilic surface are especially preferred for testing polymer-bound compounds with soluble receptors [24].

For screening, compounds are removed from the solid support at the point of attachment by cleaving a chemical linker. A variety of chemical linkers have been described, all with the common trait that they enable release of the bead surface compounds under mild conditions. For identification of specific compounds, several methods for on- or off-bead analysis are available: sequencing of peptides on single beads [25], mass spectrometry of the compounds from one isolated bead [26], chemical tags for decoding by gas chromatography [27], and new tagging strategies, such as DNA-tags [28] or radio-tags [29].

Recently, EVOTEC has developed a system for the homogeneous analysis of bead-surface interactions. This system will enable compound screening to take place without the need to remove compounds from the surface of the bead. The system also enables direct recovery of single beads for further analysis. By enabling the rapid, on-bead analysis of thousands of beads, the EVOTEC system eliminates a critical bottleneck, namely the need for deconvolution of solid-phase diversity.

#### 14.2.4 Solution-Phase Combinatorial Chemistry

Solution-phase combinatorial chemistry overcomes the restrictions of solid-phase synthesis caused by the need to attach educts to, and release products from, the solid support. The adaptation of standard synthesis procedures to solid-phase chemistry is not necessary. On the other hand, the chemist must be careful to ensure that educts have reacted almost completely, and that excess reagents are removed after each reaction step: this represents a challenge for automation and process control. A core molecule as a template with several reactive groups has been used to generate xanthenene [30] or piperazine [31] libraries, for example. An iterative process for identification can be carried out by the deletion of one of the building blocks.

A hybrid between combinatorial synthesis on solid phase and in solution phase is the liquid-phase method characterized by the application of carrier polymers completely soluble in one solvent and insoluble in another solvent, for example functionalized polyethylene glycol [32, 33].

### 14.2.5 Limitations of Combinatorial Design Strategies

Bioassays are influenced by protein-reactive compounds [34] and tolerate only limited numbers of non-related molecules. False-positive and false-negative results from screening of compound mixtures have been reported. For example, octa- and hexapeptide mixtures significantly depress binding of the hormone neuropeptide Y to its receptor when measured in a standard competition assay [35].

Many pharmaceutical companies own sample collections of up to a million individual natural or synthetic substances. Combinatorial libraries will enlarge the collections to several million compounds. This explosion in the number of samples and the associated assay costs have forced most screening strategies to resort to a pooling strategy whereby up to thousands of samples are mixed in a single well for primary screening in traditional 96-well formats. This pooling strategy makes the results of screening difficult to interpret, but it is currently the best way to meet necessary throughput levels with conventional assay formats. This is the primary reason behind the drive to miniaturization and increased throughput. The screening systems of the future must not only be faster, they must provide more reliable data and be more economical.

## 14.3 Screening Technology

The following sections describe a miniaturized HTS system designed to address the need for increased throughput, efficiency and economy in light of the combined impact of genomics and combinatorial chemistry. This system, EVOscreen, developed by EVOTEC BioSystems, was designed with multiple fluorescence read-out technologies in a miniaturized format to allow for the broadest possible application to diverse targets, compounds and assay systems.

### 14.3.1 The Ideal Read-out Technology: Confocal Fluorescence

The use of fluorescent dyes has effectively replaced radioactivity as the tagging method of choice for biological assays. The ability to make full use of the fluorescent molecule and all of its properties should be the benchmark of any fluorescence-based assay technology. While most assay strategies make use of only a single fluorescent property, the broad applicability of a screening system requires that flexibility of assay design be extended to the read-out technology by including a variety of detection modes. While most scientists may associate the use of fluorescence in biological assay systems solely with the measurement of fluorescence intensity, the measurement of additional fluorescent properties such as lifetime, polarization, and quenching can yield a wealth of information from a single measurement. This ability to collect multiple data points per measurement not only provides an internal control, but also contributes to screening efficiency by enabling rapid multi-parameter evaluation of compound–target interactions.

Confocal microscopic optical systems form the foundation of the read-out technologies employed in the EVOscreen system. Using confocal optics, sub-microliter miniaturiza-

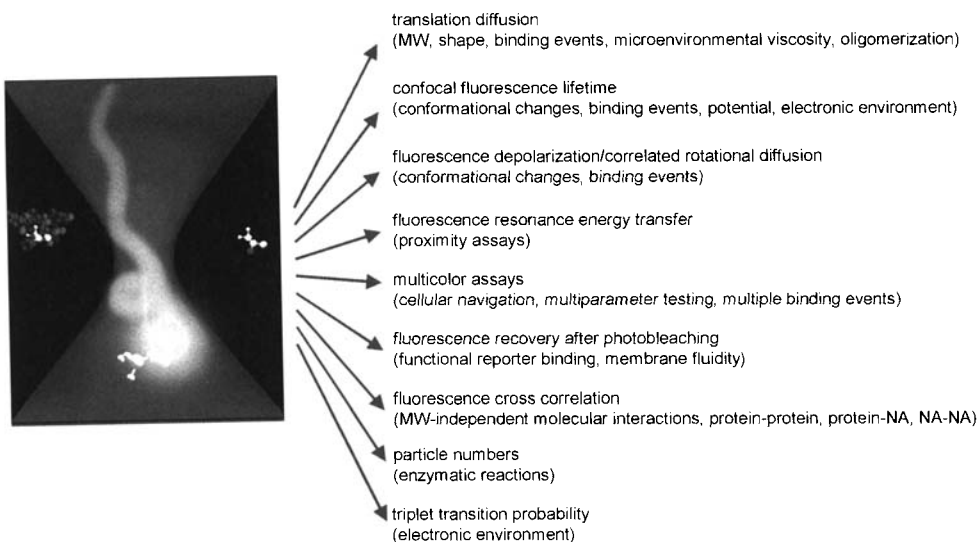
tion may be accomplished without any loss of signal quality. The highly focused confocal optics reduce the detection volume to one femtoliter. This means that the detection volume, roughly the size of a bacterial cell, is much smaller than eucaryotic cells or standard miniaturized assay volumes of 1  $\mu\text{l}$ . Whereas other methodologies often suffer from the increased contribution of surface effects and from drastically reduced signal-to-noise ratios when assay volumes are reduced, the confocal optics employed for all EVOscreen read-out technologies eliminates such concerns.

Furthermore, confocal detection technology enables the use of homogeneous assay methods, *i.e.*, those which eliminate washing or so-called ‘sandwich’ strategies. This is a great advantage in screening in that it allows biological systems to be evaluated in close to their natural *in vivo* state. Building on this confocal fluorescence detection platform, the EVOscreen system described in this report fulfils all of the demands made of a HTS system in the combinatorial age.

Confocal fluorescence detection allows biological parameters such as the following to be measured quantitatively:

- Molecular binding events: the study of the binding kinetics of molecular interactions, such as receptor/ligand, protein/protein, protein/nucleic acid, and nucleic acid/nucleic acid;
- Molecular processing: the evaluation of the progression and outcome of enzyme assays, such as aggregation, polymerization, degradation, or enzymatic modification of proteins or nucleic acids;

### Confocal Fluorescence Detection: Multifunctional Analysis



**Figure 14.3.** A survey of read-out modes for confocal fluorescence detection technologies and their applications to biological systems. NA = nucleic acid.



- Vesicular and cellular events: the study of cellular systems for the evaluation of transcriptional activity, signal transduction, membrane-bound or cytoplasmic reactions (potentially involving intracellular navigation), etc. using two-hybrid systems, for example.

Figure 14.3 lists a variety of different fluorescent parameters used as read-out modes for biological assays and their areas of application. Though this list is by no means exhaustive, it gives an indication of the wide variety of information accessible through fluorescent measurement methods.

## 14.3.2 Examples of Specific Fluorescence Detection Technologies

### 14.3.2.1 Fluorescence Correlation Spectroscopy (FCS)

FCS takes advantage of differences in the translational diffusion of large versus small molecules [36, 37]. Each molecule which diffuses through the illuminated confocal focus gives rise to bursts of fluorescent light quanta during the entire course of its journey, with each individual burst being registered. The length of each photon burst corresponds to the time the molecule spends in the confocal focus. The photons emitted in each burst are recorded in a time-resolved manner by a highly sensitive single-photon detection device. This detection method achieves single molecule sensitivity, but the fact that diffusion is a random process requires that the diffusion events for a minimum ensemble of molecules be averaged to achieve statistically reliable information. The detection of diffusion events enables a diffusion coefficient to be determined. This diffusion coefficient serves as a parameter to distinguish between different fluorescent species in solution, for example between free or bound ligand. In screening, the diffusion coefficient can be used to determine such factors as concentration or degree of binding. Hereby, confocal optics eliminate any interference from background signals and allow homogeneous assays to be performed. FCS measurements are conducted in a matter of seconds making the technology ideally suited for high-throughput applications.

### 14.3.2.2 Fluorescence Resonance Energy Transfer (FRET)

FRET is gaining increasing acceptance as a non-radioactive assay method for screening [38, 39]. FRET capitalizes on the transfer of energy from one fluorescent molecule to another. The distance between fluorescent molecules must be less than 40–50 Å [40], which is approximately the diameter of a protein molecule of 26 kDa in molecular weight. A major consideration in choosing an assay based on energy transfer is the change in distance that is induced upon ligand binding or enzyme turnover. For example, FRET can be employed for protease assays using peptide substrates labeled with two fluorescent dyes, whereby energy is transferred from one dye to the other. Upon digestion with a protease, the two fluorescent molecules are brought out of close proximity and the loss of fluorescence intensity is measured. Again, using confocal optics, homogeneous assays can be performed using this read-out technology.

### 14.3.2.3 Fluorescence Polarization (FP)

Since 96-well readers have become available, FP has gained importance for screening. The advantages of FP include the need for only one fluorescent label and the easy adaptability to homogeneous assay schemes [41, 42]. In FP, changes in the polarization of fluorescent molecules is measured. As a molecule rotates, the polarization of an attached fluorescent tag can be determined. When such a molecule binds to another, its rotational movement, and thereby the frequency in its changes in polarization, are restricted. Since the polarization shift taking place upon excitation of the fluorophore is proportional to the fraction of the total fluorophore in the bound state, receptor–ligand assays may need to be performed in the presence of excess receptor, a situation that is not always practical. Enzymes – primarily proteases, but also kinases and phosphatases – can also be assayed using FP. Compared to FRET, FP is more useful when the targets are large, since larger changes in rotational mobility can generally be observed.

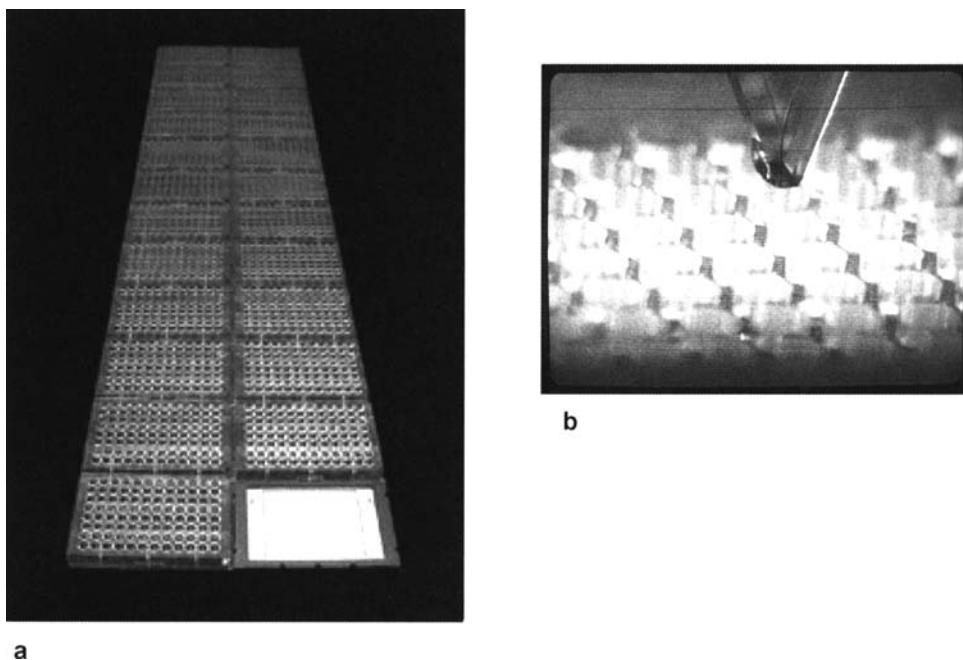
### 14.3.2.4 Fluorescence Imaging

The use of imaging technology has been enabled in drug discovery by the development of highly sensitive detection systems. Although this type of detection is based on the measurement of intensity alone, images can be acquired in quick succession, enabling rapid changes in cells to be observed. The development of a commercial high-speed, high-sensitivity imaging system [43] enabled limitations in the detection of calcium flux to be overcome (the calcium response is short-lived, < 1 s). This imaging system is able to collect parallel image data at speeds up to 10 frames per second, at sensitivities that allow detection of calcium transients in cell populations. High-sensitivity fluorescent imaging plate readers have been developed for cellular assays, particularly assays of membrane potential and intracellular calcium release [44]. The instrument uses a high-intensity laser source with narrow depth-of-field optics which helps to eliminate background signals. The combination of fluidics and imaging allows for high-throughput despite the transience of the response. Though fluorescence imaging is widely used, particularly for cellular assays, limitations such as non-specific adsorption effects and variations in signal due to small differences in filling levels restrict its applicability to solution-based assays.

## 14.3.3 EVOscreen™

EVOTEC has developed a unique platform integrating highly sensitive detection technology and high-precision liquid handling systems in a modular system which will be capable of 100 000 assays per day. The core elements of the modular system are:

- A high-performance confocal fluorescence detection unit usable in either a single channel or multichannel mode.
- Proprietary signal processing protocols based on FCS and related single molecule-based confocal fluorescence methodologies.



**Figure 14.4.** (a) A comparison of the reduction in space achieved by using miniaturized sample carriers. A total of 2000 samples may be carried by either the 23 microtiter plates seen in this picture, or by a single 2000-well nanocarrier (lower right). (b) The head of a nanodispenser situated over one of the wells of a 2000-well nanocarrier.

- A miniaturized, automated liquid handling system for nano- to low microliter volumes (including pipetting, dispensing and compound retrieval).
- A rapid, miniaturized multi-replica compound repository providing the link between traditional single-compound library formats and EVOscreen.
- A micro-separation device (HPLC) coupled to the detection system.
- A scanner and picker device for the analysis of combinatorial libraries and for functional genome analysis.

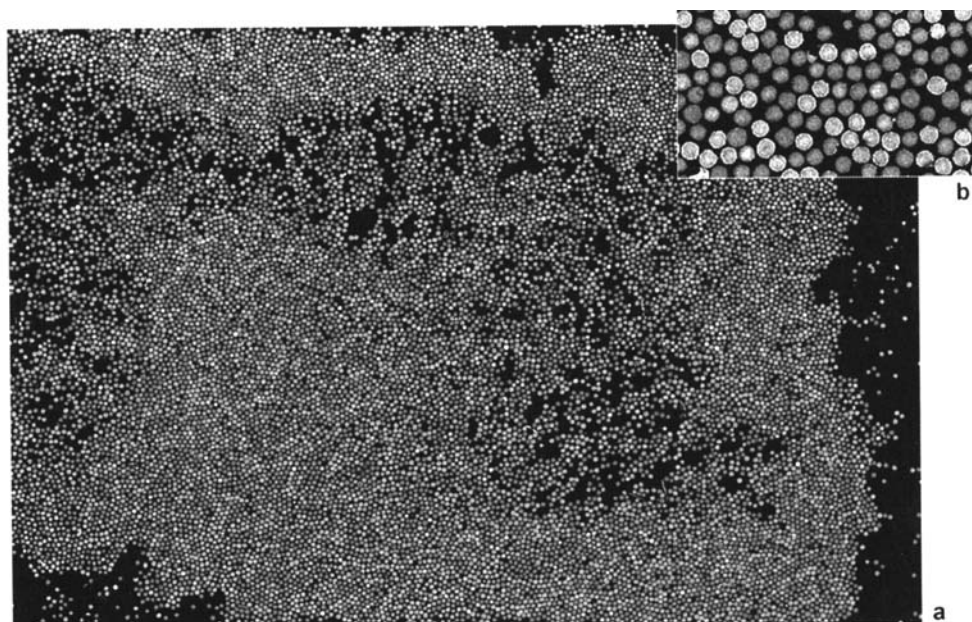
#### 14.3.3.1 Liquid Handling

The ability to miniaturize assays to the sub-microliter level is largely dependent on the precision of liquid-handling systems. Modern pipetting and dispensing systems enable components to be reproducibly added in volumes in the nanoliter range to thousands of samples per plate. The EVOscreen system includes a specially designed M2N unit which is able to reformat samples stored in any conventional format, such as 96-well plates, to

the high-density 2000-well Nanocarriers (Fig. 14.4) used for miniaturized assays. Also included is an automated storage and retrieval system providing a compound repository, as well as transport mechanisms for the many millions of compounds to be used in the screening runs. One specific feature of the EVOscreen liquid handling system is single-drop monitoring and on-line quality assurance.

#### 14.3.3.2 Scanner/Picker

EVOTEC has developed a unique assay device specifically designed to address the need for on-bead screening of combinatorial libraries. The system is designed to enable tens of thousands of beads to be screened per day. Using the same confocal optics and detection technologies described above, the bead scanner/picker enables homogeneous assays to be conducted without the need to cleave the chemical compound from the bead surface. Additionally, the system allows individual beads to be recovered for further analysis. The image of a scan of 12 000 TentaGel™ resin beads is shown in Figure 14.5.



**Figure 14.5.** (a) A scan of more than 12000 TentaGel™ beads using EVOTEC's Bead Scanner/Picker module in the presence of 1  $\mu$ M fluorescent protein conjugate. (b) A close-up of a region of the scan showing the bright fluorescent ring on the positive beads.

#### 14.3.4 Blurring the Lines between Primary, Secondary and Tertiary Screening

Drug screening has traditionally been divided into three phases: primary, secondary, and tertiary. As a result of the dramatic increase in targets and compounds to be screened, traditional screening is currently undergoing a re-evaluation. In addition to the move to miniaturization and increases in throughput, increased efficiency is also called for. One way to increase efficiency is to obtain more information at an earlier stage of drug screening.

Primary screening is largely the establishment of a structure–activity relationship (SAR). At this stage of screening, several strategies are typically employed to manage the biological testing of large chemical compound sets against multiple biological targets. A single compound per bioassay per well is the most direct – and this is precisely why throughput must be increased. The advantages are clear: no deconvolution is required and minimal potential for ‘masking’ of bioactivity exists. With one compound per well, the primary bioassay provides extensive SAR data, while negative bioassay data also provide additional information for subsequent lead optimization activities. For example, using FCS, not only can a SAR be determined between the compound and target, but the concentration can also be determined in a single measurement. This enables dose–response relationships to be evaluated immediately during screening. Miniaturization and automation are required to reduce the cost of this approach compared to the alternative strategy of compound pooling. While the pooling of 3–10 compounds per bioassay has been utilized to assay large compound sets quickly and efficiently, the primary disadvantage is the need for subsequent deconvolution of positive read-outs, and the potential for masking of one compound’s activity by others. By maintaining the one-compound-per-well configuration, the maximum amount of information can be obtained per compound.

In secondary screening, multifunctional testing for parameters such as selectivity, toxicity, and dose–response is conducted. By combining miniaturization, automation and multiple detection modes, this process can be moved forward in the drug screening process. New read-out technologies will enable several parameters to be measured simultaneously, meaning that assays – rather than compounds – will be ‘pooled’ in a single sample. Just as conventional strategies lead to the pooling of compounds to increase throughput, new detection technologies enable far more information to be obtained in a single measurement, leading to an increase in information and a shortening of screening time. This will make lead evaluation feasible in early stages of the drug discovery process.

Novel types of assay can also be designed for tertiary screening: the profiling of compounds to obtain valuable biological information, such as bioavailability and cellular uptake. To predict whether a molecule has the potential for in-vivo pharmacologic activity, it is necessary to determine whether the molecule persists in the body long enough at concentrations sufficient to exert the intended pharmacological effect. Such evaluations hinge on assessments of absorption, distribution, metabolism, excretion and pharmacokinetics (ADME/PK). If these evaluations are made in the early stages of drug discovery, resources can be appropriately channeled to those molecules that are likely to succeed in terms of in-vivo efficacy. Moreover, where a series of lead compounds is identified

through in-vitro pharmacological testing, AMDE/PK provides a means of rank-ordering the compounds. Such tertiary screens are typically carried out in test animals or in living cells. The new detection technologies described in this report are being used in functional cellular assays to explore signal transduction, transcriptional control and even cellular viability. The availability of human liver tissues, cell culture systems (e.g., Caco-2) as a model of the intestinal mucosa, recombinant drug-metabolizing enzymes, and cytochrome P<sub>450</sub> isoform-selective substrates and inhibitors have provided the tools for effective conduct of such in-vitro studies.

### 14.3.5 Using Primary Cells

Because of the difficulty of working with living cells, as well as the difficulty in automating functional assays, cellular assays have been relegated to smaller screens much later in the discovery process. The need to increase the screening speed and the type of information derived means that assays using living cells are now even being used in primary screens. Most assay strategies for use in living cells rely on genetically manipulated (otherwise known as recombinant) cells. However, such over-expressing recombinant cell lines frequently have altered physiological properties compared to primary or non-recombinant cells, particularly with regard to regulatory mechanisms of signal transduction. Thus they cannot be considered to be ideal for this purpose.

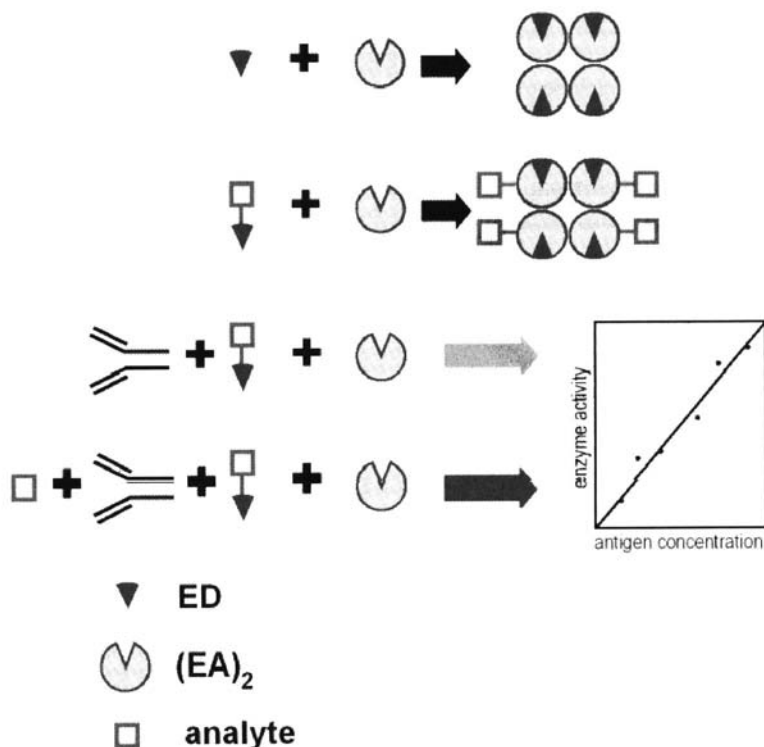
For cellular assays, miniaturized systems, such as EVOscreen, are capable of providing a superior solution: the combination of sensitivity and miniaturization enables primary cells to be employed. The restrictions resulting from the availability of primary cells are overcome by the very small number of cells needed for each assay (100–1000 cells/nano-well).

## 14.4 Practical Applications

### 14.4.1 Association Kinetics of Gene Product Fragments derived from *E. coli* $\beta$ -Galactosidase

Fluorescence correlation spectroscopy (FCS) was used to elucidate the association kinetics of lacZ fragments derived from *E. coli*  $\beta$ -galactosidase. This example demonstrates the power of FCS to provide mechanistic information by allowing the interaction to be observed at the molecular level. In general, it demonstrates how FCS may be used as a powerful tool for kinetic measurements in the life sciences.

The  $\beta$ -galactosidase analog consists of two subunits: a large dimeric polypeptide (200 kDa), denoted enzyme acceptor (EA)<sub>2</sub>, and a small polypeptide (20 kDa), denoted enzyme donor (ED). (EA)<sub>2</sub> and ED are enzymatically inactive but spontaneously associate to give enzymatically active tetramers. In CEDIA™ assays, [45, 46] the hapten or analyte is covalently linked to the ED, and an analyte-specific antibody is used to inhibit the assembly of the enzymatically active tetramers. Analyte in a patient's serum competes with the analyte in the analyte–ED conjugate for antibody, modulating the amount of active  $\beta$ -

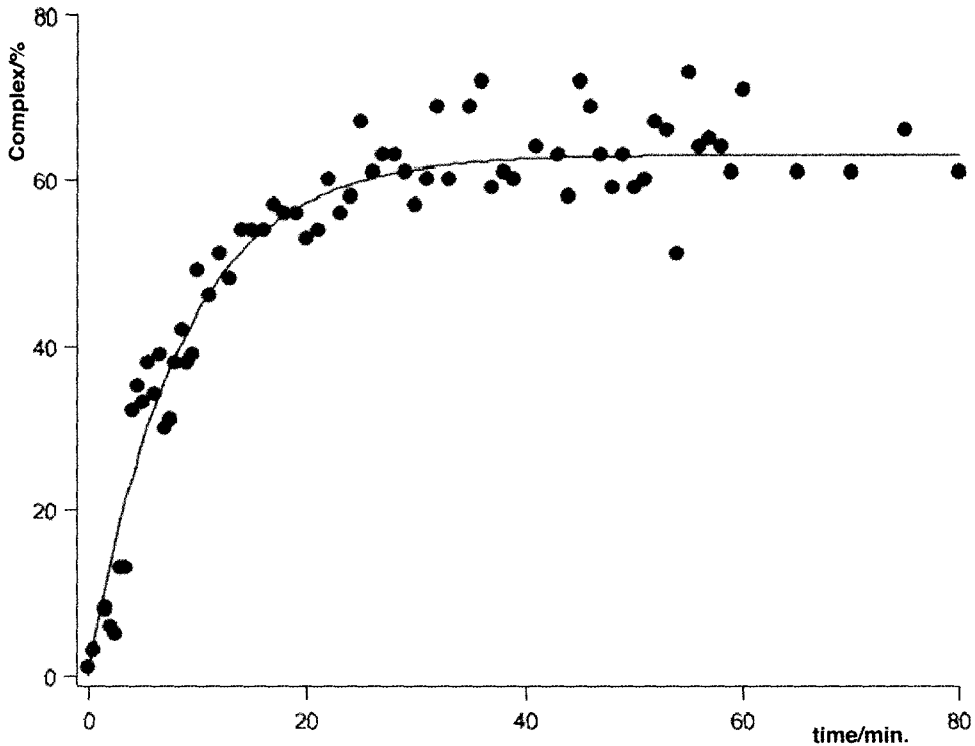


**Figure 14.6.** CEDIA™ assay. An analyte is attached to an ED molecule in such a way that the analyte–ED conjugate recombines spontaneously with (EA)<sub>2</sub> to yield active β-galactosidase [45]. The addition of analyte-specific antibodies inhibits spontaneous enzyme assembly [46]. Analyte in the patient’s serum modulates enzyme assembly so that the signal generated by substrate turn over is directly proportional to analyte concentration.

galactosidase formed. The signal generated by enzyme substrate is therefore directly proportional to the analyte concentration in the patient’s serum. (Fig. 14.6).

Characterization of the tetramer association process is required in order to investigate contaminants which may cause perturbation of the assay. Further optimization of CEDIA™-like assays requires insight into the reaction mechanism. It is important to know whether an enzymatically inactive (EA)<sub>2</sub> variant is not capable of binding ED, or has only been altered at its active site. So far, only indirect measurements of the association kinetics at very low ED concentrations have been performed by measuring the turnover of β-galactosidase substrate. Such measurements, however, are not capable of determining whether a lack of substrate turnover is rooted in a lack of subunit association or a loss of enzyme activity in a correctly assembled form of β-galactosidase. By using FCS, this constraint has been overcome.

A detailed study of association was performed in order to elucidate the reaction mechanism prior to enzymatic conversion of the substrate. The data give strong evidence for a fast pre-equilibrium state followed by a slower association into a stable complex. De-



**Figure 14.7.** Association kinetics of 5 nM ED and 9.9  $\mu\text{M}$   $(\text{EA})_2$ . The percentage of  $(\text{EA})_2/\text{ED}$  complex is plotted as a function of time. The kinetics of this reaction can be analyzed from FCS experiments. The data were fitted to a single exponential function.

tailed elucidation of the reaction mechanism has been described elsewhere [47]. Here, we will describe the set-up of the FCS assay and the data obtained to enable elucidation of the reaction mechanism.

The association of ED-Cy5 with native and heat-inactivated lyophilized dimeric EA was measured using FCS. It could be shown that the inactivated  $(\text{EA})_2$  variant was able to bind ED. The dynamics of the binding was similar to that of native  $(\text{EA})_2$  to ED-Cy5. In the case of the inactivated EA dimer, we conclude that the active catalytic center was converted into an inactive conformation, whereas the binding domain was not affected.

#### 14.4.1.1 Association Kinetics

Binding experiments were carried out with ED and  $(\text{EA})_2$  concentrations close to those used under assay conditions described for the CEDIA™ system. The time course of



the association was followed using time-resolved FCS. All binding experiments were performed under 'pseudo first-order' conditions with at least a 10-fold excess of  $(EA)_2$  in order to simplify the evaluation of the association kinetics. A typical time course of an association kinetics is shown in Figure 14.7. Even at the lowest ED concentration (1 nM), binding could be observed. The degree of binding at chemical equilibrium did not change with  $(EA)_2$  concentration, indicating a high binding constant for ED- $(EA)_2$  and that ED is saturated even at the lowest  $(EA)_2$  concentration. These results agree with those of Langley et al. [48], who report a binding constant of  $1.2 \times 10^9 \text{ M}^{-1}$  for the binding of the ED-analog ( $\alpha$ -complementing peptide CB2) and the  $(EA)_2$ -analog (M15  $\beta$ -galactosidase-residue). About 55 % of ED were bound by native  $(EA)_2$ . This means that a considerable percentage of ED is not able to be bound, at least under these experimental conditions. This could be due to conformational changes or isomers of ED which are unable to be bound by  $(EA)_2$ .

Association experiments were carried out using 1, 5 and 10 nM ED respectively, and 9.9  $\mu\text{M}$   $(EA)_2$  (Table 14.1). The time constants for the kinetics did not differ significantly in the three cases, which is characteristic of a bimolecular association mechanism, with the kinetics being determined by the species in over-abundance ( $(EA)_2$  in this study). The corresponding time constant was calculated to be  $8.8 \pm 1.7 \text{ min}$  in all cases. Several association experiments carried out on different days proved the reproducibility of time constants (Table 14.1).

In other experiments, the concentration of ED was kept constant at 5 and 10 nM, respectively, whereas the concentration of  $(EA)_2$  was varied from 0.27 to 9.9  $\mu\text{M}$ . The higher the  $(EA)_2$  concentration, the faster the association reaction. The reciprocal of  $\tau$  for the association kinetics depends on  $(EA)_2$  concentration according to a simple bimolecular association mechanism:

$$\frac{[\text{ED} - (\text{EA})_2]}{[(\text{ED})_0]} = 1 - e^{\frac{(-t)}{\tau}}$$

The slope of the linear regression corresponds to the association rate constant with the intercept corresponding to the dissociation rate.

**Table 14.1.** Time constants of ED- $(EA)_2$  association kinetics with 9.9  $\mu\text{M}$   $(EA)_2$  and 1, 5 and 10 nM ED, respectively.

[ED] (nM)	Time constant (t) (min)/min.
1	8.2
1	7.0
5	8.2
5	12
5	7.9
10	9.9
	$8.8 \pm 1.7$

### 14.4.2 Comparison of Native (EA)<sub>2</sub> and Inactivated (EA)<sub>2</sub>

It is known that inactivated (EA)<sub>2</sub>, in contrast to the native form, has no enzymatic activity. Until now, it was unclear whether this was due to altered binding characteristics or a direct alteration of enzymatic activity. The mean binding degree – calculated from FCS measurements – for inactivated (EA)<sub>2</sub> was 46% and that for native (EA)<sub>2</sub> 50%, respectively. The time constants of association kinetics with inactivated (EA)<sub>2</sub> were slightly larger than with native (EA)<sub>2</sub>. The deduced association rates for experiments with native (EA)<sub>2</sub> and ED was 207 M<sup>-1</sup> s<sup>-1</sup>, whereas the association constant for inactivated (EA)<sub>2</sub> and ED was slightly smaller at about 140 M<sup>-1</sup> s<sup>-1</sup>. Since the FCS measurements revealed similar binding characteristics for both the native and the inactivated forms of (EA)<sub>2</sub>, binding seems not to be significantly affected.

### 14.4.3 FCS Adaptation

The adaptation of the CEDIA™ assay to an FCS-compatible assay is not problematic. Moreover, because FCS-based assays are homogeneous in nature, no time-consuming purification steps or quality controls with respect to enzymatic activity are necessary. FCS enables kinetics to be measured under native conditions in the nM range, and is a rapid and reliable method for the on-line analysis of kinetic reactions. The FCS technology also allows for rapid measurements (in the range of seconds) of enzyme kinetics in homogeneous solutions without interfering with chemical equilibrium. The formation of the β-galactosidase tetramers has been measured for both the inactivated and the active form of (EA)<sub>2</sub>. The binding of the two reaction partners, ED and (EA)<sub>2</sub>, is the first step of this complex reaction cascade. By using FCS, we could show that the binding characteristics of the enzymatically inactive form were not influenced. The association reaction could be easily followed, and comparable degrees of binding were measured in all experiments.

Comparison of native and inactivated (EA)<sub>2</sub> revealed similar association kinetics for both forms. Hence, it can be concluded that inactivation of the protein does not affect the binding domain, but rather the catalytic site. Therefore, in order to optimize (EA)<sub>2</sub>, the conformation of the catalytic site must be stabilized rather than the binding domain.

The binding of two or more molecules to one another is a property that is used universally to ascertain information about biological events. FCS provides direct access to binding events in real time – and also provides the means to immediately determine an array of pertinent biochemical information. Compared to conventional CEDIA™ assay systems, an FCS-based CEDIA™-like assay will be more robust and not dependent on full enzymatic activity.

## 14.5 Summary

As stated earlier, the nature of screening for new drugs is undergoing a fundamental change. The exponential increase in the number of compounds available through the application of combinatorial strategies coupled with the increases in therapeutic targets

made possible through genomics means that screening must reach a new level of efficiency. The need to simultaneously shorten the drug discovery process and allow millions of new compounds to be tested against hundreds of new targets has led to a complete re-evaluation of screening technology. Miniaturization and automation cannot be avoided if screening is to remain economical.

A maximum of information must be gleaned from each assay. This means multiple read-out modes must be used to assess a wide variety of parameters in a single sample. As *in vitro* assays lend themselves to be miniaturized and integrated in a high throughput screening format, the new compounds can be directly tested for biological significance – and data obtained to help design the next iteration of lead optimization.

## 14.6 Acknowledgements

The authors would like to thank Dr. Eloisa Lopez-Calle for reviewing an early form of the manuscript, and for helpful discussions.

## 14.7 References

- [1] M. A. Gallop, R. W. Barrett, W. J. Dower, S. P. A. Fodor, E. M. Gordon, *J. Med. Chem.* 37 (1994) 1233–1251.
- [2] J. S. Früchtel, G. Jung, *Angew. Chem. Int. E. Engl.* 29 (1996) 17–42.
- [3] J. A. Ellmann, *Acc. Chem. Res.* 29 (1996) 132–143.
- [4] L.J. Beeley, D. M. Duckworth, *Drug Discovery Today*, 1(1996) 474–480.
- [5] A. Persidis, *Nat. Biotechnol.*, 16 (1998) 393–394.
- [6] A. W. Lloyd *Drug Discovery Today*, 2 (1997) 397–398.
- [7] G. L. Verdine, *Nature* 384 Suppl. (1996) 11–13.
- [8] S. Borman, *C&EN*, April 6, 1998.
- [9] S. P. A. Fodor, J. L. Read, M. C. Pirrung, L. Stryer, A. T. Lu, D. Solas, *Science*, 251 (1991) 767–773
- [10] E. M. Gordon, M. A. Gallop, D. V. Patel, *Acc. Chem. Res.* 29 (1996) 144–154.
- [11] D. Madden, V. Krchnak, M. Lebl, *Persp. Drug Discovery Des.* 2 (1995) 269–285.
- [12] P. Willett, *A Practical Guide to Combinatorial Chemistry* (Eds.: A. Czarnik, S. H. DeWitt) ACS Books, Washington DC (1997) pp 17–48.
- [13] E. J. Martin, J. M. Blaney, M. A. Siani, D. C. Spellmeyer, A. K. Wong, W. H. Moos, *J. Med. Chem.* 38 (1995) 1431–1436.
- [14] T. Pötter, H. Matter, *J. Med. Chem.* 41 (1998) 478–488.
- [15] R. A. Rivero, M. N. Greco, B. E. Maryanoff, *A Practical Guide to Combinatorial Chemistry* (Eds.: A. Czarnik, S. H. DeWitt) ACS Books, Washington DC (1997) pp 281–307.
- [16] J. F. Cargill, M. Lebl, *Curr. Opin. Chem. Biol.* 1 (1997) 67–71.
- [17] G. Wießmeier, D. Hönicke, *J. Micromech. Microeng.* 6 (1996) 285–289.
- [18] W. Ehrfeld, *Microreaction Technologies: Proceedings of the First International Conference.* (1998) Springer Verlag, Berlin.
- [19] F. Garcia, J. D. Henion, *Anal. Chem.* 64 (1992) 985–990.
- [20] C. C. Leznoff, *Acc. Chem Res.* 11 (1978) 327–335.
- [21] A. Furka, F. Sebestyén, M. Asgedom, G. Dibo, *Abstr. 14<sup>th</sup> Congr. Biochem. Prague Czechoslovakia* (1988) Vol 5, p 47.
- [22] H. An, B. D. Haly, P. D. Cook, *J. Med. Chem.* 41(1998) 706–716.
- [23] H. M. Geysen, T. J. Mason, *Bioorg. Med. Chem. Lett.* 3 (1993) 397–404.
- [24] J. K. Chen, S. L. Schreiber, *Angew. Chem. Int. Ed. Engl.* 34 (1995) 953–969.
- [25] K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, R. J. Knapp, *Nature* 364 (1991) 82–84.

- [26] C. Brummel, I. N. W. Lee, Y. Zhou, S. J. Benkovic, N. Winograd, *Science* 264 (1994) 399–402.
- [27] M. H. J. Ohlmeyer, R. N. Swanson, L. W. Dillard, J. C. Reader, G. Asouline, R. Kobayashi, M. Wigler, W. C. Still, *Proc. Natl. Acad. Sci. USA* 90 (1993) 10922–10926.
- [28] S. Brenner, R. A. Lerner, *Proc. Natl. Acad. Sci. USA* 82 (1992) 5381–5383.
- [29] K. C. Nicolaou, X.-Y. Xiao, Z. Parandoosh, A. Senyei, M. P. Nova *Angew. Chem. Int. E. Engl.* 34 (1995) 2289–2292.
- [30] G. W. Shipps, K. E. Pryor, J. Xiang, D. A. Skyler, E. H. Davidson, J. Rebek, *Proc. Natl. Acad. Sci. USA* 91 (1994) 11833–11838.
- [31] C. D. Garr, J. R. Peterson, L. Schultz, A. R. Oliver, T. L. Underliner, R. D. Cramer, A. M. Ferguson, M. S. Lawless, D. E. Patterson, *J. Biomolec. Screening* 1 (1996) 179–186.
- [32] M. Mutter, E. Bayer, *Angew. Chem. Int. E. Engl.* 13 (1974) 149–152.
- [33] D.J. Gravert, K.D. Janda, Organic synthesis on soluble polymer supports: Liquid-phase methodologies. *Chem. Rev.* 97 (1997) 489–510.
- [34] G. M. Rishton, *Drug Discovery Today* 2 (1997) 382–384.
- [35] K. H. Wiesmüller, A. G. Beck-Sickinger, H. G. Ihlenfeld, H. Wieland, K. Udaka, P. Walden, G. Jung *Peptides, Proc. 13<sup>th</sup> Amer. Peptide Symposium* (Eds.: P. T. P. Kaumaya, R. S. Hodges) Escom, Leiden (1996) pp 998–1000.
- [36] M. Eigen, R. Rigler, *Proc. Natl. Acad. Sci. USA* 91 (1994) 5740–5747.
- [37] R. Rigler *J. of Biotechnology* 41 (1995) 177–186.
- [38] T. T. Yang, S. R. Kain, P. Kitts, A. Kondepudi, M. M. Yang, D. C. Youvan *Gene* 173 (1996) 13–17.
- [39] P.R. Selvin *Methods in Enzymology* 246 (1995) 300–334.
- [40] C. R. Cantor, P. R. Schimmel, *Biophysical Chemistry*, San Francisco, Freeman WH, 1980, 451–454
- [41] W. J. Chekovich, R.E. Bolger, T. Burke *Nature* 375 (1995) 254–256.
- [42] D. M. Jameson, W. H. Sawyer *Methods in Enzymology* 246 (1995) 283–300.
- [43] Akong M, Siegel R, Vasserman E, Row B, Karlton D, McNeil J, Varney M, Stauderman K, Velicelebi G. *Soc. Neurosci. Abstr.* 21 (1995) 577
- [44] K. Schroeder, B. Neagle, *J. Biomol. Screening* 1 (1996) 75–80.
- [45] D. R. Henderson, S. B. Friedman, J. D. Harris, W. B. Manning and M. A. Zoccoli *Clin. Chem.* 32/9 (1986) 1637–1641.
- [46] W. A. Coty, R. Loor, M. J. Powell and P. L. Khanna *J. Clin. Immunoassay* 17 (1994) 144–150.
- [47] F.-J. Meyer-Almes, K. Wyzgol, M.J. Powell *J. Biophys. Chem.* (in press).
- [48] K. E. Langley, A. V. Fowler, I. Zabin *J. Biol. Chem.* 250 (1975) 2587–2592.

This Page Intentionally Left Blank

# 15 High-Throughput Experimentation in Catalysis

*Ferdi Schüth, Christian Hoffmann, Anke Wolf, Stephan Schunk,  
Wolfram Stichert and Armin Brenner*

## 15.1 Introduction

Combinatorial approaches have been well known in the pharmaceutical industry for many years. The first report that similar approaches might be useful also for inorganic materials was made in a review by Gordon et al. [1], and the first reports about highly parallelized synthesis and characterization of inorganic materials appeared a few years later. As various elements in the case of inorganic materials differ from combinatorial chemistry in pharmaceutical research, it is more appropriate to use the term 'high-throughput experimentation' (HTE) to describe related developments in materials science.

The first publications in this field focused on the identification of novel magnetoresistive materials [2], phosphorescent materials [3] and superconductors [4]. These areas of application have in common, that the characterization of the relevant physical property is relatively straightforward, albeit possibly very sophisticated, if parallelized to a high degree. Following these seminal papers, several groups devoted much effort to the development of high-throughput techniques in other areas. Certainly, one of the most promising application areas is catalysis, as catalysts are used in about 90% of all modern synthesis processes in the chemical industry, and contribute substantially to the efficiency and thus to the profitability of these processes. Improvements in catalysts often translate directly into substantial savings, and hence tremendous efforts are invested world-wide in industrial and academic laboratories, both to improve existing catalysts and to develop novel ones. As this is a slow process, HTE clearly has a large potential in catalysis research. The following discussion will be focused on HTE approaches in heterogeneous catalysis, since: (i) this is much the bigger part in industrial catalysis; and (ii) homogeneous catalysis is strongly related to combinatorial chemistry in the drug industry. In both cases, well-defined molecular units are synthesized and the properties are screened in the liquid phase. Although reaction conditions in homogeneous catalysis might be more demanding, synthesis and analysis seems to be much more straightforward compared to heterogeneous catalysis, as several successful examples already show [5, 6].

## 15.2 General Considerations

### 15.2.1 Motivation for and Problems of HTE Approaches in Catalysis

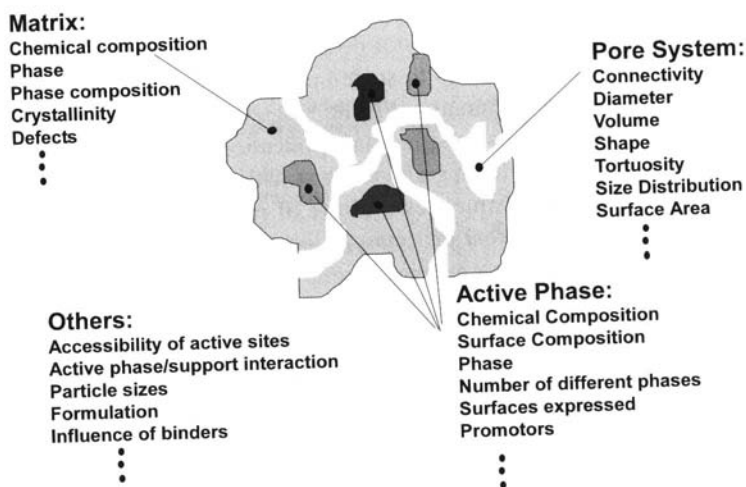
The motivation for the use of HTE approaches in catalysis research is very strong. The development of novel catalysts is often still very much driven by trial and error. Although a very well-developed set of tools exists for the optimization and incremental improvements of catalysts, completely new formulations or a totally novel application of a known formulation are found rather by serendipity than by rational design. High-throughput approaches will most probably not substitute the conventional process of catalyst development – at least not in the near future – but rather be a first, additional stage in the catalyst discovery process. High-throughput methods at this stage have not been developed to give the same information as extensive catalyst testing in the conventional mode. Therefore, HTE techniques are suitable to discover fundamentally new active compounds, with subsequent conventional development.

In addition, catalysts are often multicomponent systems. Partial oxidation catalysts for propene oxidation or ammoxidation can contain five or more elements as major constituents or promoters. The ammonia synthesis catalyst ICI 35-4 contains  $\text{Fe}_3\text{O}_4$ , 0.8 %  $\text{K}_2\text{O}$ , 2.0 %  $\text{CaO}$ , 0.3 %  $\text{MgO}$ , 2.5 %  $\text{Al}_2\text{O}_3$ , 0.4 %  $\text{SiO}_2$ , and traces of  $\text{TiO}_2$ ,  $\text{ZrO}_2$ ,  $\text{V}_2\text{O}_5$  [7]. To develop such a composition in the conventional manner is extremely difficult, and in fact, Mittasch and his co-workers at BASF in the early decades of this century needed approximately 20000 experiments to develop a catalyst for ammonia synthesis which has been used almost unchanged until today. In principle, HTE is certainly suitable to facilitate such a discovery process. In fact, there is nothing fundamentally new in this approach to catalyst development, except that nowadays we have the tools available to automate and parallelize the process used by Mittasch et al. at the beginning of this century [8].

Last, but not least, there is a strong economic driving force to apply HTE techniques in catalysis. There are several processes in the chemical industry for which no useful catalyst exists at all. Examples for such processes are the direct oxidation of propene to propene oxide, the synthesis of alpha olefins from alkanes, or direct activation processes for methane at moderate temperatures. These processes would be operated at a scale of several hundred thousand tons per year with a suitable catalyst, and the possible annual revenues involved are in the multimillion dollar range. For such processes, the financial investment necessary for a HTE catalyst discovery program is worth the risk. About 20 reactions can be identified which fall into this category, and with internal knowledge of industrial companies with regard to the raw material situation, this figure would likely increase to about 100 attractive processes.

However, although the possible benefits discussed above exist, so too do strong obstacles to the use of HTE techniques in catalysis. For the above-mentioned processes, it is certainly also possible that HTE approaches will not result in the discovery of a totally new catalyst, as world-wide research efforts of several hundred person-years have already been invested. Thus, the risk of failure is considerable.

There are also many technical problems in applying HTE approaches to catalyst development. A catalyst, especially one for a heterogeneously catalyzed reaction, is an extremely complex system (Fig. 15.1). Looking only at the active phase, the nature of this is



**Figure 15.1.** Schematic representation of the components of a heterogeneous catalyst and the parameters affecting the performance.

determined not only by the chemical composition, but also by factors such as the surface composition, which might differ substantially from the bulk composition, the crystallographic phase, the type of surface expressed, the size of particles, and many others. In addition, a catalyst does not only consist of the active phase, but also of a support, characterized by the matrix and the pore system, and of other constituents, like binders. Moreover, the different constituents of a catalyst might act together in various ways. The support and the active phase in supported noble metal catalysts often do not interact at all, but in other cases can strongly interact, as, for instance, in the case of the strong metal support interaction (SMSI). In contrast, a molecular compound which is used as a drug is determined fully by the chemical composition and structure alone. If a sufficient quantity of this compound has been synthesized, it can usually be analyzed by the full range of modern analytical tools, and its identity established unambiguously. In addition, if the compound is prepared in a combinatorial experiment, the synthesis procedure can usually be backtracked by tagging techniques, and thus the identity of the compound can be determined. This is extremely difficult in case of a heterogeneous catalyst. The synthesis pathway to a particular solid can strongly affect the properties of the resulting solids. For instance, in the synthesis of zeolites, a different order in mixing of the reagents can lead to different crystallographic phases. For zirconia, which is useful as a catalyst in many respects, differences in pH during precipitation, or the times of aging of the hydroxide precursor, can affect phases and particle sizes which, in turn, strongly influence catalytic properties [9]. The properties of heterogeneous catalysts are therefore governed by many more parameters than those of molecular species, and the analysis of a heterogeneous catalyst therefore needs to be much more extensive.

This creates a substantial problem as the synthesis of a heterogeneous catalyst is not easily scaled-up. A procedure which leads to an active catalyst on the milligram scale is not necessarily feasible on the gram scale. This means that it might not be easy to produce



a sufficient amount of catalyst for the analysis of chemical composition, structure, porosity, distribution of elements, etc., and therefore it is not possible to feed the information from the analysis back into the synthesis process to scale-up the synthesis. Solving this problem will possibly involve a lot of development work and make it difficult to translate a discovery directly from the HTE process to an industrial application.

Another difficulty lies in the meaningful testing of a catalyst. Other than the textbook definition of a catalyst as an agent which influences the rate of reaction without being consumed or changed itself, a catalyst in a real process can – and usually will – change drastically. Before it operates at steady state, a heterogeneous catalyst often needs to be formed, which can take as long as weeks. On the other hand, a catalyst often deactivates or ages during the catalytic process, by which activity is lost. Aging can be a very slow process which occurs over several years, but a catalyst might also be deactivated in a matter of seconds, as is the case for the FCC catalyst, which needs subsequently to be regenerated. An HTE evaluation of a prospective catalyst needs to take these points into account. There are two different approaches for this: If the HTE process is so efficient, that hundreds of thousands of formulations can be synthesized and tested in a year, one might afford to neglect forming and deactivation. In contrast, a slower HTE procedure should take these factors into account, and therefore each catalyst in a library should be tested at different times during a run. In an HTE catalyst development project one would most probably use a two-stage approach, with first a very highly parallelized approach in which forming and deactivation is neglected, while changes of the catalyst during the tests must be analyzed in a second step of evaluation. In addition, the performance of a process does not depend on the catalyst alone. The process is a complex interplay between reaction conditions, reactor, feedstock and catalyst, and it is usually not possible to optimize one part alone. The HTE process will be able to match – to a limited degree – the catalyst, reaction conditions and feedstock, but adaptation of the reactor will not be possible. Optimization of a heterogeneous catalyst normally involves all four factors mentioned above, and is not complete without taking into account the reactor. Nowadays, optimization to increase conversion and selectivity by 1% in technical processes which already are running at a high level of optimization would be a not unexpected request of a HTE project.

A last important difference to combinatorial approaches in the pharmaceutical industry is the fact that reaction conditions in catalysis research are frequently much more demanding. While testing of drugs is usually carried out at room temperature in aqueous environments and at ambient pressure, a heterogeneously catalyzed process quite often is operated in the temperature range around 500 °C, at pressures between ambient and several hundred bars, and in corrosive environments. A test system has to be adapted to such conditions, which clearly can be quite difficult.

However, we firmly believe that the problems discussed above can in principle be solved, and the possible benefit from even one novel catalytic process for the production of base chemicals resulting from an HTE development would already justify the effort. It should be clearly stated, though, what one can expect and cannot expect from a HTE project in heterogeneous catalysis at present. The main points are summarized in Figure 15.2. From these points it becomes clear that the HTE discovery process does not substitute conventional catalyst development, but rather adds a stage before the conventional development to search possible active formulations on a broader basis.

What is to be expected?	What is not to be expected?
<ul style="list-style-type: none"> <li>● screening on a much broader basis compared to conventional development</li> <li>● finding of fundamentally new active compounds more probable</li> <li>● highly reproducible synthesis and testing of formulations</li> </ul>	<ul style="list-style-type: none"> <li>● in the first stage no optimization of catalysts</li> <li>● limited possibilities to match catalyst and reaction engineering</li> <li>● no scale-up of catalyst synthesis</li> <li>● no scaleup of process</li> </ul>

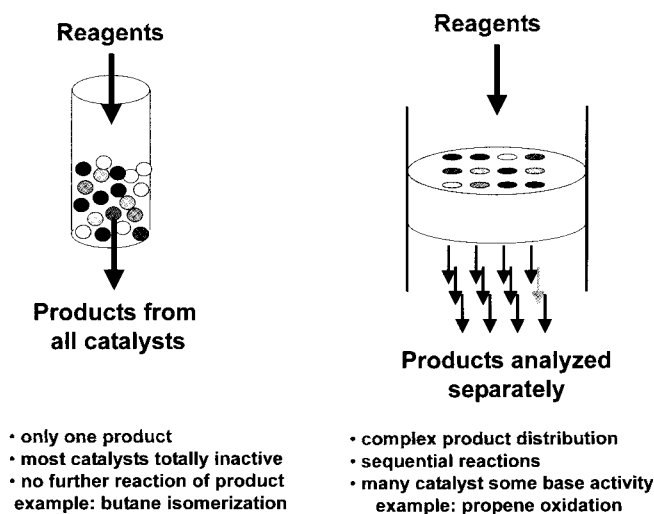
**Figure 15.2.** Results which can and cannot be expected from a HTE project in heterogeneous catalysis.

### 15.2.2 Strategies for Library Design and Testing

In HTE catalyst development, similar problems arise as in drug discovery by combinatorial techniques. One must ensure that libraries are sufficiently diverse in order to scan a broad range of parameters. However, as mentioned above, the parameter space for a catalyst formulation is much larger compared to that for molecular entities. Therefore, in addition to chemical diversity with respect to the constituting elements, diversity in the methods of preparation is also necessary. A catalyst prepared by a sputtering or CVD technique will have much different properties when compared to one synthesized by precipitation or impregnation. An advanced laboratory for HTE catalyst development therefore should have access to automated synthesis of materials by all means which are used nowadays in catalyst synthesis. In a first, very broad screening stage, however, it might be more useful to rely on a synthetic strategy which is very quick and highly parallelized, rather than cover the whole parameter space of synthesis pathways.

For choosing suitable chemical compositions of the libraries, two extreme strategies seem to be possible: One approach would focus fully on such elements in the periodic table which are known to be active for a certain class of reactions. For a partial oxidation reaction, one would use vanadium, molybdenum, bismuth, antimony, etc., to create the library. However, for the interesting catalytic processes in which an HTE approach would be most useful, many hundred person-years have already been spent in the development process, though without success. This means that many of the obvious combinations of elements have already been tested. A systematic scanning of the composition range might lead to active catalysts, but the chance to find a fundamentally new catalyst seems to be limited.

The second extreme strategy would be to neglect totally anything which is known in heterogeneous catalysis and randomly to scan combinations of elements in the periodic table. However, this approach does not seem to be very useful, since by this procedure one would, for instance, combine elements that form volatile compounds under reaction conditions which would then ruin the whole library and not be useful under industrial conditions. In addition, the price and availability of certain compounds might restrict the fully randomized approach.



**Figure 15.3.** Principal methods of approaching screening of the catalytic activity of a heterogeneous catalyst.

We consider that the best approach is one which combines the two extremes. Formulations should in most cases be based on elements known to be active in catalysis in a broad sense, but with the addition of non-obvious components in order to create non-obvious formulations. Only such a strategy would create the chemical diversity which increases the chances to discover something fundamentally new.

Similar approaches exist in the testing of catalysts as do in drug discovery. It would be possible to screen the different catalysts at one time (cocktail screening), and then to isolate the active ones by dividing the catalysts; alternatively, they could be screened one at a time (parallel screening) (Fig. 15.3). In most cases in heterogeneous catalysis – at least in gas-phase reactions – parallel screening will be the method of choice. A catalyst bed with different catalyst pellets over which the same feed is passed, is only useful in reactions where: (i) there is only one or very few products; (ii) most catalysts are totally inactive; and (iii) the product does not react further. If one of these conditions is not fulfilled, the test would give such complex results that further conclusions would be almost impossible. A partial oxidation reaction, for instance, usually has a complex product spectrum, as many different formulations have some basic activity, and many catalysts would also be good total oxidation catalysts. If different formulations were run simultaneously, the activity of the whole ensemble would cover one particularly good catalyst. A cocktail screening would, for instance, be suitable for testing butane isomerization catalysts, since this reaction usually is fairly selective – although some cracking usually is also observed – and needs very strong acidity which one finds in only a few catalysts. A carefully designed library could therefore be advantageously screened in one reactor. Another problem in gas-phase reactions is the fact that tagging of catalysts to track differently prepared particles would be difficult. Molecular tags, as used in drug discovery, usually do not have sufficient thermal stability. Inorganic tags are therefore necessary, and these are much more

difficult to design. For liquid-phase reactions, cocktail screening could be very useful, if spatially resolving detection techniques are used. Since mixing of the fluid is much slower in liquids than in gases, the contributions of a single catalyst pellet could be identified if the analytical technique is sufficiently fast. This was elegantly demonstrated by Taylor and Morken, who used infra-red (IR) thermography to detect the hottest – and therefore most active – catalyst beads in an acylation reaction in the liquid phase [10].

Most gas-phase reactions of interest, however, have a complex product distribution, can be sequential, and are to some extent catalyzed by many different catalysts; examples are the partial oxidation reactions. For this class of reactions, a parallel testing would be much better than cocktail screening, as in this case promising catalysts could be directly identified. The drawback of parallel screening, however, is the more complex reactor design and the fact that a parallel analysis or, at least, a relatively fast sequential analysis is necessary.

### 15.3 Approaches to Synthesis

Although at present the bottleneck in the HTE approach to catalysis is most probably that of testing to obtain meaningful catalytic data, high-throughput synthesis of potential catalyst materials is an important aspect of the whole process. In order to exploit fully the whole parameter space, one cannot rely on only one method of synthesis, but must acquire capabilities to produce possible catalyst materials by all routes used conventionally, but now adapted to the high-throughput mode, and probably also additional routes designed especially for the HTE process.

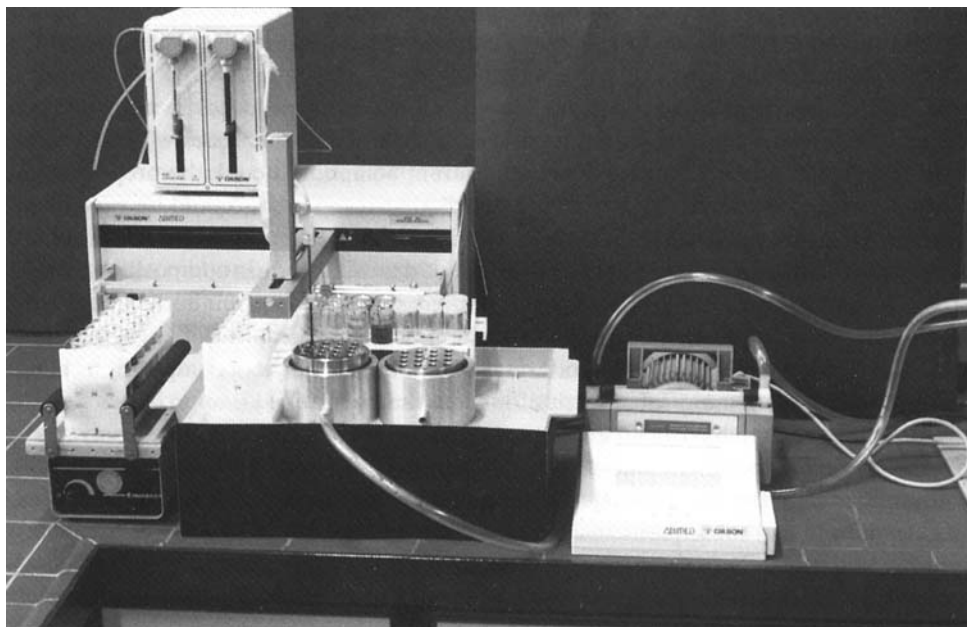
Several examples already exist in the literature for HTE approaches to inorganic synthesis. In one of the seminal papers, evaporation/sputtering techniques with adjustable masks are described, by which an array of different compounds could be prepared on a support material [2]. Using such techniques, different components could be evaporated simultaneously or subsequently, and converted to a desired compound in subsequent steps. In a variation of this approach, one does not deposit a certain composition in a desired location on a substrate, but uses a gradient sputtering technique from different sources. By correctly placing the sources of the precursors, compositional gradients are achieved on a substrate. Gradient sputtering to synthesize a range of compositions in one step has been used in many laboratories for some time before HTE approaches were labeled as such [11], and have been found useful to produce in a simple way a variety of materials. In a HTE experiment, one does not even necessarily need to know the exact composition of each of the locations on a substrate. A group at Lucent technologies used gradient sputtering and subsequent analysis of the dielectric constant by a mercury capillary probe measuring the capacitance of a specific area [12]. Only those regions which appeared to have promising properties were then analyzed to determine their exact composition.

Using sputtering or evaporation techniques will normally result in materials which are quite different from most conventional catalysts in that they expose only their geometric surface area and usually will not have any porosity. Therefore, synthesis approaches have been developed which allow the preparation of materials which are closer to a conven-

tional catalyst material. The easiest approach is simply to deliver precursor solutions manually from a pipette to the substrate. This approach has been chosen by Reddington et al. to create the smaller libraries used in their study [13]. Such a procedure, however, is no longer practical if the library size exceeds a certain threshold, or if the amounts of catalyst to be deposited become extremely small. Automated delivery techniques need to be developed to solve such problems, one example being the use of a modified ink-jet printer head, as described elsewhere [13, 14]. However, the latter library was not used to analyze the catalytic behavior of the members, but to discover novel luminescent materials.

More demanding are hydrothermal reactions, for example to synthesize zeolites, though these can also be extensively parallelized. The first publication concerned a semi-parallel zeolite synthesis procedure, where the gel was loaded into a relatively small Teflon vessel, of which 96 were arranged on one block [15]. As in most zeolite syntheses the hydrothermal treatment is the most time-consuming step, parallelizing this step immediately saves a considerable amount of time. The next important step was reported recently by W.F. Maier and co-workers, who used a spatially resolving X-ray diffractometer by AXS to analyze automatically a library of possible zeolites in which the individual members were present in only minute amounts, *i.e.*, on the milligram scale [16].

In our own laboratory, a multipurpose synthesis system was developed which allows all types of impregnation techniques to be carried out, as well as ion exchange and precipita-



**Figure 15.4.** The unit for catalyst synthesis based on a Gilson automated dispenser. From left to right, precipitation module with shaker; dispenser; preparation and filtration module made of stainless steel, in which post-treatment of the precipitate is possible. The dimensions and spacings of the wells in the preparation module are identical to those of the reaction module, so that simultaneous transfer of all catalysts is possible.

tion or co-precipitation. The system is based on a Gilson automated dispenser which has been modified to suit our needs (Fig. 15.4). In particular, the precipitation step is somewhat difficult to implement, as it normally involves relatively large amounts of liquids. Precipitation reactions are therefore carried out in relatively large vessels which are continuously shaken to prevent settling of the precipitate. The suspension is then pipetted onto a 16-well filtering unit which has the same dimensions as the catalytic reactor (see below) and washed after the mother liquor has been filtered off. The solids thus prepared can be post-treated after synthesis and then simultaneously filled into the reactor for the catalytic test. At present, this system is being used to synthesize high-activity, gold-based catalysts for oxidation reactions, in which the synthesis and precipitation conditions strongly affect the catalytic behavior.

## 15.4 Approaches to Testing

Catalyst testing is the most crucial problem in any HTE approach to catalysis, as in many reactions a certain amount of catalyst seems to be necessary to obtain meaningful catalytic data. Thus, it must always be established that any trends observed in a high-throughput experiment are reproduced in a conventional unit. The ultimate proof that a catalyst which appears promising in a HTE test must be obtained in the conventional manner, as the final performance of a catalyst in a process is governed by an intimate interplay between catalyst, reaction conditions and reactor – and all three factors cannot be simultaneously optimized in a HTE experiment.

In the pharmaceutical industry, the procedure in discovering new leads is a multistage process, in which first a highly unselective (though rapid) test is used to identify compounds which warrant closer inspection. The next screening stage is then more selective, albeit more time-consuming, and so on until a potential drug reaches the stage of clinical trials, or even commercialization. A similar approach seems appropriate in the development of novel catalysts by a HTE project. In the first stage a very broad – but probably relatively unselective – screening will be necessary. The second stage will resemble a conventional catalytic experiment, but it will be massively parallelized. The next stage will be test set-ups as in miniplants, or a catalytic microreactor which at present is the first step in industrial catalyst development. HTE would therefore typically place two additional steps before the conventional development process, but in doing so will scan an immensely broadened parameter space.

The very first method used to analyze catalytic activity in a parallel fashion employed IR thermography [17]. In that first publication, conventionally prepared catalyst pellets were placed on a supporting wafer that was located in a reaction chamber. The top of the reaction chamber was sealed with an IR-transparent window (sapphire), and a hydrogen/oxygen mixture was passed over the catalysts. Catalytic activity was detected via the heat of this highly exothermal reaction, which was analyzed using a thermographic imaging device. Using this technique, the noble metal-containing samples for which high activity was expected were indeed found to be the most active.

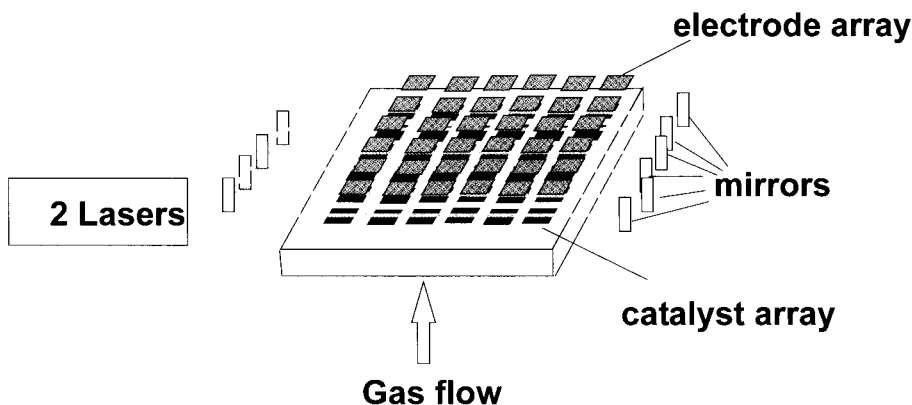
Similar thermographic techniques have been used subsequently in a more elaborate fashion. Taylor and Morken analyzed the heat of reaction released in an acylation reac-

tion [10]. The catalysts in this case were heterogenized homogeneous catalysts on polymer beads, the reaction being carried out in solution. One problem of thermography in solution is the fact, that the bulk of the solution screens the heat released at an individual catalyst bead. This problem was elegantly circumvented by using chloroform as the solvent. As the polymer beads are less dense than the chloroform, the beads float on the surface of the solution and the thermographic image thus represents the actual heat released by the reaction. Individual tests on the hottest beads revealed that highly active acylation catalysts were bound to these beads.

In heterogeneously catalyzed gas-phase reactions, one of the problems encountered in the first publication was the fact that only relatively large thermal signals could be detected. This problem was solved by the work of Holzwarth et al. [18], who used a background subtraction technique to reduce the detection limit to differences of about 0.1 K. With this set-up it was possible to analyze the activity of several metal-doped, amorphous, mixed metal oxides in total oxidation reactions of hydrocarbons.

Although the results of these studies are impressive, thermography suffers from one serious disadvantage in that it does not provide any selectivity information. This is not a problem for reactions where only one product is expected, as in CO-oxidation or the  $H_2/O_2$ -reaction. However, most of the commercially highly interesting reactions for which no satisfactory catalytic solution yet exists are partial oxidation reactions. In such reactions, the hottest catalyst identified in the thermographic experiment is probably the worst, as the most exothermic process is the total combustion of the hydrocarbon to carbon dioxide and water. Thermography is most likely a good technique for outline and preliminary screening to detect compositions which have any activity at all, and there is clearly a great need for more strongly discriminating assays in a more carefully controlled test.

Several of such tests have been developed, some of which are highly parallelized, and some of which use fast sequential analysis. A highly parallelized test has been developed for the discovery of novel methanol fuel cell catalysts [13]. In the anode reaction in a

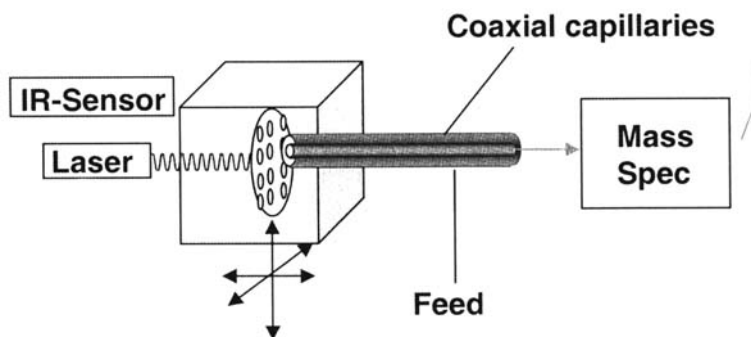


**Figure 15.5.** Schematic drawing of the screening unit using resonance-enhanced multiphoton ionization with detection of the formed ions by an electrode array. For an explanation, see the text.

methanol fuel cell, six protons are released for every methanol molecule converted. Reddington et al. used a probe which changes its luminescence if protonated. Changes in the luminescence behavior thus signify an active catalyst for the anode reaction in a methanol fuel cell, and by using this technology, the composition of a novel fuel cell catalyst may be discovered.

A parallelized analysis technique for a gas-phase reaction was developed recently by Senkan [19], who used laser-induced resonant multiphoton ionization (REMPI) to detect benzene formed by the dehydrogenation of cyclohexane. REMPI makes use of the selective ionization of molecules by first exciting a molecule by a first laser, and then irradiating with a second laser resonantly into the excited state to ionize the molecule (Fig. 15.5). The laser light is passed over the whole catalyst array by means of mirrors. The ionized molecules in the gas plumes after passing the catalyst materials are detected at electrodes mirroring the catalyst array. A current is measured only at those electrodes where benzene is produced at the catalyst placed upstream of the electrode. This technique will probably be useful in selected cases, where excited states of different possible products are well separated which, however, is not the case for many products of partial oxidation reactions. In addition, concern must be expressed about the electrodes being contaminated by coke or other high-boiling residues which might accumulate there with time, and change the detector response.

A selective and relatively quick – although not parallelized – method of analysis was developed by the Symyx team using mass spectrometry (Fig. 15.6) [20]. This technique is suited for small amounts of catalyst deposited on a flat substrate. The reagents flow towards a catalytically active spot via the outer of two coaxial capillaries, and the products are brought into a mass spectrometer by the inner capillary. Heating of the catalyst spot is carried out from the back via a laser, and the temperature is also monitored at the back of the substrate via an IR sensor. The whole substrate with the catalyst library can be moved in  $x$ -,  $y$ -, and  $z$ -directions to position the capillaries. This set-up has been used to analyze catalysts for the CO-oxidation [20] and for oxidative ethane coupling [21], and it was claimed that, when compared to an industrially optimized catalyst, an improvement had been achieved.

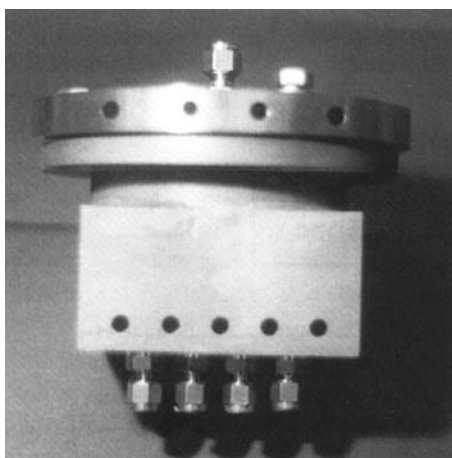


**Figure 15.6.** Schematic drawing of the screening unit using the scanning mass spectrometer approach. For an explanation, see the text.



For the second step of optimization, it is likely that multiple pass reactor technology will be used, where several wells filled with catalysts are attached to a common feed and a common analysis. This reduces costs, as the gas-handling system and analysis are often the most expensive parts of the equipment. More importantly, however, the evaluation times have been shortened substantially, as bringing a reactor up to reaction temperature, reaching the steady state, and cooling down again after the test is often very time-consuming. The analysis times themselves are comparatively short, so that sequential analysis is often not a major problem. If, therefore, a massive parallelization of the previously mentioned steps of heating, etc. is possible, the actual time for the analysis is of major concern only if a daily throughput of more than 100 catalysts is required. Gas chromatographic (GC) analysis has a typical cycle time of about 15 min, which means that in a system operating around the clock, 100 analyses can be made per day using only one gas chromatograph. If multiple analysis per reaction channel is desired to monitor the formation and deactivation of catalysts, then more than one gas chromatograph might be linked to the same multiple-pass reactor. In addition, if catalysts are screened for only one desired product, an attempt should be made to optimize the GC program such that this compound is well separated and is not influenced by the separation of other products. This could reduce cycle times substantially, assuming that isothermal analysis programs can be developed.

Although multiple-pass reactors have been known for a relatively long time [22], parallelization in such reactors was often not pronounced. In our laboratory, a 16-well multiple-pass reactor has been operating routinely for about six months (Fig. 15.7), while a 48-pass reactor is in the prototype stage. Reaction gases are passed simultaneously through all channels, but only the flow from one channel at a time is switched to the analytical instrument by an array of 16 three-way valves in the case of the 16-well reactor, and a specially designed sampling unit in case of the 48-well reactor.



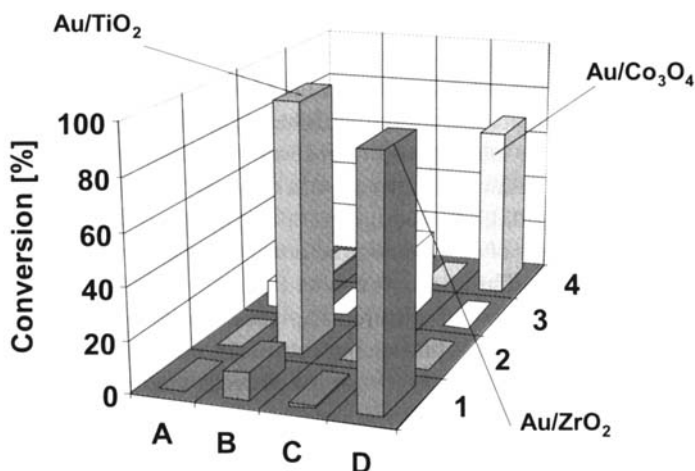
**Figure 15.7.** The 16-well parallel reactor, side view (actual size ~15 cm). The gas inlet is at the top; the outlets from the individual wells are at the bottom. The capillaries and three-way valves attached to the bottom in the assembled reactor are not visible here.

If multiple-pass reactors are developed, then several problems must be taken into account:

- The reactor must be isothermal over all wells. This can be attained by careful placement of heating elements and the use of a reactor material with high thermal conductivity. The reactor operating in our laboratory is equipped with several well-placed heating cartridges, and is made of brass, which allows for temperature deviations between individual wells of at most 1 K at a level of 200 °C without further insulation. If highly exothermic reactions are analyzed, the heat of reaction within a well at very high conversion might also affect neighboring wells. These effects can be taken into account either by careful data analysis, or by a second run with differently placed catalysts. However, when the amounts of catalyst used and the gas phase concentrations are selected carefully, the heat of reaction can be kept sufficiently low to avoid cross-talking between channels by thermal conduction.
- In order to allow comparison of catalysts, flow rates through all wells must either be identical, or be monitored individually. Different flow rates could result if the catalyst particles are packed at different densities, as this leads to different pressure drops over each catalyst bed. This can be avoided if capillaries are placed in the offstream of each well, which provides the major flow resistance. Differences in pressure drop over different catalyst beds are then relatively irrelevant, as they are small compared to the total pressure drop over each well.
- Back-mixing in the entrance of the reactor must be avoided. This is especially important if catalyst activities are low, and linear flow rates have thus also to be kept low in order to reach conversions which can be analyzed reliably. Slow linear flow rates will lead to cross-talking between wells via the gas phase, as diffusive back-mixing is faster than the linear flow through the wells. This problem was encountered in our first reactor design, but was solved by placing a diffusor plate with very narrow bores above each well into the entrance section of the reactor. Linear flow rates in these narrow holes were so high that diffusive back-mixing was reduced to zero.

Although these are the most important considerations in designing a parallel reactor module, another important point is the analysis. While in principle each analytical instrument is suitable and can be connected to the exit of such a multiple-pass reactor, one must ensure that the instrument works with relatively small amounts of sample, as well as low flow rates. For the model reaction under investigation in our reactor, the CO-oxidation, non dispersive IR is used. As CO<sub>2</sub> concentrations are relatively high under our conditions, the analysis chamber can be kept short, and purge times are therefore also short. Analysis times are around 4 minutes, this being determined mainly by the purge times of the tubes and the sample chamber. However, any analytical techniques, such as mass spectrometry, GC, etc., can in principle be used in connection with this set-up.

We used this reactor set-up to evaluate highly active room temperature CO-oxidation catalysts based on gold. These systems were discovered by Haruta and co-workers [23], but it is rather difficult to reproduce samples with similar performance based on the publications, as the activity of the final catalysts is determined by many parameters. The bottleneck in synthesizing active catalysts has been testing time, as the full catalytic test takes about a day, if temperatures other than only room temperature are to be investigat-



**Figure 15.8.** Conversion of CO to CO<sub>2</sub> at room temperature measured in the 16-well parallel reactor. Catalysts are arranged in a 4×4 matrix; 1% CO in air, gas hourly space velocity 25000 ml h<sup>-1</sup> geat.

ed. We therefore choose to test conventionally synthesized catalysts in the parallel reactor to increase throughput. Figure 15.8 illustrates one of the best results obtained for a set of 16 different catalysts which were evaluated simultaneously. Of these, three had a high room temperature activity, *i.e.*, Au/TiO<sub>2</sub>, Au/ZrO<sub>2</sub>, and Au/Co<sub>3</sub>O<sub>4</sub>, which had been deposited or co-precipitated and pretreated in a special way. The performance of the catalysts was checked using a conventional single-pass microflow reactor, and the conversions were found to be within 10–15% at each temperature compared to results obtained in the multipass reactor. The throughput in the evaluation of catalysts in this system was increased by about a factor of 15, while still obtaining comparable data as in conventional testing. The reaction examined here is certainly a very simple one, but the results show that testing times can be shortened considerably, even with a relatively uncomplicated and straightforward experimental set-up. More complex reactions, such as the methane partial oxidation, are currently under investigation in our laboratory.

## 15.5 Conclusions

The discussion above shows that, although several obstacles remain in the field of HTE approaches towards catalyst development, there seems to be a large potential for such techniques in the catalysis industry. Nonetheless, time will show whether results obtained on the scale of the HTE experiment can be validated on the conventional scale, or even in industrial processes. We are, however, optimistic that this will be the case, and that over the next five to ten years HTE will become a standard tool in catalyst development.

## 15.6 Acknowledgements

This research was funded by the BMBF under contract number FF/VF310/AF/ml. The authors also gratefully acknowledge many discussions concerning this topic with colleagues from industry and academia who, unfortunately, are so numerous that they cannot be mentioned individually here.

## 15.7 References

- [1] E.M. Gordon, R.W. Barrett, W.J. Dower, S.P.A. Fodor, M.A. Gallop, *J. Med. Chem.* **37**, 1386 (1994)
- [2] G. Briceno, H. Chang, X. Sun, P. G. Schultz, X. D. Xiang, *Science* **270**, 273, (1995)
- [3] X. D. Sun, K. A. Wang, Y. Yoo, W. G. Wallace-Freedman, C. Gao, X. D. Xiang, P. G. Schultz, *Adv. Mater.* **9**, No. **13**, 1046, (1997)
- [4] X. D. Xiang, X. Sun, G. Briceno, Y. Lou, K. A. Wang, H. Chang, W. G. Wallace-Freedman., S. W. Chen, P. G. Schultz, *Science* **268**, 1738, (1995)
- [5] M.T. Reetz, M.H. Becker, K.M. Kühling, A. Holzwarth, *Angew. Chem.* **110**, 2792, (1998)
- [6] K. Burgess, H. J. Lim, A. M. Porte, G. A. Sulowski, *Angew. Chem.* **108**, 192, (1996)
- [7] M.V. Twigg, "Catalyst handbook - second edition", Chapter 8, page 392,
- [8] R. Schlögl, *Angew. Chem* **110**, 2467, (1998)
- [9] W. Stichert, F. Schüth, *J. Catal.* **174**, 242, (1998)
- [10] S.J. Taylor, P.J. Morken, *Science* **280**, 267 (1998)
- [11] J. Hanak, *J. Mater. Sci.* vol. **5** (11), 964, (1970)
- [12] R.B. van Dover, L.F. Schneemeyer, R.M. Fleming, *Nature* **392**, 162 (1998)
- [13] E. Reddington, A. Sapienza, B. Gurau, R. Viswanathan, S. Sarangapani, E. Smotkin, T. Mallouk, *Science* **280**, 1735 (1998)
- [14] X. D. Sun, K. A. Wang, Y. Yoo, W. G. Wallace-Freedman, C. Gao, X. D. Xiang, P. G. Schultz, *Adv. Mater.* **9**, No. **13**, 1046, (1997)
- [15] D. E. Akporiaye, I. M. Dahl, A. Karlsson und R. Wendelbo, *Angew. Chem.* **110**, 629, (1998)
- [16] J. Klein, C.W. Lehmann, H.-W. Schmidt, W.F. Maier, *Angew. Chem.* **110**, 3557, (1998)
- [17] F. C. Moates, M. Somani, J. Annamalai, J. T. Richardson, D. Luss, R. C. Willson, *Ind. Eng. Chem. Res.* **35**, 4801, (1996)
- [18] A. Holzwarth, H.-W. Schmidt, W.F. Maier, *Angew. Chem.* **110**, 2788, (1998)
- [19] S.S. Senkan, *Nature* **394**, 350 (1998)
- [20] P. Cong, D. Giaquinta, S. Guan, E. McFarland, K. Self, H. Turner und W. H. Weinberg, Proceedings to *Process Miniaturization: 2<sup>nd</sup> International Conference on Microreaction Technology*, March 9–12 (1998), New Orleans, Louisiana, 118–123
- [21] H. Weinberg, Presentation at UMIST, September 1998
- [22] L. Kiezel, M. Rutkowski, Z. Tomasik, *Chem. Stosow. A*, **12** (3), 407, (1968)
- [23] M. Haruta, *Catal. Today* **36**, 153, (1997)

This Page Intentionally Left Blank

# 16 FT-Infrared Spectroscopy and IR-Microscopy On-Bead Analysis of Solid-Phase Synthesis

*Holger Bandel, Wolfgang Haap and Günther Jung*

## 16.1 Introduction

In recent years, combinatorial chemistry has proved to be one of the best approaches for acceleration in drug development [1–6]. Solid-phase organic synthesis (SPOS) [7–9] plays a most important role for production of combinatorial libraries.

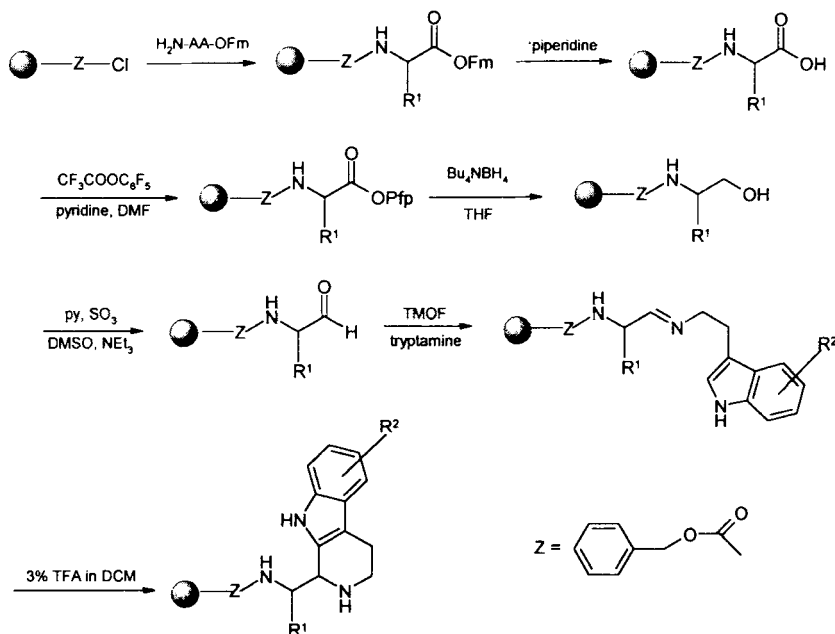
The combinatorial synthesis of organic compounds with low molecular weight is preferably carried out on a polymer support, as this allows automation to be carried out much more easily than reactions in the liquid phase. The main advantage of the SPOS is that time-consuming cleaning procedures of intermediate products are omitted. However, in order to apply common monitoring methods, such as TLC, HPLC or GC, samples of the polymer-bound products must be cleaved and analyzed in time-consuming procedures. Consequently, for the efficient optimization of reaction conditions, no chromatographic methods for continuous reaction monitoring are possible on the solid support. For this reason, non destructive spectroscopic methods as infrared (IR) and nuclear magnetic resonance (NMR) spectroscopy [10, 11] are applied, and allow the characterization of intermediate products in solid-phase synthesis. In particular, Fourier transform (FT)-IR spectroscopy allows rapid monitoring of the reaction via absorptions of characteristic functional groups and quantification of the resin-bound molecules.

Because of the simple sample preparation, FT-IR spectroscopy [12, 13] has become an increasingly popular method in SPOS for reaction monitoring. For IR spectra in the transmission mode [14, 15] or with ATR-[16], FT-Raman-[17, 18] and DRIFTS-[19] methods, only 1–2 mg of resin beads are required for routine measurements. With FT-IR microscopy [20–24], it is even possible to obtain spectra from single resin beads or spatially resolved spectra [25].

## 16.2 Analytical Methods Using FT-IR Spectroscopy

### 16.2.1 KBr Pellet Method

A common routine method for the control of solid-phase reactions is the use of FT-IR spectra from potassium bromide samples. For this, potassium bromide is mixed with resin beads and sintered under high pressure. The sample contains approximately 2% of the polymer support, and can be measured in any conventional IR-spectrometer.

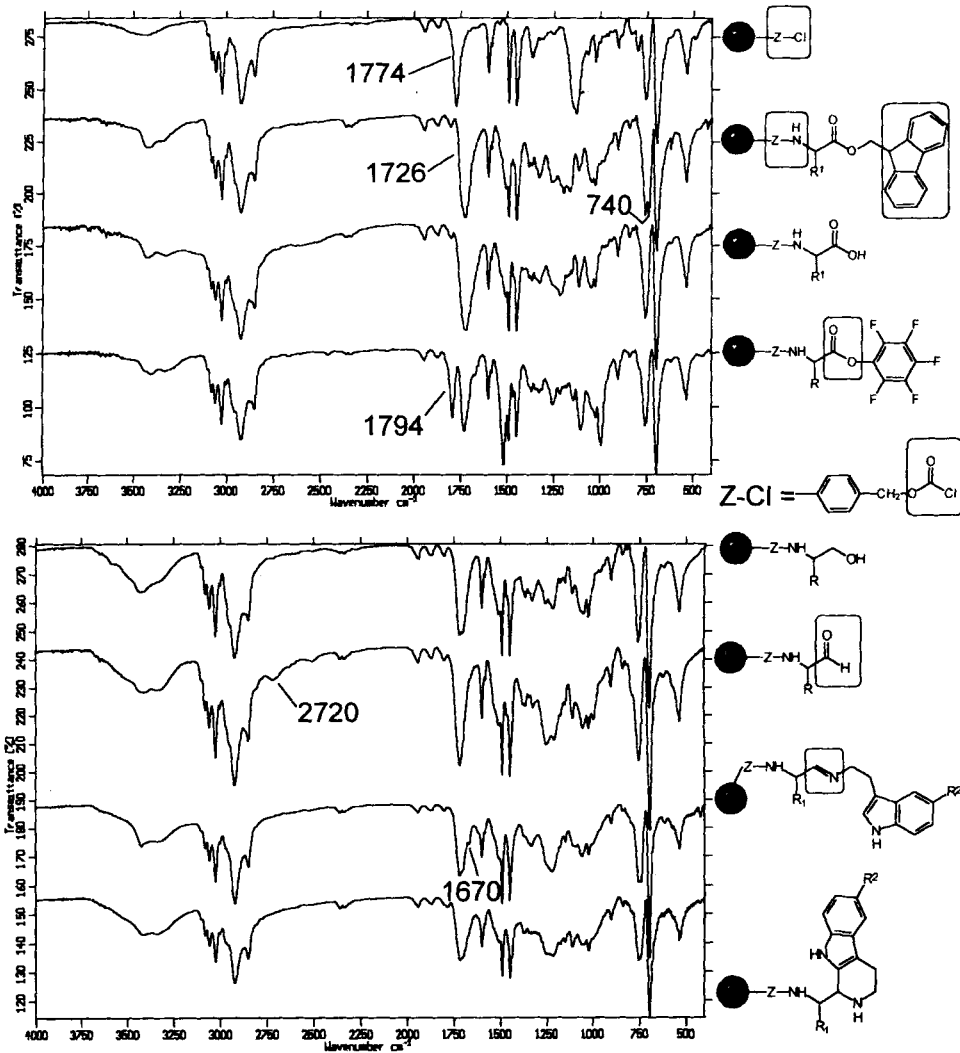


**Figure 16.1.** Solid-phase synthesis of 1,2,3,4-tetrahydro- $\beta$ -carbolines. (Reproduced from [26].)

The IR-spectroscopic control of a multi-step organic synthesis is illustrated on the example of the SPOS of an alkaloid compound. A polymer-bound benzoyloxycarbonyl anchor group on polystyrene resin is used as the starting material for the synthesis (Fig. 16.1). Each reaction step of the seven-step synthesis of 1,2,3,4-tetrahydro- $\beta$ -carbolines [26] can be verified by IR-spectroscopy (Fig. 16.2).

The PS-benzoyloxycarbonyl resin (Fig. 16.2) shows at  $1774\text{ cm}^{-1}$  the typical carbonyl absorption which disappears by loading the resin with amino acid fluorenylmethyl esters. This can be detected by a new carbonyl absorption at  $1726\text{ cm}^{-1}$  and the characteristic C—C absorption band of the fluorene at  $740\text{ cm}^{-1}$ , which disappears by treating the resin with piperidine. The successful reaction to the Pfp-ester can be detected by the additional carbonyl absorption band at  $1794\text{ cm}^{-1}$ . The following reduction step to the amino alcohol is proofed by the disappearance of this band. The oxidation to the aldehyde and the formation of the imine is shown by the characteristic absorptions of new functional groups. [ $\nu(\text{C—H, aldehyde}) = 2720\text{ cm}^{-1}$ ], imine ( $\nu(\text{C=N}) = 1670\text{ cm}^{-1}$ ]. The successful Pictet–Spengler cyclization is proofed by the disappearance of the imine band.

Besides the monitoring of reactions, which is possible through the appearance or disappearance of characteristic IR absorptions, the study of reaction kinetics (see Sections 16.2.2 and 16.2.3) is also possible with the KBr pellet method. The disadvantage of this method is the time-consuming sample preparation, each sample requiring about 10 minutes.

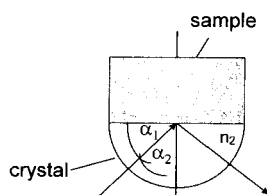


**Figure 16.2.** Control of the synthesis via FT-IR-spectroscopy (KBr-method). From top to bottom: polymer-Z-Cl-resin (P-Z-resin,  $\nu(\text{C}=\text{O}) = 1774 \text{ cm}^{-1}$ ); P-Z-fluorenyl methyl ester ( $\nu(\text{C}=\text{O}) = 1726 \text{ cm}^{-1}$ ,  $\delta(\text{C}-\text{C}, \text{fluorene}) = 740 \text{ cm}^{-1}$ ); P-Z-amino acid (fluorene absorption disappeared); P-Z-amino acid pentafluorophenylester ( $\nu(\text{C}=\text{O}) = 1794 \text{ cm}^{-1}$ ); P-Z-amino alcohol (Pfp-ester absorption disappeared); P-Z-amino aldehyde ( $\nu(\text{C}-\text{H}, \text{aldehyde}) = 2720 \text{ cm}^{-1}$ ); P-Z-imine ( $\nu(\text{C}=\text{N}) = 1670 \text{ cm}^{-1}$ , aldehyde absorption disappeared); P-Z-carboline (imine absorption disappeared). (Reproduced from [26].)



## 16.2.2 ATR-Spectroscopy

ATR (Attenuated Total Reflection) spectroscopy [27] was invented and applied mainly for investigations of surfaces. This method entails an experimental set-up which reflects the IR beam via a mirror system to a crystal with a high refractive index. Normally, germanium (refraction index  $n = 4$ ), silver chloride ( $n = 2$ ) or diamond ( $n = 2,4$ ) crystals are used. The crystal is standing in direct contact with the surface of the sample (Fig. 16.3).



**Figure 16.3.** The principle of FT-ATR-IR spectroscopy.

An IR beam passing through a prism made of a material with a high refractive index  $n_2$  (germanium  $n = 4$ ) under the angle of total reflection ( $n_1 < n_2$ ) is totally internally reflected.

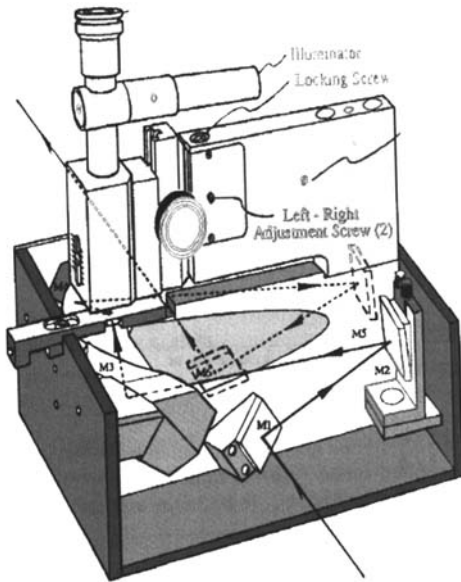
If sample material is in contact with the totally reflecting surface of the prism, an evanescent wave in the sample extends beyond the reflecting interface and the evanescent wave will be attenuated in infrared regions. The intensity of this wave decays exponentially with the distance from the surface of the ATR crystal. Due to the fact that the electromagnetic field passes only a few micrometers of the sample, this method is insensitive to sample thickness and therefore useful for analysis of strong absorbing or thick materials.

Influencing factors for FT-ATR-IR-spectroscopy are as follows:

- difference of the refraction indices of the ATR crystal and the sample at the wave length  $\lambda$
- angle  $\alpha$
- optical contact between ATR crystal and sample
- illuminated area of the crystal
- penetration length of the surface wave

Due to physical reasons, only regions of the sample near to the surface are detected. Furthermore, the relations of the intensities of the absorption bands in ATR spectra are different from those in transmission spectra. The intensity of the spectra in long-wave regions is higher due to the fact that the penetration distance is longer. Normally, quantitative comparisons of spectra are not possible. Only the relation of intensities or integrals of absorption bands can be applied for quantifications when the absorption bands for evaluation are close together.

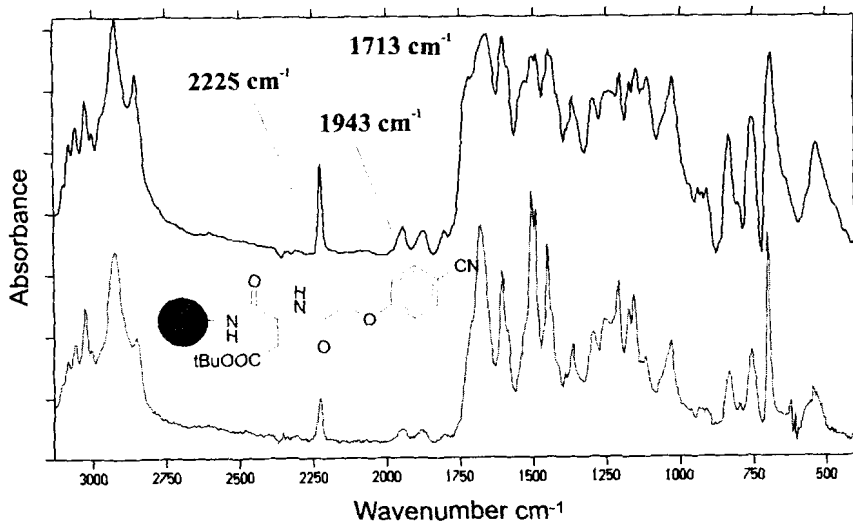
The installation of Harrick's Scientific SplitPea-ATR unit [28] in a common FT-IR spectrometer enables the recording of ATR spectra. This unit contains a microscope for fixing and adjusting the sample on the ATR crystal. The IR beam is reflected via a mirror optic into the germanium crystal (Fig. 16.4).



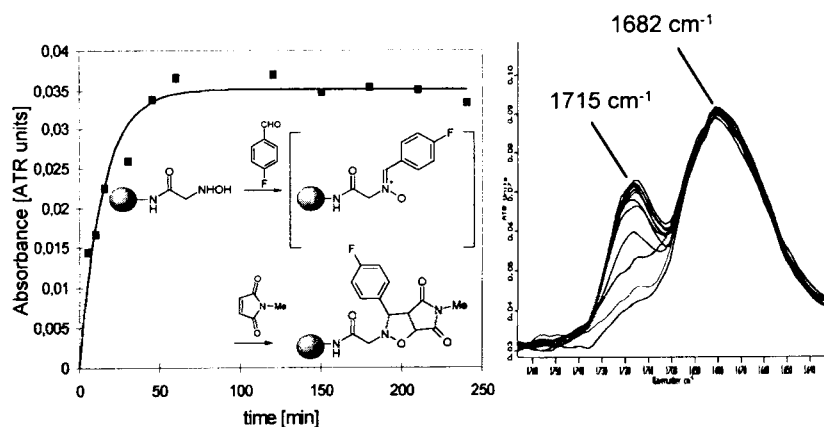
**Figure 16.4.** Harrick Scientific's SplitPea ATR unit.

Figure 16.5 shows the comparison of the IR-spectrum of an aryl ether immobilized on polystyrene resin using the KBr pellet method with a spectrum taken with the ATR method [29]. With regard to the resolution of the spectrometer, the absolute positions of the absorptions are identical.

The main advantage of ATR spectroscopy is the possibility of carrying out measurements without sample preparation, as the resin beads are pressed directly onto the ATR



**Figure 16.5.** FT-IR spectra of a polymer-bound aryl ether. Top: Transmission spectrum using the KBr-pellet method. Bottom: ATR spectrum.



**Figure 16.6.** Kinetics of the reaction to isoxazolidines on a polystyrene matrix determined by ATR-spectroscopy. The relative peak intensities at  $1715\text{ cm}^{-1}$  are applied versus reaction time. The spectra were normalized using the constant amid absorption at  $1682\text{ cm}^{-1}$ . The half life  $t_{1/2} = 10,3\text{ min}$ . and the steady state = 103 min was determined with  $f(t) = a[1 - \exp(-bt)]$ .

crystal. Furthermore, the SplitPea unit allows reactions to be followed on pins, because they can be analyzed directly [30]. Using a constant wavelength, it is possible to create depth profiles of resin beads or pins by varying the angle of incidence. Moreover, using the ATR method, sharper absorption bands are received and, in contrast to the KBr pellet method, less matrix effects occur.

An example for the determination of the kinetics of a solid-phase reaction via ATR-spectroscopy is shown in Figure 16.6. The cyclo-addition reaction leading to substituted isoxazolidines is initialized, and resin samples are taken out of the reaction mixture after certain time intervals [31]. ATR spectra were recorded to determine the relative intensities of the carbonyl vibration ( $\nu(\text{C}=\text{O}) = 1715\text{ cm}^{-1}$ ). The spectra were normalized using the amide vibration ( $\nu(\text{C}(\text{O})\text{NH}_2) = 1682\text{ cm}^{-1}$ ).

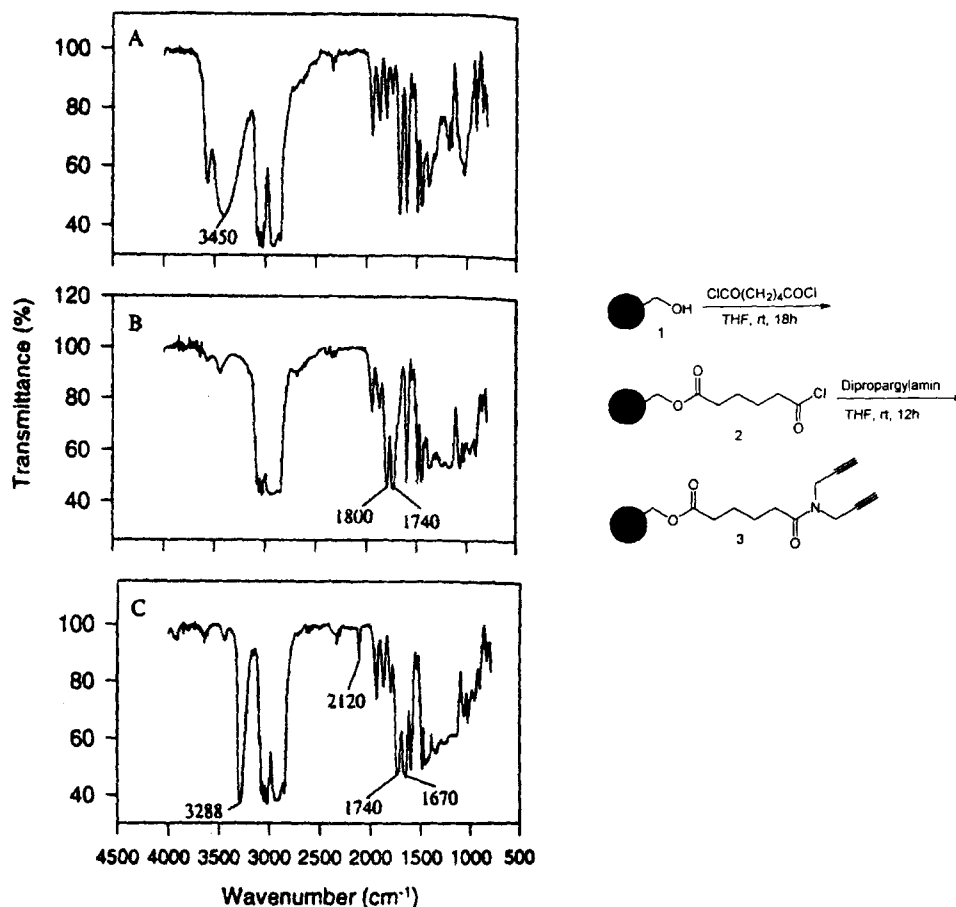
Less than 1 mg of resin is sufficient for the measurement; indeed, by using a micro ATR unit it is possible to examine single resin beads. The authors determined the optimal reaction times for five different solid-phase reactions on a routine basis [22].

## 16.2.3 FT-IR Microscopy

### 16.2.3.1 Single Bead Reaction Monitoring

The methods for monitoring reactions on-bead can also be performed with the help of IR microscopy [32–35]. A few resin beads are distributed on a KBr window, and flattened. The IR beam is focused on the resin bead for transmission measurement.

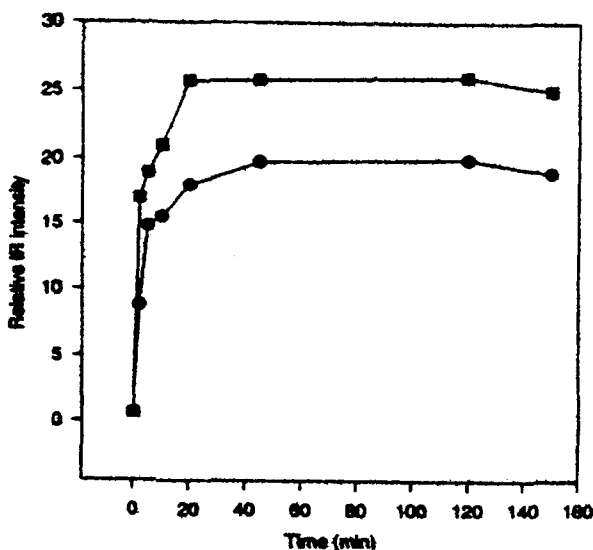
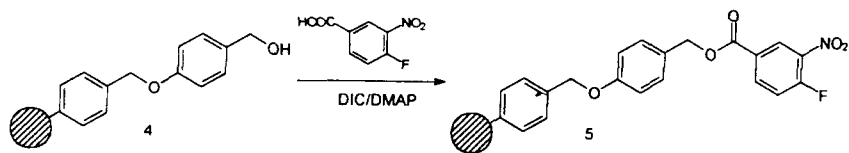
The first use of this method for the monitoring of solid-phase reactions was described by Yan et al. [20]. Polystyrene hydroxymethyl resin as starting material was reacted with adipinic acid chloride and dipropargyle amine (Fig. 16.7).



**Figure 16.7.** Monitoring of the formation of *N,N*-dipropargyle amide via FT-IR-microscopy. (A) Polystyrene hydroxy methyl resin ( $\nu(\text{O}-\text{H}) = 3450 \text{ cm}^{-1}$ ); (B) resin-bound moiety 2 ( $\nu(\text{C}(\text{O})-\text{Cl}) = 1800$ ,  $\nu(\text{COOR}) = 1740 \text{ cm}^{-1}$ ); (C) resin-bound moiety 3 ( $\nu(\text{alkyne C-H}) = 3288$ ,  $\nu(\text{C}\equiv\text{C}) = 2120$ ,  $\nu(\text{COOR}) = 1740$  and  $\nu(\text{CONR}_2) = 1670 \text{ cm}^{-1}$ ). (Reproduced from [20].)

The successful transformation can be verified by the IR spectra shown in Figure 16.7. The characteristic OH vibration at  $3450 \text{ cm}^{-1}$  for the starting material shows spectrum A, which disappears after the treatment with the acid chloride. Therefore the newly formed ester ( $1740 \text{ cm}^{-1}$ ) and the acid chloride ( $1800 \text{ cm}^{-1}$ ) are detected (Fig. 16.7B). The reaction of the acid chloride with dipropargyle amine is indicated by the absorption of the corresponding amide ( $1670 \text{ cm}^{-1}$ ), the disappearance of the vibration from the acid chloride and the C-H valence vibration at  $3288 \text{ cm}^{-1}$  of the hydrogen atoms at the C-C triple bond. This bond is detected by its characteristic absorption at  $2120 \text{ cm}^{-1}$ .

IR-microscopy allows the monitoring of solid-phase reactions during the time course analogous to FT-ATR-IR methods described above. Therefore, resin beads are removed from the reaction mixture at certain time intervals, washed and analyzed. The relation of

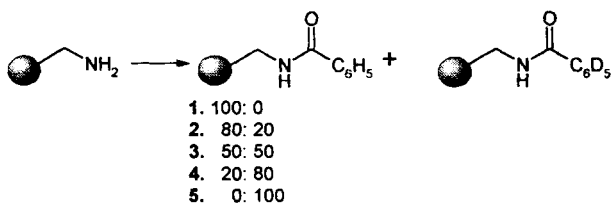
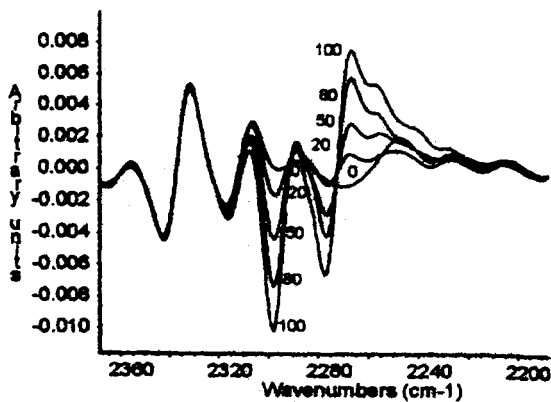
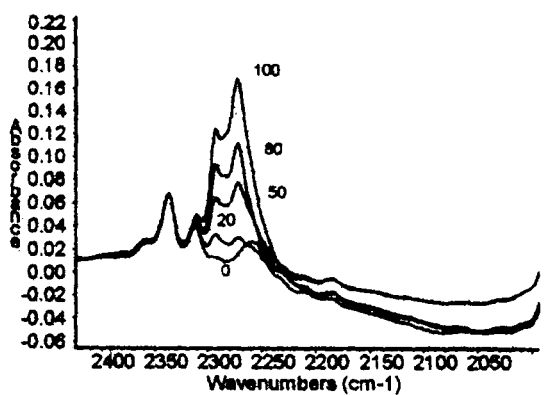


**Figure 16.8.** Time-resolved FT-IR-microscopy. The relative intensities of the carbonyl vibration ( $1722/1943\text{ cm}^{-1}$ ) and the N—O vibration ( $1540/1943\text{ cm}^{-1}$ ) of compound **5** are applied versus time. (Reproduced from [20].)

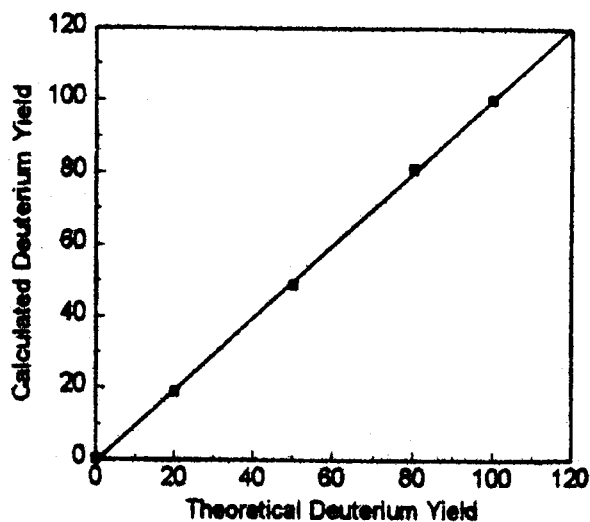
the intensities of the carbonyl- and nitro-absorption bands of compound **5** (Fig. 16.8) were calculated and applied, together with a polymer absorption ( $1943\text{ cm}^{-1}$ ), versus time. This method does not allow an absolute quantification of the resin-bound compound, but a determination of the maximal reaction progress.

The quality of the IR spectra depends on the size of the resin beads. The best results for sample measurement in transmission are received by smaller beads with a maximum diameter of 50  $\mu\text{m}$ . IR-measurements of bigger beads produces overloaded IR spectra. This can be explained by the longer IR beam passing through the bead and interfering with  $\text{CO}_2$  and humidity. When those resin beads are flattened in order to reduce the path of the IR beam through the bead, the quality of the spectra is improved.

Qualitative and quantitative analyzes of resin-bound molecules with any type of IR-spectroscopy is often limited by varying absorption coefficients and an overlapping of different vibrations. The use of deuterated reactants increases the selectivity of IR analysis, due to the fact that the region of C-D stretching vibration absorptions are in the range between  $2300$  and  $2200\text{ cm}^{-1}$ . This region is free of other IR absorptions. An example of quantitative IR analysis is the benzoylation of amino methyl polystyrene resin [23]. Fig-



**Figure 16.9.** Benzoylation of amino methyl resin with various degrees of deuterated reactant (bottom). Five different concentrations of non deuterated/deuterated reactants were used. Top: Aromatic C-D vibration with 0–100% deuterated reactant. Middle: C-D derivative spectra with 0–100% deuterated reactant. Concentration of deuterated reactants: **1**, 0%; **2**, 20%; **3**, 50%; **4**, 80%; **5**, 100%. (Reproduced from [23].)



**Figure 16.10.** Calculated loading versus theoretical loading of resin with deuterated reactant. (Reproduced from [23].)

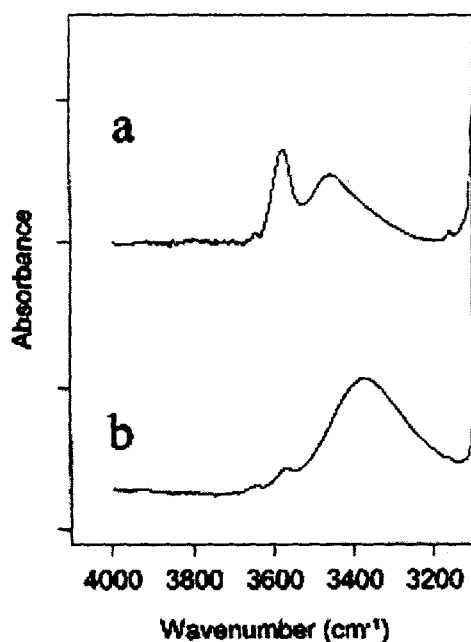
Figure 16.9 shows the IR-absorption spectra of the benzoylated resin with a theoretical loading between 0 and 100% of  $d_5$  benzoic acid. The decrease of the aromatic C-D absorption at  $2290\text{ cm}^{-1}$  can be clearly observed. For calibration, the absorption bands of the spectra (Fig. 16.9, top) they were normalized with respect to the absorption of the polystyrene resin between  $2356$  and  $2342\text{ cm}^{-1}$ . A good correlation of the C-D absorption and the theoretical loading is the result of this investigation (Fig. 16.10).

In principle, this method can be transferred to any other reaction, but the calibration step has to be repeated. According to the authors, it is possible to determine the amount of *Boc-d<sub>9</sub>*-lysine in a mixture of polymer-bound lysines.

The use of deuterated reactants permits the real-time monitoring of solid-phase reactions via IR-microscopy. A few resin beads were trapped in a flow-through cell which can be charged with solvents and reagents using an HPLC pump. These beads were analyzed during the amidation of amino methyl polystyrene resin with deuterated benzoic acid, demonstrating that the reaction was complete after 12.7 min [36, 37].

### 16.2.3.2 Examination of the Interaction between Resin-Bound Reactive Groups via IR-Microscopy

Resin-bound alcohols with different steric environments were examined for hydrogen-bonding interactions [38]. Figure 16.11(a) shows a part of the absorption spectra of *p*-(benzyloxy)benzyl alcohol resin (Wang resin). A splitting of the absorption band to a sharp band at  $\sim 3580\text{ cm}^{-1}$  and a broad band at  $\sim 3420\text{ cm}^{-1}$  is observed. Several experiments were performed to prove whether the sharp band originates from free hydroxy



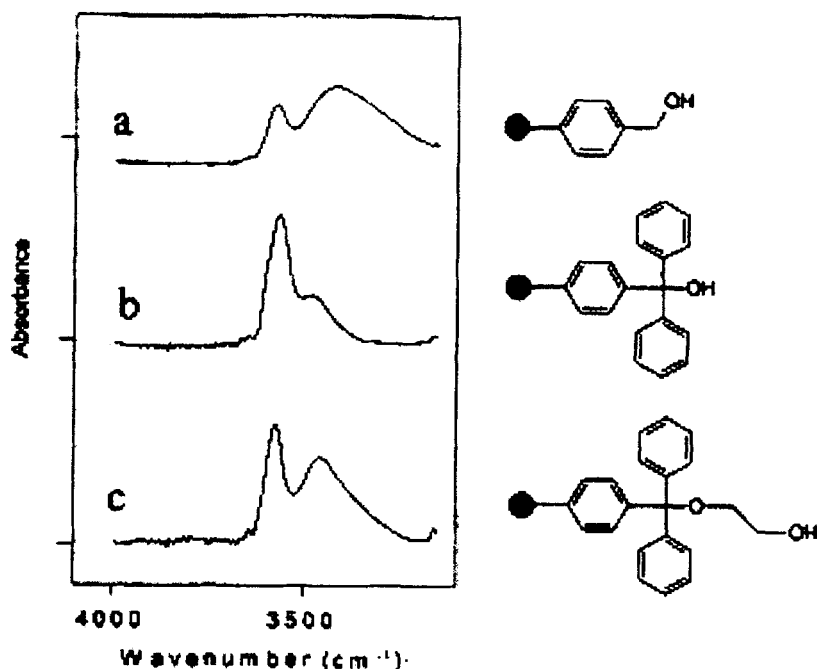
**Figure 16.11.** IR spectra from a single resin bead: (a) Untreated Wang resin. (b) Wang resin treated with DMSO for 2 min and dried under vacuum for 5 min. The O-H stretching vibration at  $3580\text{ cm}^{-1}$  is shifted to  $3420\text{ cm}^{-1}$  by treating the resin with DMSO. (Reproduced from [38].)

groups, and the broad band is due to hydroxy groups which are in mutual interaction or due to hydrogen bonding with water in the air. Confirming that the sharp absorption at  $\sim 3580\text{ cm}^{-1}$  is a result of free hydroxy groups, the resin was treated with the hydrogen bonding-accepting molecule DMSO, and dried under vacuum. The existence of free hydroxyl groups which are in dynamic equilibrium with hydrogen-bonded hydroxy groups was demonstrated by this conversion. Figure 16.11(b) shows the resulting spectrum with the near-disappeared band at  $\sim 3580\text{ cm}^{-1}$ .

To solve the problem whether the broad band is from hydroxy groups in mutual interaction, or from hydrogen bond between hydroxy groups and water in the air, hydroxylated resins with different bulky groups near to the OH group were examined. The increasing steric hindrance inhibits hydrogen bonding between neighboring hydroxy groups, but not the interaction with water molecules.

A comparison of the three spectra in Figure 16.12 shows that hydroxymethyl polystyrene resin has the broadest absorption band at  $3420\text{ cm}^{-1}$  and the smallest absorption at  $3580\text{ cm}^{-1}$ . The relationship of the bands in the spectra of trityl alcohol resin is inverted, indicating the presence of a large portion of free hydroxy groups. The introduction of the  $\text{C}_2$ -chain between trityl and hydroxy groups partially reversed the status and produced an increase of the absorption at  $3420\text{ cm}^{-1}$ . This correlation between different bulky substituents and different amounts of free hydroxy groups demonstrates that the origin of the broad band is the association of neighboring hydroxy groups via hydrogen bonds.

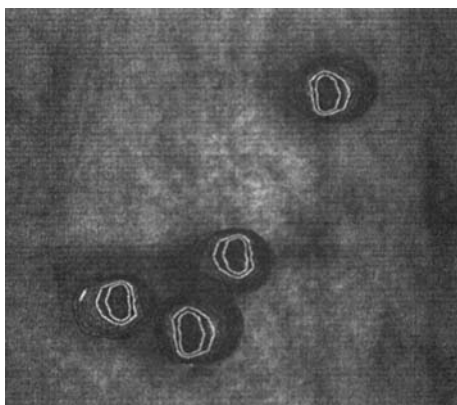




**Figure 16.12.** Structure of the linker group and single bead IR spectra. (a) Wang resin. (b) Trityl alcohol resin. (c) Ethylene glycol bound to trityl resin. The intensity of the hydroxy absorption at  $3580\text{ cm}^{-1}$  changes by the substituent near to hydroxy group. (Reproduced from [38].)

### 16.2.3.3 FT-IR Mapping: A New Tool for Spatially Resolved Characterization of Polymer-Bound Combinatorial Compound Libraries with IR-Microscopy

With the techniques that have now been established for analyzing single resin beads such as NMR spectroscopy, matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry, and FT-IR spectroscopy, only polymer-bound single compounds could be analyzed so far. In combinatorial chemistry, however, techniques are needed with which a large number of compounds can be analyzed simultaneously. Owing to its high sensitivity, the simplicity of the apparatus, and the high sample throughput, IR microscopy is suited for parallel analysis. The combination of a FT-IR microscope with a motor-driven x-y stage allows the automatic spectral mapping of surfaces and objects embedded in them. If single resin beads are embedded in a KBr window by pressing, they can be visualized with this measuring arrangement in the form of an IR map. The superpositioning of the visible video image of several resin beads with the corresponding IR map (Fig. 16.13) shows the good correlation of both types of representations. In this case, for the IR reconstruction the typical polystyrene combination vibration at  $1942\text{ cm}^{-1}$  was chosen. Through mapping of larger areas, hundreds of embedded resin beads can be detected and the compounds bound to them identified.



**Figure 16.13.** UV/visible video image and IR reconstruction of four polystyrene resin beads with a diameter of about 80  $\mu\text{m}$ . The IR map was obtained with a motor-driven x-y stage in combination with an IR microscope. For the reconstruction of the resin beads the integral of the polystyrene combination vibration at  $1942\text{ cm}^{-1}$  was used (Reproduced from [25]).

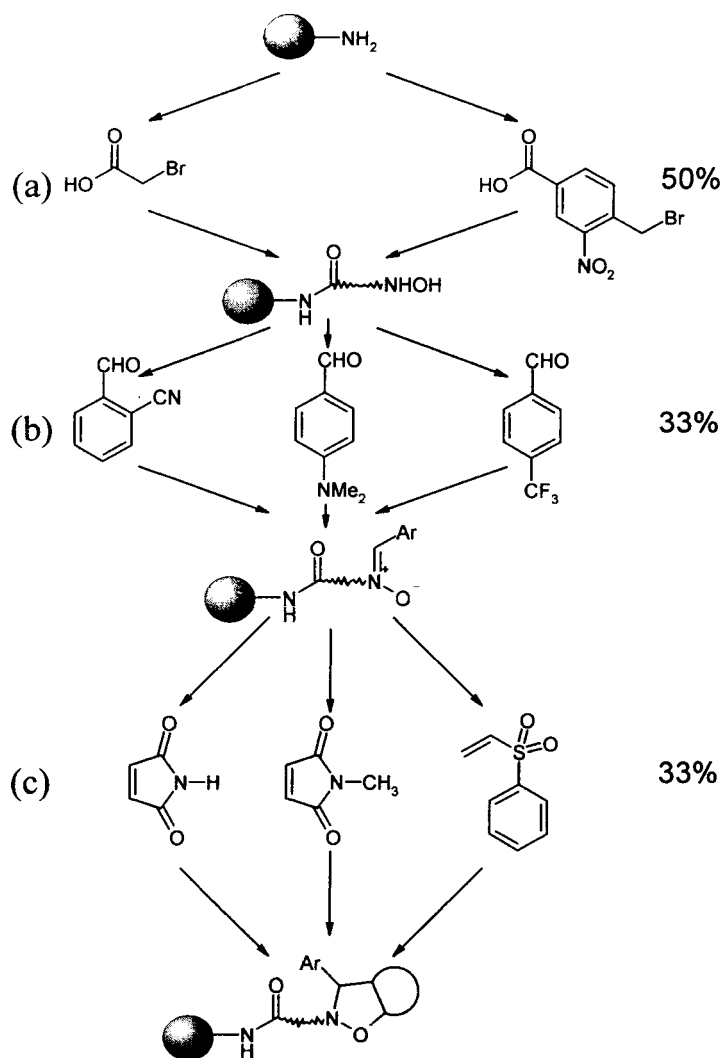
To demonstrate the potential of this method, a model library of isoxazolidines on Rink amide resin was synthesized by the split-and-combine method [25] (Scheme 16.1). The polymer-bound library obtained consists of a total of 18 compounds. If one considers the different building blocks of the synthesis, a nitro group or the carboxamidomethyl residue is found in 50% each of the compounds, a cyano, N,N-dimethylamino, or trifluoromethyl group in 33% each; and a succinimide, N-methylsuccinimide, or sulfone group in 33% each.

To verify that this library was successfully synthesized, a portion of the resin was removed, and the compounds were cleaved from the polymer with trifluoroacetic acid and analyzed with HPLC electrospray ionization mass spectrometry. All the compounds shown could be identified. On cyclo-addition of the maleimides, endo/exo isomers and – in the case of the vinylsulfones – regioisomers were obtained. This led to complex chromatograms of isomeric compound mixtures; co-elution often occurred. Quantification of the single compounds was not possible, and the peak area in the UV trace (detection at 214 nm) did not correlate with those of the total ion current chromatogram in the mass spectrum.

The isoxazolidine library prepared according to Scheme 16.1 was embedded in a KBr window by pressing. An area of  $3\times 3\text{ mm}^2$  was automatically mapped. When the polystyrene absorption at  $1942\text{ cm}^{-1}$  was used to visualize the IR map, all resin beads on the surface were detected (Fig. 16.14).

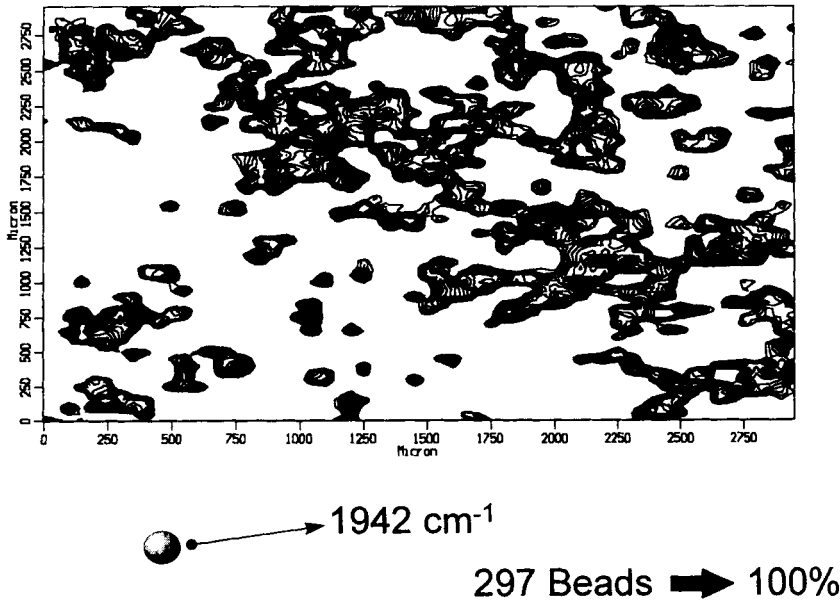
Through a choice of different absorption bands for visualization, those resin beads carrying a specific functional group can be selectively ‘fished out’ from this library of 297 resin beads. For example, the resin beads with the N-methylsuccinimide group are extracted spectroscopically from the resin bead mixture by visualization by means of the carbonyl absorption at  $1715\text{ cm}^{-1}$ . In the scanned area there are a total of 112 resin beads (39.1%) with this functional group (Fig. 16.15) Analogous IR maps could be obtained for further functional groups (Fig. 16.16).

The IR analysis of polymer-bound molecules is made difficult by the strong absorption of the polymer matrix itself. In the IR spectrum containing polymer and anchor absorptions, for example, signals for neither the N,N-dimethylamino nor the carboxamidomethyl groups can be recognized. However, as all further functional groups can be clearly

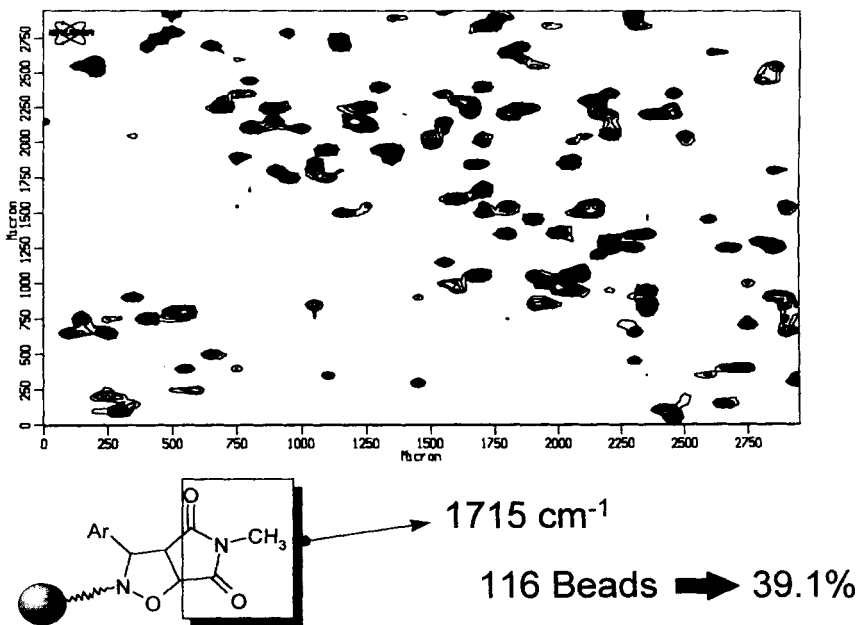


**Scheme 16.1** Solid-phase synthesis of isoxazolidines according to the split-and-combine method. (a) Distribution of the resin into two equal portions, coupling of bromocarboxylic acids with *N,N'*-diisopropylcarbodiimide (DIC), combination of the resin, substitution with hydroxylamine. (b) Distribution of the resin into three equal portions, condensation with three different aromatic aldehydes to the corresponding nitrones, combination of the resin. (c) Distribution of the resin into three equal portions, cyclo-addition with three different dipolarophiles to isoxazolidines, combination of the resin.

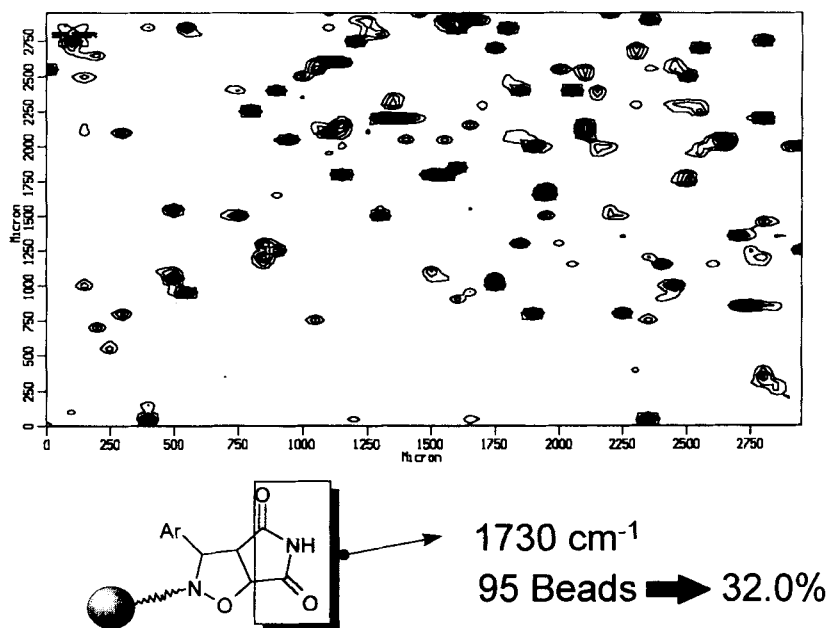
differentiated one from another, the number of resin beads carrying a specific group is given by the exclusion principle. If this method is applied to all resin beads, the distribution of all functional groups can be determined by single counting of the resin beads by data processing programs (Table 16.1). Despite the relatively small amount of resin beads in the scanned regions, no significant deviations from the expected statistical distribution



**Figure 16.14.** IR map of an area ( $3 \times 3$  mm) of a KBr window in which resin beads of the isoxazolidine library are embedded. By application of the polystyrene absorption at  $1942 \text{ cm}^{-1}$  all resin beads were detected. White region: KBr window; gray region: resin beads. There is a total of 297 resin beads in the scanned area.



**Figure 16.15.** 'N-Methylsuccinimide' resin beads extracted from the library by visualization over the carbonyl absorption at  $1715 \text{ cm}^{-1}$ .



**Figure 16.16.** ‘Succinimide’ resin beads extracted from the library by visualization over the carbonyl absorption at  $1730\text{ cm}^{-1}$ .

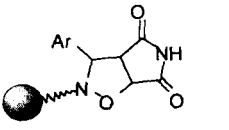
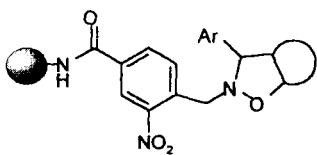
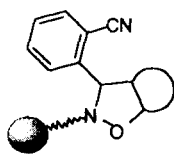
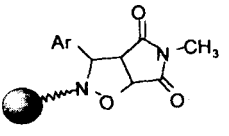
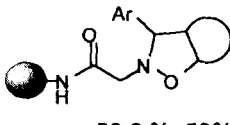
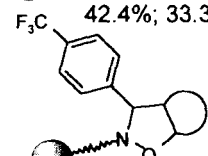
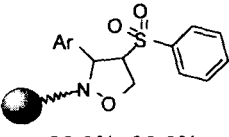
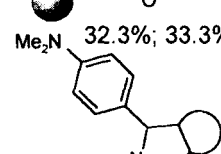
of the individual functional groups are observed. Although the carboxamidomethyl, cyano, and N-methylimide groups are somewhat over-represented, their contributions are on the same order of magnitude.

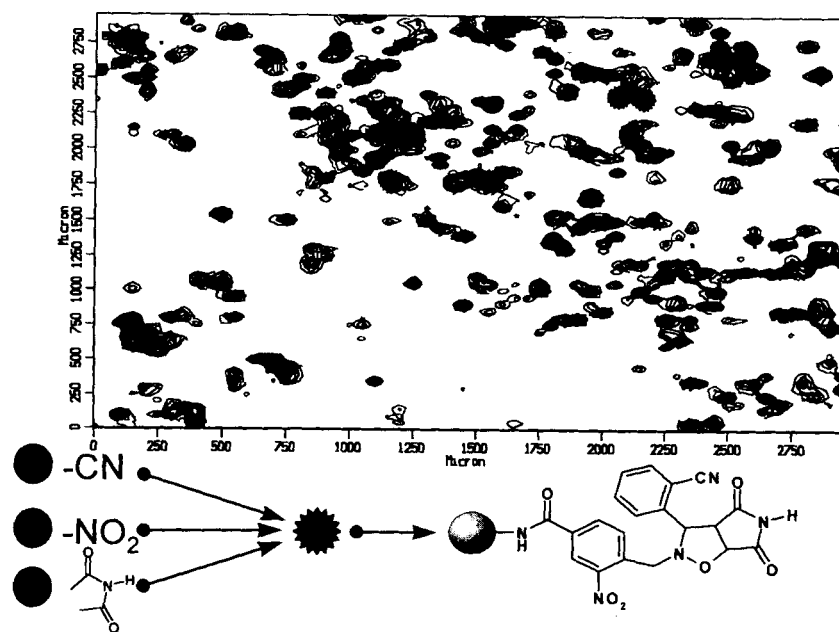
When applied together with a LC-MS analysis, IR mapping allows a statistical investigation and thus a rapid quality control of the library synthesized by the split-and-combine method. The decisive advantage of this technique is that IR analysis, in contrast to mass spectrometry, can be carried out without destruction of the sample, and with spatial resolution.

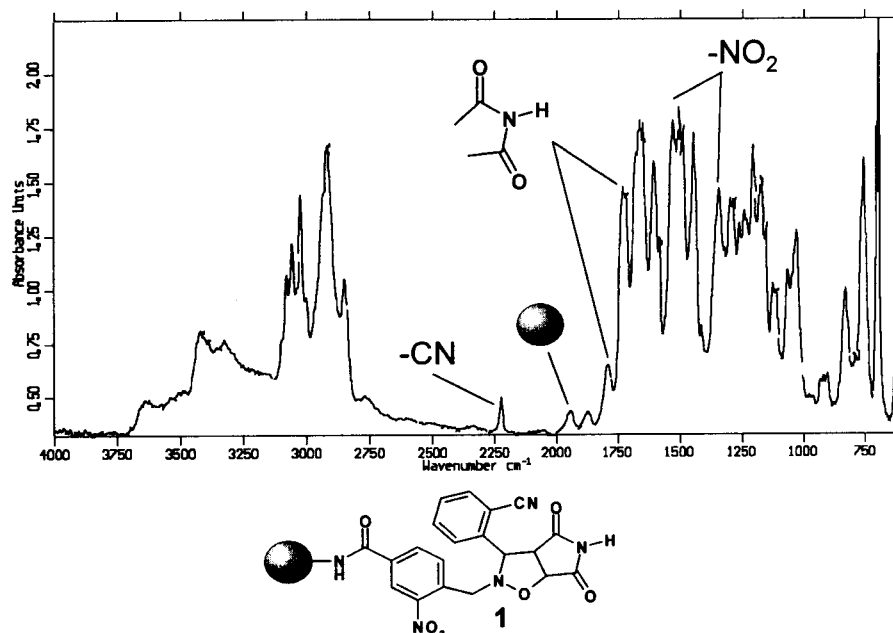
Furthermore, a direct identification of defined resin-bound molecules is possible through superpositioning of the IR maps. If, for example, the maps of the imide ( $1730\text{ cm}^{-1}$ ), nitro ( $1360\text{ cm}^{-1}$ ), and cyano absorptions ( $2225\text{ cm}^{-1}$ ) are placed on top of each other, the resin beads that carry exclusively compound **1** can be identified with spatial resolution in the overlapping regions (Fig. 16.17, stars) of the functional groups. The FT-IR spectra of the individual resin beads extracted from the maps clearly confirm attachment of **1** to the resin bead (Fig. 16.18).

IR microscopy analysis allows the spatially resolved and destruction-free characterization and quality control of polymer-bound compound libraries. Furthermore, by superpositioning of different IR maps, compounds immobilized on individual resin beads can be identified. Through improved embedding techniques, even larger areas with many resin beads can be mapped very quickly, thus improving the statistical evaluation by image analysis.

**Table 16.1.** Statistical distribution (%) (theoretical values given in parentheses) of resin beads with specific functional groups.

Functional group amount [%]; (theoret.)	Functional group amount [%]; (theoret.)	Functional group amount [%]; (theoret.)
 32.0%; 33.3%	 41.4%; 50%	 42.4%; 33.3%
 39.1%; 33.3%	 58.6%; 50%	 32.3%; 33.3%
 28.9%; 33.3%		 25.3%; 33.3%

**Figure 16.17.** ‘FT-IR tagging’ by superposition of the IR maps of imide ( $1730\text{ cm}^{-1}$ ), nitro ( $1360\text{ cm}^{-1}$ ), and cyano absorptions ( $2225\text{ cm}^{-1}$ ). Resin beads on which compound **1** is immobilized were identified through overlapping regions of all IR maps.



**Figure 16.18.** FT-IR spectrum of an individual resin bead carrying one defined compound **1**, demonstrating the one-bead-one-compound principle of a split-mix synthesis of isoxazolidines.

This method can thus provide a valuable contribution to the development of miniaturized syntheses on individual resin beads. The high analytical power of combining IR mapping with MS and HPLC-MS analysis, in particular high-resolution FT-ICR-MS (see Chapter 20), means that tagging concepts can be abandoned.

## 16.2.4 Other IR-Spectroscopy Methods

The following spectroscopic methods can be also applied for on-bead analysis; however, they are not suitable for routine high-throughput analysis but more for special research projects.

### 16.2.4.1 DRIFTS (Diffuse Reflectance Infrared Fourier Transform) Spectroscopy

The portion of the reflected IR beam caused by diffuse reflectance is examined. This is the part of the radiation which interacted with the sample. The optics collect this scattered IR light and direct it to the detector. The required amount of sample material, quality of the spectra, time for measurement and the costs for the equipment are comparable with the KBr pellet method.

#### 16.2.4.2 Photoacoustic Spectroscopy [40]

Modulated infrared energy is absorbed by the sample. The material heats and cools in response to modulated infrared light. This response is converted into a pressure wave, which communicates with a surrounding gas. The pressure change is detected by an acoustic detector in the enclosed sample chamber. The necessary amount of sample material is less than 10 mg. The measuring time is about a few minutes, and the quality of the spectra is similar to that of the DRIFTS and ATR methods. The costs for the equipment are higher. A comparison of photoacoustic spectroscopy with all various FT-IR methods for the use in combinatorial chemistry was made by Yan et al. [12, 13].

#### 16.2.4.3 FT Raman Spectroscopy

This method is based on inelastic scattering of photons which exchange energy with the sample. Because of the different selection rules for FT-IR and FT-Raman spectroscopy, additional information for the resin-bound moiety can be obtained. In particular, the gain of information is helpful for the characterization of symmetric molecules. Strong absorptions can be obtained for vibrations of functional groups with high polarizability and low polarity. The intensity of the absorption is also enhanced by conjugation of multiple bonds. The spectra can be acquired by measurement of less than 1 mg of resin beads, or using FT-Raman microscopy with a single bead. The cost of the FT-Raman accessory is higher than that of FT-IR equipment.

### 16.3 Conclusions

FT-IR spectroscopy has become an established method in solid-phase chemistry, because it is a fast, sensitive and convenient method for monitoring resin-bound reactions with relatively low equipment costs.

KBr pellet or single bead samples are suitable for the purpose of solid-phase reaction monitoring and the study of reaction kinetics. For the KBr-embedded samples, a higher amount of resin material and longer sample preparation time is required, but these are compensated by the low costs for the measurement with KBr-pellets. The compromise here is the use of an ATR-instrumentation. With this technique the throughput of samples is much higher at low costs for additional equipment. ATR-IR spectra quality is at the same level than for the common methods mentioned before.

The use of single bead IR spectroscopy is a valuable method for real-time monitoring of reaction kinetics or investigations of small amounts of resin beads.

The most interesting examination method with IR microscopy is the spatially resolved mapping of resin-bound compound libraries. A non destructive statistical characterization of resin-bound libraries is possible by this method. However, the investment for IR-microscopy is relatively high, especially when used in combination with image analysis.



## 16.4 References

- [1] G. Jung, A. G. Beck-Sickinger, *Angew. Chem., Int. Ed. Engl.*, **1992**, *31*, 367–375.
- [2] L. A. Thomson, J. A. Ellman, *Chem. Rev.*, **1996**, *96*, 555–600.
- [3] F. Balkenhohl, C. Bussche-Hunnefeld, A. Lansky, C. Zechel, *Angew. Chem.*, **1996**, *108*; 2436–2488; *Angew. Chem., Int. Ed. Engl.*, **1996**, *35*, 2289–2337.
- [4] Früchtel, J. S.; Jung, G., *Angew. Chem.*, **1996**, *108*, 19–46; *Angew. Chem., Int. Ed. Engl.*, **1996**, *35*, 17–42.
- [5] Jung, G., (Ed.): *Combinatorial Peptide and Non-peptide Libraries*. Verlag Chemie, Weinheim **1996**.
- [6] E. M. Gordon, M. A. Gallop, D. V. Patel, *Acc. Chem. Res.*, **1996**, *29*, 144–154.
- [7] R. B. Merrifield, *J. Am. Chem. Soc.*, **1963**, *85*, 2149–2154.
- [8] R. Brown, *Contemp. Org. Synth.*, **1997**, *4*, 216–219.
- [9] P. H. H. Hermkens, H. C. J. Ottenheijm, D. C. Rees, *Tetrahedron*, **1997**, *53*, 5643–5678.
- [10] S. K. Sakar, R. Garigipati, J. L. Adams, P. A. Keifer, *J. Am. Chem. Soc.*, **1996**, *118*, 2305–2306.
- [11] M. Pursch, G. Schlotterbeck, L.-H. Tseng, K. Albert, W. Rapp, *Angew. Chem.*, **1996**, *108*, 3034–3036; *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 2867–2869.
- [12] B. Yan, *Acc. Chem. Res.*, **1998**, *31*, 621–630.
- [13] B. Yan, H. U. Gremlich, S. Moss, G. M. Coppola, Q. Sun, L. Liu, *J. Comb. Chem.*, **1999**, *1*, 46–54.
- [14] C. Chen, L. A. A. Randall, R. B. Miller, A. D. Jones, J. M. Kurth, *J. Am. Chem. Soc.*, **1994**, *116*, 2661–2667.
- [15] J. I. Crowley, H. Rapoport, *J. Org. Chem.*, **1980**, *45*, 3215–3219.
- [16] W. J. Haap, T. Redemann, J. S. Früchtel, *GIT*, **1997**, *4*, 97–99.
- [17] J. Ryttersgaard, B. D. Larsen, A. Holm, D. H. Christensen, D. F. Nielsen, *Spectrochim. Acta, Part A*, **1997**, *53*, 91–95.
- [18] B. D. Larsen, D. H. Christensen, A. Holm, R. Zillmer, O. F. Nielsen, *J. Am. Chem. Soc.*, **1993**, *115*, 6247–6251.
- [19] T. Y. Chan, R. Chen, M. J. Sofia, B. C. Smith, D. Glennon, *Tetrahedron Lett.*, **1997**, *38*, 2821–2826.
- [20] B. Yan, G. Kumaravel, H. Anjaria, A. Wu, R. C. Petter, C. F. Jewell, J. R. Wareing, *J. Org. Chem.*, **1995**, *60*, 5736–5738.
- [21] B. Yan, G. Kumaravel, *Tetrahedron*, **1996**, *52*, 843–848.
- [22] B. Yan, J. B. Fell, G. Kumaravel, *J. Org. Chem.*, **1996**, *61*, 7467–7472.
- [23] K. Russell, D. C. Cole, F. M. McLaren, D. E. Pivonka, *J. Am. Chem. Soc.*, **1996**, *118*, 7941–7945.
- [24] D. E. Pivonka, K. Russell, T. Giero, *Appl. Spectr.*, **1996**, *50*, 1471–1476.
- [25] W. J. Haap, T. Walk, G. Jung, *Angew. Chem.*, **1998**, *110*, 3506–3509; *Angew. Chem., Int. Ed. Engl.*, **1998**, *23*, 3311–3314.
- [26] T. Redemann, *Dissertation*, Universität Tübingen, **1997**.
- [27] H. Günzler, H. Böck, *IR-Spektroskopie*. Verlag Chemie, VCH, Weinheim, **1990**.
- [28] N. J. Harrick, M. Milosevic, S. L. Berets, *Appl. Spectrosc.*, **1991**, *45*, 944–949.
- [29] W. J. Haap, *Dissertation*, Universität Tübingen, **1998**.
- [30] H.-U. Gremlich, S. L. Berets, *Appl. Spectrosc.*, **1996**, *50*, 532–536.
- [31] W. J. Haap, D. Kaiser, T. B. Walk, G. Jung, *Tetrahedron*, **1998**, *54*, 3705–3724.
- [32] B. Yan, L. Wenbao, *J. Org. Chem.*, **1998**, *63*, 4092.
- [33] S.S. Rahman, D. J. Busby, D. C. Lee, *J. Org. Chem.*, **1998**, *63*, 6196–6199.
- [34] B. Yan, H. Gstach, *Tetrahedron. Lett.*, **1996**, *37*, 8325–8328.
- [35] B. Yan, Q. Sun, J. R. Wareing, C. F. Jewell, *J. Org. Chem.*, **1996**, *61*, 8765–8770.
- [36] D. E. Pivonka, K. Russell, T. Giero, *Appl. Spectr.*, **1996**, *50*, 1471–1476.
- [37] D. E. Pivonka, T. R. Simpson, *Anal. Chem.*, **1997**, *69*, 3851–3855.
- [38] B. Yan, Q. Sun, *J. Org. Chem.*, **1998**, *63*, 55–58.
- [39] A. Furka, F. Sebestyén, M. Asgedom, G. Dibo, *Int. J. Pept. Protein Res.*, **1991**, *37*, 487–493.
- [40] F. Gosselin, M. Di Renzo, T. H. Ellis, W. D. Lubell, *J. Org. Chem.*, **1996**, *61*, 7980–7984.

# 17 Mass Spectrometric Analysis of Combinatorially Generated Compounds and Libraries

*Roderich Süßmuth, Axel Trautwein, Hartmut Richter,  
Graeme Nicholson and Günther Jung*

## 17.1 Introduction

The search for new drugs has received new impulses from combinatorial synthesis. Much research has been undertaken on making and understanding of syntheses. However, despite some exceptions, *e.g.*, in the field of mass spectrometry [1, 2], in reviews and monographs analytical aspects are dealt shortly and in vague terms. In particular, mass spectrometry plays an important and decisive role in the analysis of combinatorially generated compounds. This is also due the explosive developments in the electronics, automation and data handling which have increased the capabilities of mass spectrometers enormously with respect to sensitivity, resolution and sample throughput.

The past years have seen the development of two important ionization techniques: electrospray ionization (ESI) and matrix-assisted laser desorption (MALDI), which have been used especially in conjunction with quadrupole, time of flight (TOF) and ion cyclotron (IC) analyzers. Both ionization techniques have revolutionized the essence of mass spectrometry itself, radically increasing the upper mass limit and thus the applicability of the mass spectrometric technique and indirectly affecting the endeavors for improved resolution and sensitivity. With these techniques, the analysis of peptides, proteins and oligonucleotides has become possible. The aim of this chapter is to demonstrate the capabilities of these techniques, especially electrospray mass spectrometry in conjunction with HPLC techniques, for the analysis of combinatorially generated compounds and libraries, and to exemplify and discuss both the potentials and limits of these analytical techniques.

## 17.2 Significance of Mass Spectrometry in the Analytical Concept of Combinatorial Chemistry

### 17.2.1 Decisional Pathway for Combinatorial Synthesis and Analysis

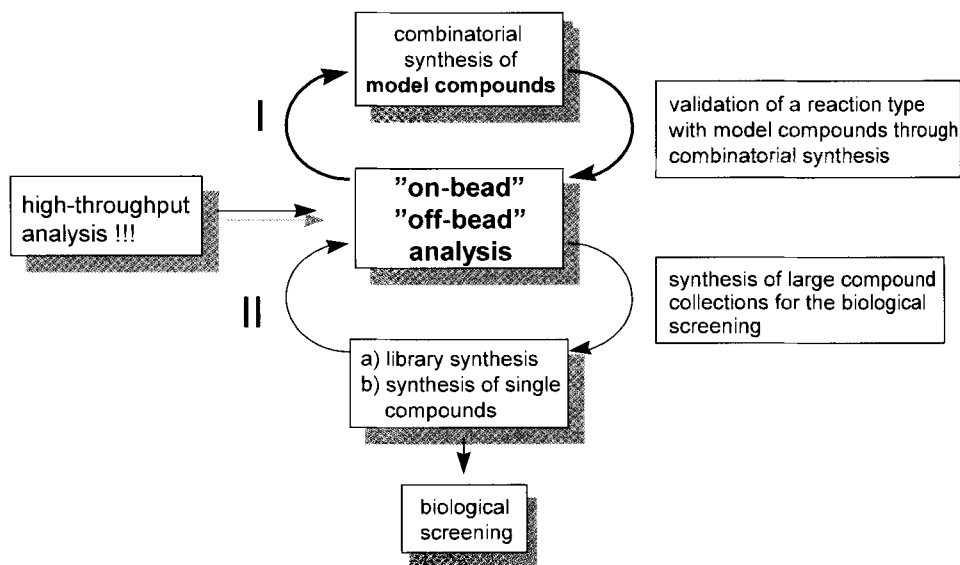
There is an enormous wealth of diverse methodologies and methods at our disposal to comprehend and to manage the concept of combinatorial chemistry. In this article however, we wish to focus our discussion on molecules synthesized by organic synthesis. The majority of combinatorial compounds hitherto produced have been synthesized on solid support, either as collections of single compounds or as libraries. The choice between these strategies varies from case to case and depends on individual preferences. In gener-

al, the compounds are synthesized on a small scale, generally in the range of a few milligrams or even less. This amount must suffice both for the analytical investigations and for a number of bioassays.

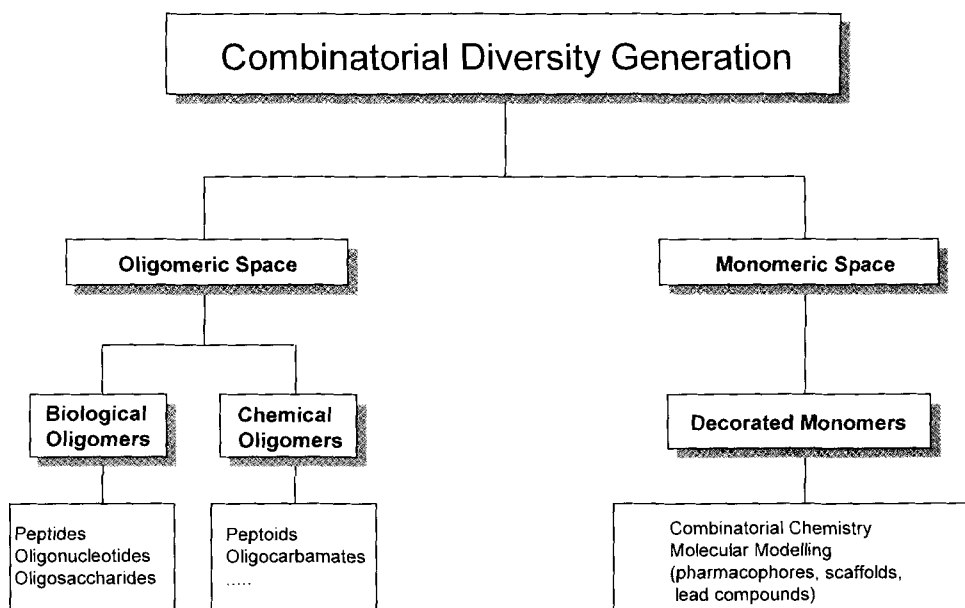
In Figure 17.1, a typical decisional pathway leading to the development of drugs by combinatorial chemistry is shown in schematic form. This scheme is valid irrespective of whether the target compound or class of compounds has been proposed by molecular modeling experiments, or is based upon known structures. In cycle I, the reaction is validated analytically for a few model compounds. Validation will generally involve the use of IR-spectroscopy and mass spectrometry, usually in conjunction with a separation technique (*e.g.*, HPLC). Additionally, in the case of single compounds, an NMR analysis is performed using one- or two-dimensional NMR techniques to confirm the validated synthesis representatively on one model compound. Therefore, an appropriate use of NMR would appear in the final verification of structure where only single representatives from a collection of hundreds of compounds are investigated.

In cycle II, the decision must be made as to whether single substances or libraries should be synthesized before these are then submitted to biological screening. Also, for the analysis of the large numbers of compound collections which can be produced after validating the reaction conditions, mass spectrometry is the method of the first choice. However, an aspect which is not considered in this contribution is the use of mass spectrometry for screening purposes [2].

Compounds generated by the combinatorial approach can be classified according to different aspects. One possible and meaningful attempt is to go for combinatorial diversity (Fig. 17.2) [3] by considering all conceivable organic molecules as a **combinatorial space**. The latter can be divided into an oligomeric and a monomeric space. It is important



**Figure 17.1.** Development of drugs by combinatorial chemistry.



**Figure 17.2.** Combinatorial diversity generation.

to realize that these considerations also have consequences for the mass spectrometric analysis of these compounds.

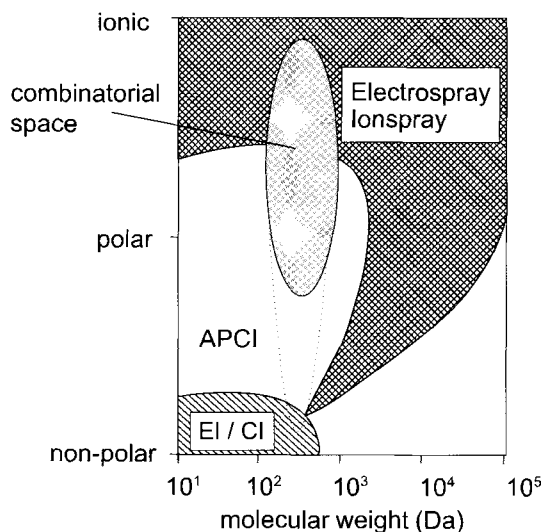
The compounds of the **oligomeric space** comprise a limited set of recurring units, building up a sequence which is defined by its length and order. We can say that nature explores the oligomeric space, with peptides or proteins as prime examples. With respect to peptides and oligonucleotides, the analysis of oligomeric structures from the natural and synthetic pool by means of ESI and MALDI techniques are well established [4]. The potential of these methods can be demonstrated on the example of synthetic peptide libraries [5] which might be considered as the first decisive anticipation and realization of combinatorial ‘thinking’. The combinatorial approach indeed represents a scientific revolution and a significant shift of paradigm [6]. The turning towards biopolymer mimetics (*e.g.*, peptoids, oligocarbamates, etc.), to overcome problems of biological instability and bioavailability represents the beginning of this shift of paradigm. Analysis of these compounds is preferably performed with the above-mentioned mass spectrometric techniques (ESI/MALDI) [7], and can be considered as well established.

Unlike oligomeric structures, molecules within the **monomeric space** consist of a lead structure, with diversity being introduced through variation of the substituents. Bioavailable non-oligomeric molecules below a molecular mass of 600–700 Da have become major targets of synthetic and pharmacological research. This becomes apparent from the annual reports on drug sales on the world market. Most of the pharmacological ‘bestsellers’ are representatives of this monomeric class of compounds. A large number of drugs bear polar or ionic functional groups, *e.g.*, alcohols, phenols, amines, amides, carboxylic acids, nitro groups, halides and heterocyclic components.

There is no doubt that for the analysis of all these compounds, mass spectrometry is the first choice to obtain information about the molecular masses of the investigated compounds. Mass spectrometry offers high information content with extremely low sample consumption. In principle, any mass spectrometric method can be employed. From the diverse ionization techniques which are available for the analysis of combinatorially synthesized compounds, especially the ESI-technique but also the MALDI-technique have gained most common acceptance. The reasons for this development are various and are discussed in the following sections.

### 17.2.2 Mass Spectrometric Techniques for Combinatorial Compound Analysis

In the following section we consider mass spectrometric analysis with respect to the above-mentioned target properties of combinatorially generated compounds. Figure 17.3 shows the areas of application appropriate to the various hyphenated separation and mass spectrometric techniques, with respect to molecular mass and polarity of the molecules to be investigated [8]. With some restrictions, this schematic diagram is also applicable to mass spectrometry without the use of coupling techniques. The diversity of combinatorially generated compounds, which is the issue of our considerations, can mainly be located primarily in the area of compounds with polar to ionic properties. Let us briefly consider some of the common mass spectrometric techniques and their relevance to combinatorial chemistry. Readers who are already familiar with this subject can proceed immediately to Section 17.2.3.



**Figure 17.3.** Application range of ionization techniques for combinatorial chemistry.

### 17.2.2.1 Ionization Techniques

A range of these techniques is available:

- **Electron impact (EI)** is the most common ionization technique for organic molecules. Ionization is accomplished by bombardment of the sample with energetic electrons. It is restricted to volatile substances. The radical ions generated readily undergo fragmentation, providing additional structural information to the molecular mass. However, fragmentation may be so intense that molecular ions cannot be detected. In such cases **chemical ionization (CI)** [9] proves a useful complementary method. Ionization is accomplished by the transfer of a charged species (*e.g.*, a proton) from a reactant gas plasma to the sample molecule. The transfer of energy in this process is low and formation of the protonated molecular ion is generally favored. CI is usually performed on volatilized samples, although Baldwin and McLafferty developed a CI method (desorption chemical ionization; DCI) [10] for non volatile substances. Both EI and CI are well-suited to coupling with GC, which shares the same restrictions regarding volatility. However, the involatility of the majority of combinatorial libraries, either on account of their high molecular mass or their high polarity, severely limits the use of EI and CI in combinatorial analysis.
- By **fast atomic bombardment (FAB)** [11], ions of polar compounds with molecular mass up to 10000 Da may be generated. The sample is dissolved in a non volatile matrix (*e.g.*, nitrobenzoic acid) and ionized by bombardment with neutral gas atoms, *e.g.*, argon. Continuous-flow FAB (CF-FAB) [12] allows the use of the FAB-ionization technique in conjunction with an LC-coupling.
- In **Matrix-assisted laser desorption (MALDI)** [13], the sample is allowed to crystallize together with an organic acid. Ionization is achieved by UV-laser shots being absorbed by the matrix with subsequent proton transfer. MALDI allows the ionization of polymers and proteins with molecular masses of up to 500000 Da. The major restriction of the MALDI-technique is its incompatibility with separation methods.
- For samples dissolved in ionic or polar solvents, the spray ionization techniques are of particular importance. Differences between spray ionization techniques lie mainly in the environment in which the spray is created. **Electrospray ionization (ESI)** is operated at atmospheric pressure at a temperature which is close to ambient. In electrospray, the dissolved sample is introduced via a capillary onto the needle tip. The needle tip is held at a voltage of about 3–6 kV relative to the counter ion electrode. The polarity of this voltage may be reversed to switch between positive and negative ion measurement. The applied high electric field induces a charge accumulation at the liquid surface located at the end of the capillary, this causing the liquid to disperse and to form highly charged droplets which rupture to form smaller droplets from which ions are finally desorbed. Formation of the spray can be supported by a coaxial nebulizing air flow (Ion-spray, IS) [14]. A major advantage of the ESI technique is the ‘soft’ nature of the ionization process and the possibility to perform LC- and CZE-couplings. The available mass range of typical quadrupole analyzers allows direct detection of molecules within the mass range of 100–4000 Da, or indirectly by detection of multiply charged ions to a molecular mass of 200000 Da.

- **Atmospheric pressure chemical ionization (APCI)** [15] involves the primary formation of ions by corona discharge in a solvent spray under atmospheric pressure. A solution of the sample is nebulized pneumatically and the droplets formed are vaporized by heating (120 °C). The resulting sample gas is chemically ionized by the transfer of protons from the primary reactant ions. The APCI technique is widely used in the analysis of drugs or metabolic studies. Detection is limited to a molecular mass of about 1000 Da.

Preferably, electrospray ionization (ESI) is used in combination with quadrupole mass filters [16, 17], whereas MALDI is commonly used in combination with the time-of-flight (TOF) analyzer [18]. The relatively simple construction of these two types of analyzers and the resulting price advantage has led to their replacing the traditional magnetic sector instruments as the workhorses of mass spectrometric analysis. A more recent development is the ion-cyclotron-resonance (ICR) analyzer [19] which can be used for both ES- and MALDI-ionization.

### 17.2.3 Analysis of Peptide- and Oligonucleotide-Libraries

In the following sections we present some examples of mass spectrometric analysis that are relevant to combinatorial synthesis. Unless otherwise mentioned, ESI is performed in combination with the quadrupole analyzer, whereas MALDI is used in combination with TOF-instrumentation.

First analyses of combinatorial libraries using ES-MS were performed on peptide libraries by Metzger et al. [4, 20]. These investigations vary from 48-compound libraries to libraries theoretically containing 24 576 components. For small-compound libraries, HPLC-MS and tandem experiments [5] led to an exact library characterization including sequence information and side-product detection. It was found that the ionization yield of each single component in a peptide library is proportional to its concentration. This observation resulted in Gaussian-like ion-distribution patterns (fingerprint spectra) which were also calculated by a software program developed for this purpose [21]. Other library characterizations make use of single ion monitoring (SIM) [22] enabling an exact analysis of compound libraries containing  $10^5$  peptides. The utility of FT-ICR mass spectrometry for the exact mass determination and characterization of a library has already been shown at an early stage of the development of this technique using ES-ionization. High-resolution measurements using FT-ICR enabled the assignment of molecules with isobaric nominal masses from libraries containing between 19 components [23] and about 350 components [24].

A special applicability of MALDI is the single bead analysis of peptides ( $\sim 1000$  Da) with a TOF-analyzer [25] also using different resin linkers [26]. Improved resolution of mass spectra was achieved with the development of reflectron devices and post source decay for a direct structural readout of single compounds [27]. For direct ionization and desorption in the MALDI source, photosensitive linkers were developed [28, 29].

With respect to the limited time which can be invested in order to characterize a large library, the low-resolving analyzer techniques in most cases only provide rough, finger-

print-like qualitative information. For time reasons, an accurate library analysis with SIM can only be used for the characterization of some more or less arbitrarily chosen peptides.

#### 17.2.4 Analysis of Non-Oligomeric Compounds and Libraries

It must be remembered, that the above-mentioned oligomeric compound libraries comprise only a part of the diversity space within combinatorial chemistry. The limited number of coupling procedures and the resulting compounds also restrict the problem area for mass spectrometric analysis. For non-oligomeric structures, the large number of reagents and reaction types available generates a much greater variety of core or scaffold structures and requires more flexible analytical methods. Thus, the applicability of mass spectrometry for an extended analysis of a greater variety of compounds had to be demonstrated.

##### 17.2.4.1 Direct Analysis of Non-Oligomeric Compound Libraries

Direct ES-MS characterization of a combinatorial library has been performed on the example of modified xanthene derivatives [30]. To circumvent the analysis of  $10^4$  to  $10^5$  possible compounds, small representative sublibraries rather than the whole library were synthesized and investigated [31]. The results obtained were then extrapolated to the synthesis of the target library. A xanthene library was also used for ES-FT-ICR-MS investigation [32]. A more recent contribution deals with the calculation of elemental compositions from FTICR mass spectra. However, the mass accuracy required increases exponentially with the mass of the analyte [33]. Therefore, elemental compositions are calculated from the exact measurement of fragment ions performing  $MS^n$ -experiments [34].

Non-oligomeric compound libraries (200–600 Da) have also been investigated by MALDI [35]. Although the linear TOF analyzer is not well-suited to the measurement in the low mass ranges because of inherent problems with resolving power and interfering matrix effects, this contribution demonstrates the possibility of using MALDI in the analysis of combinatorial compounds.

##### 17.2.4.2 Coupling of Separation Techniques with Mass Spectrometry

An additional dimension to library analysis is introduced when mass spectrometry is coupled with common separation techniques, for example liquid chromatography (LC) and capillary electrophoresis (CE). While these couplings are compatible with spray ionization techniques such as ESI or APCI, they more or less exclude the use of the MALDI technique. Several contributions deal with LC-ES-MS [36, 37] and CE-ES-MS-coupling [38]. For molecules with isobaric nominal masses, MS/MS-experiments are performed to confirm the identity of a library component. Separation and analysis of compound mixtures may also be performed by GC-MS [39, 40]. As a supplement to the more common-



ly used HPLC-MS and CZE-MS techniques, GC-MS with its high separation potency and low running costs is the method of choice for hydrophobic or lipidic compounds, but also in cases where the volatility of the analytes can be raised sufficiently by derivatization.

### 17.2.5 'On-Bead' versus 'Off-Bead' Analysis

Parallel to the synthesis of library compounds on solid supports, the direct analysis (*i.e.* 'on-bead') is an attractive concept. Leaving aside the methods of on-bead-infrared-spectroscopy [41] using attenuated total reflection (ATR) and gel-phase NMR-techniques [42, 43], we wish to consider the use of on-bead analysis in mass spectrometry. The requirements of the ionization process restrict on-bead analysis to the MALDI technique. However, also for MALDI-analysis, in so-called direct monitoring studies [44–46], the compounds are cleaved from the bead before the actual ionization. True on-bead analysis under MALDI conditions is only possible with photolytically cleavable linkers, and this technique has been dealt with in several publications [28, 29] in which the compounds are both cleaved and ionized simultaneously within the MALDI source with a single laser shot. The wavelength of the MALDI-laser must however correspond with the wavelength required to cleave the compounds from the resin. Therefore, on-bead analysis represents a method for special analytical problems which is limited to MALDI.

### 17.2.6 High-Throughput-Systems for Sample Analysis and Purification

#### 17.2.6.1 High-Throughput Analysis

The ability of combinatorial chemistry to generate a myriad of compounds in a very short period of time sets demands upon mass spectrometry to produce a maximum of information in a minimum of time. For mass spectrometric analysis, the high-throughput task is related to single compound collections and libraries. While MALDI is in principle also capable of high-throughput operation, this question is best addressed by ES-MS. Using flow-injection systems, both an autosampler or an autosampler-HPLC unit are suitable for direct coupling to the mass spectrometer.

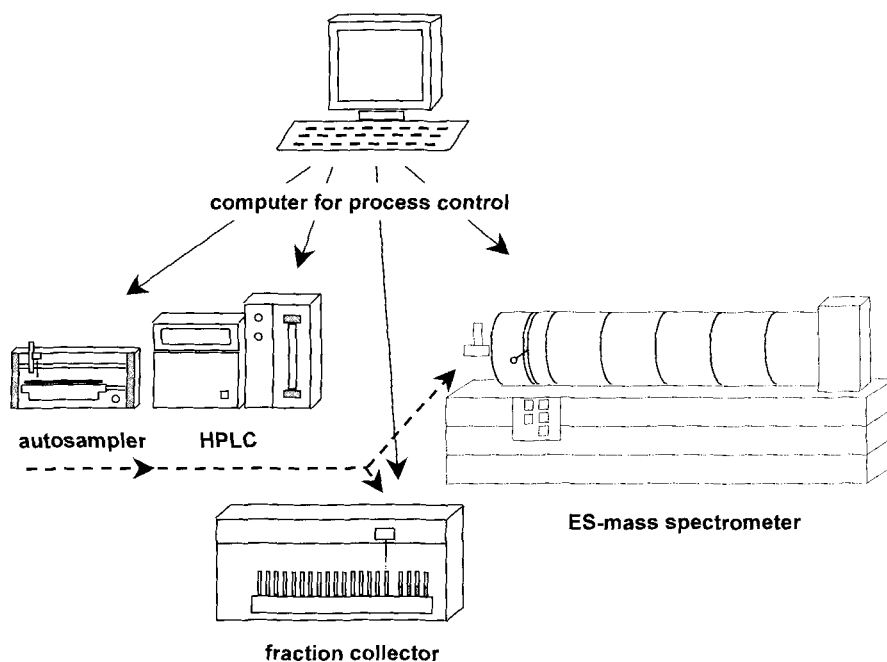
The computer-controlled automation of simultaneous LC-MS and LC-MS/MS analysis extends the range of possibilities in high-throughput analysis. This was demonstrated on the sequencing of peptides from mixtures with computerized sequence interpretation of the fragment ions obtained [47]. The difficulty here however, is the differing collisional energy required for the different substances to obtain optimal decay. Moreover, one must be aware of the vast amount of additional data accumulating from MS/MS experiments. Its employment for high-throughput tasks is only conceivable where software-supported interpretation of spectra is available (see Section 17.3.1.4.). However, even peptide spectra interpretation programs cannot resolve all ambiguities in the sequence determination. At present, the combination of ESI and the quadrupole analyzer technique would appear the most versatile tool for high throughput, particularly with regard to sample preparation, handling and the capability of combination with LC and CE.

Attempts to adapt MALDI-TOF to allow high-throughput analysis [48] have been made, but drawbacks still exist regarding pending sample preparation, matrix effects and the intrinsic inhomogeneity of the sample/matrix crystals combined with the positional resolution of the laser beam. Nevertheless, *e.g.* for oligonucleotides, it has been shown that autosampler arrays can provide rapid and reproducible results [49].

Also, with FT-ICR, automated MALDI measurements have been performed providing acquisition and evaluation of high-resolution spectra in both MS and MS/MS modes at a rate of 20 samples per hour [50]. Accuracy was in the 5 ppm range and the activation energy was held constant. An internal mass reference using the method of multiple ionization techniques simultaneously (MITS) [51] avoids problems of ion discrimination.

### 17.2.6.2 High-Throughput Sample Purification using MS-Detection

At an early stage in the history of combinatorial synthesis it was apparent that product purity in excess of 80% is not always attainable. In fact, purities of 30–40% are quite common for several types of reaction. However, especially for single compounds from compound collections, high-throughput purification systems using HPLC have been developed. The degree of purity is first ascertained by analytical HPLC, and products having a purity below a pre-set threshold are subjected to this fully automated sample purification on a semipreparative or preparative scale. In first technical approaches, samples were



**Figure 17.4.** Automated HPLC-ES-MS sample purification system for preparative yields.

subjected to analytical HPLC with UV-detection alone [52], or also to parallel MS- and UV-detection [53]. The analytical data obtained were later used to guide high-throughput purification by UV-detection alone.

For the high-throughput purification process, ES-MS is an extremely attractive and useful method of detection, complementary to or instead of the commonly used UV-detector. In a second improved approach, the preparative HPLC-system with automated sample injection was directly coupled via stream splitting to an ES mass spectrometer and a fraction collector [54], with all units being guided and controlled by a computer (Fig. 17.4). For each run, the expected molecular mass of the desired compound is entered in a file list of the guiding software program. If the pre-set  $m/z$ -ratio is detected by the mass spectrometer, the fraction collector is guided to collect this fraction, while the rest of the eluant is led directly to the solvent waste. High flow rates (50 ml/min) and short columns provide rapid separations, allowing purification of up to 200 samples per day. With this method also compounds of low yield from well-established reactions or interesting by-products can be isolated. This approach is economically advantageous as the moderate costs associated with a bench-top single quadrupole mass spectrometer are lower than those of developing and performing other, more efficient, syntheses. Although this recent development still suffers from some problems with software control, its further development can be expected to make future use of such a system highly attractive.

### 17.2.6.3 Automated Software-Aided Spectra Evaluation

Effective processing of measured samples in high-throughput analysis can only be achieved with software-aided evaluation and interpretation of spectra. Especially in industrial laboratories the rapid and immediate data exchange required is achieved with the help of local networks [55]. Such concepts comprise both automated sample measurement and automated data processing, including background noise subtraction and averaging of significant spectra. However, for some applications especially in more experimentally oriented laboratories, mass spectrometers in open access mode are used [56]. In such systems manipulation by the user is blocked during phases of automated continuous sample analysis.

## 17.3 Application of ES-MS in Combinatorial Chemistry Analysis

The ability of combinatorial chemistry to generate a myriad of compounds in a very short period of time sets demands upon the analytical techniques to produce a maximum of information in a minimum of time. A prime candidate for the analysis of combinatorially generated compounds is electrospray ionization–mass spectrometry (ES-MS) using a quadrupole mass analyzer. In the following section we wish to examine the practical aspects of ES-MS, and exemplify these on the routine analysis of several combinatorially generated compounds and libraries. Furthermore, attention will be paid to high-throughput and computer-assisted sample evaluation.

### 17.3.1 ES-MS-Analysis of Single Compounds

When working on combinatorial synthesis, differences in analyzing combinatorial compounds rather than peptides or oligonucleotides become apparent. These arise from the greater variety of reagents used and different reaction types which require an adaptation of several reaction parameters, *e.g.*, reaction time, temperature, etc. A further reason is the development of a huge number of different resin types and linkers tailored for combinatorial synthesis, but which were not necessary in peptide chemistry.

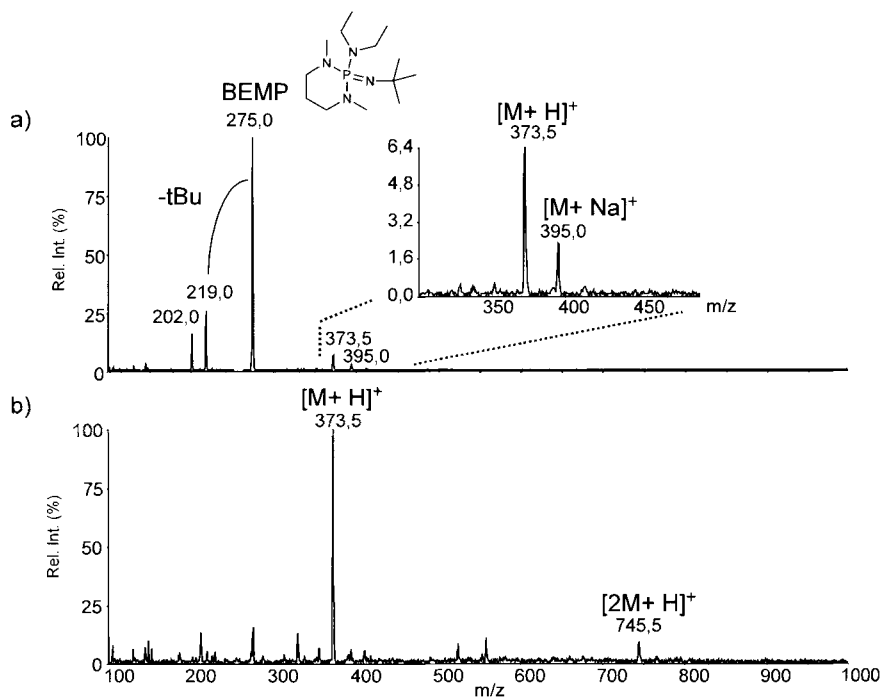
#### 17.3.1.1 Influence of Resin Types on MS-Analysis

A problem which arises after cleavage not only from polystyrene-based polymeric carriers is the presence of impurities derived presumably from linker or resin components [57]. The occurrence of these impurities hampers routine product analysis and may have considerable influence also on the screening assay. In the mass spectrometric analysis, they can compete with poorly ionizable products, thus inhibiting a reliable analysis. At worst, they may even preclude detection of the expected compound, leading to a misinterpretation of the synthesis results. It has been found that the extraneous signals in the mass spectra are primarily dependent upon the type of resin used. Differences between analogous resins from different suppliers, or even different charges are also observed. As yet, it has not been possible to characterize these impurities with the methods of mass spectrometry and NMR.

#### 17.3.1.2 Influence of Reagents on MS-Analysis

In combinatorial synthesis the variety of reagents used has a considerable influence on the mass spectrometric analysis, as demonstrated in the following examples. Figure 17.5 shows mass spectra of two analogous synthesis products which differ only in the base used (Fig. 17.5(a), BEMP; Fig. 17.5(b), KOtBu). The base peak in the upper spectrum arises from BEMP, despite the fact that the resin was extensively rinsed and that the presence of phosphazene base is not apparent in the HPLC chromatogram. This reflects the extremely high ionization yield of BEMP. In addition, fragment ions of BEMP (*e.g.*, 219 *m/z* arising from the loss of tBu) are also observed. Although the relative intensity of the molecular ion of the desired product is only about 10% that of the base peak arising from BEMP, the purity of this product is in fact 90%. Such readily ionized reagents, although present in low concentrations, can easily give a misleading impression as to the true sample composition. In some cases the ionization yield of the product may be so low that it cannot be detected at all. Therefore, if the analyzed compounds bear protonatable as well as deprotonatable functional groups, it is advisable routinely to perform the MS analysis in both the positive and negative scan modes.

It is well known from biological isolates, that buffer additives, *e.g.* borate, can completely inhibit the detection of ions. Another effect is clustering of alkali metal ions with the



**Figure 17.5.** ESI mass spectra of a reaction on solid support using two different bases. (a) BEMP; (b) KOtBu.

analyzed sample molecules. Clustering also occurs with bases in positive ion mode or acids in the negative mode. This becomes apparent through the presence of spurious signals which may be misinterpreted as side reactions. These effects can often be observed with TFA in the negative ion mode. The dominant signals at  $m/z$  113 and 227 arise from traces of the cleavage reagent TFA still present after lyophilization of the sample. For measurements in the negative ion mode, the use of TFA in the solvent system should be avoided. It is preferable to use formic acid or acetic acid since with these acids the yield of ionization in the negative mode is maintained. Some of the reagents leading to mass spectrometric complications are summarized in Table 17.1.

**Table 17.1.** Examples of reagents and buffer additives disturbing sample detection.

Reagents	Positive mode	Negative mode
Bases	DBU/phosphazenebases/...	–
Buffer additives	TEA ( $m/z$ 101) borate/phosphate $\text{NH}_4^+$ ; $\text{Na}^+$ ; $\text{K}^+$ -salts	TFA ( $m/z$ 113/227) borate/phosphate

### 17.3.1.3 Influence of Fragmentation on MS-Analysis

Although ES-MS represents a very soft ionization technique, with the molecular ion being generally very prominent, molecules of low molecular mass show a tendency to fragment. These fragments can be of help in determining molecular structure. Table 17.2 shows examples of frequently occurring typical fragmentations in ES-MS. The amount of fragmentation can vary within the compound class, and depends on the orifice voltage and the scan mode (positive or negative ions). In cases where insufficient fragmentation occurs for conclusive elucidation of the molecular structure of a substance, additional fragments can be generated. However, this technique requires either pure samples for collision-induced decay (CID) in the ionization region or a multiple stage mass spectrometer for CID with argon in the collision quadrupole. MS/MS-measurements cannot be easily automated, as time is required to adjust mass spectrometric parameters in order to optimize fragmentation.

**Table 17.2.** Examples of frequently occurring fragmentations in ESI-MS.

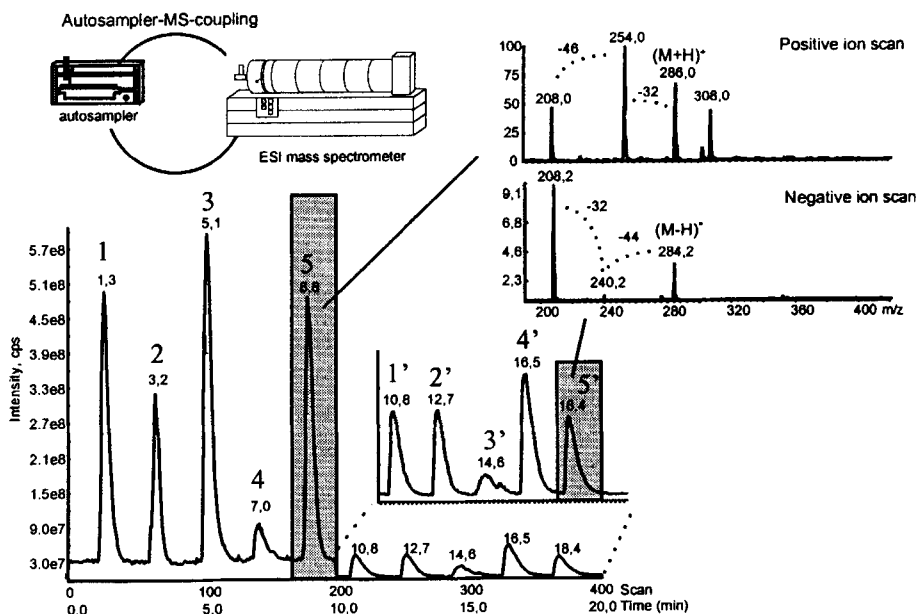
Compound class	Molecule loss	Mass loss
Carboxylic acid amides	NH <sub>3</sub> ; NCHO	- 17/ - 43
Carboxylic acids	OH; CO <sub>2</sub>	- 17/ - 44
Alcohols	OH	- 17
Ethers (e.g., methoxy/ethoxy-group)	HOCH <sub>3</sub> /HOCH <sub>2</sub> CH <sub>3</sub>	- 32/ - 46
Amines	NH <sub>3</sub>	- 17

### 17.3.1.4 High-Throughput Analysis

A typical ES-MS autosampling procedure with five samples is shown in Figure 17.6. The *x*- and *y*-axes represent time and total ion intensity respectively. After injection of a sample, the ion current increases above that arising from the solvent alone. Spectra can then be extracted from the ion chromatogram. Sample injection and detection can be performed in 2-minute-intervals or even less, thus allowing about 300 samples to be analyzed in 10 hours.

In this example, samples 1–5 were recorded in positive ion scan mode, while samples 1'–5' are the related samples recorded in the negative ion scan mode. Although the concentration is the same in each sample, different ionization yields are observed in both scan modes. Ionization in negative scan mode is generally weaker by a factor of about ten. Sample 4 cannot be detected in the positive mode, despite the small increase of ion current which results from artifacts introduced during cleavage from the resin. On the other hand, sample 3' shows no ionization in negative mode. The other samples are ionizable in both modes, indicating the presence of both protonatable and deprotonatable functional groups in these compounds.

The use of both positive and negative scan modes allows detection of both protonatable and deprotonatable molecules. Not in all cases do the synthesized molecules contain



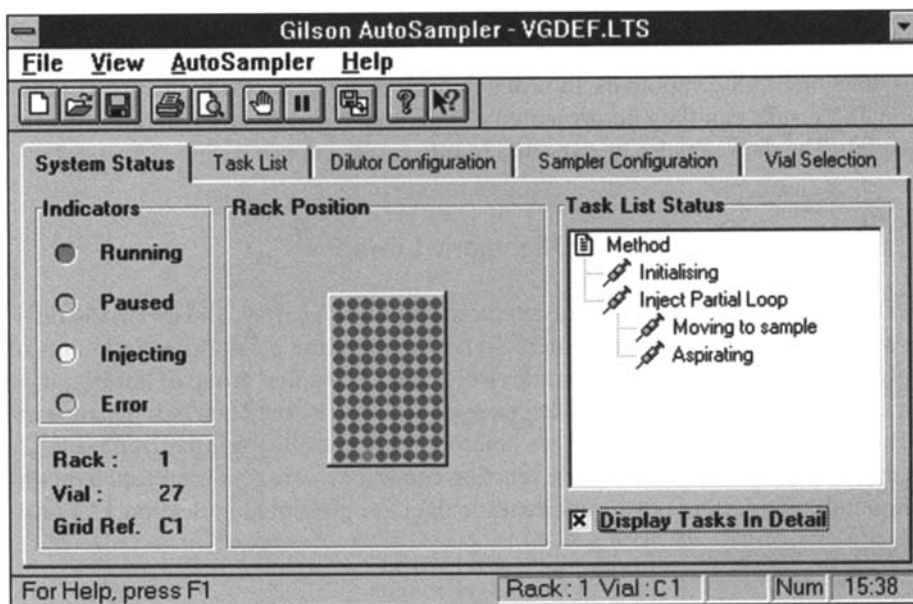
**Figure 17.6.** Total-ion chromatogram of an autosampler run in positive ion mode (samples 1–5) and in negative ion mode (samples 1'–5').

protonatable functional groups. In the case of solid-phase organic synthesis (SPOS) using *e.g.*, Wang-resins, it is more likely to obtain deprotonatable molecules. Although the chemist can easily decide on the basis of the molecule structure which mode has to be chosen, in most cases it is advisable to use both scan modes.

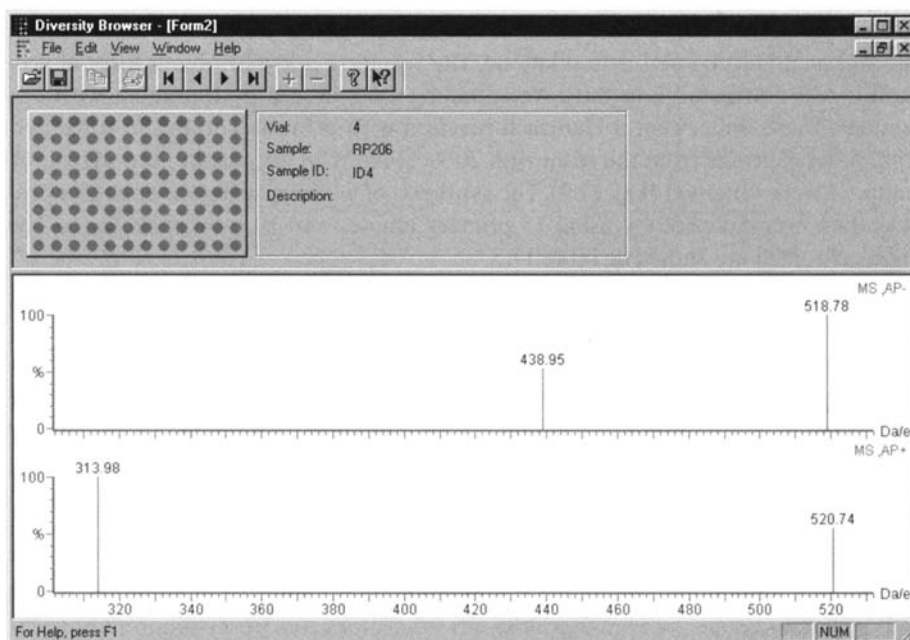
#### 17.3.1.4.1 Automated Software-Aided Data Evaluation

In addition to automated sample injection, considerable effort has been directed towards the development of automated sample evaluation. The final goal is a set of programs and file formats which accompany a collection of compounds from the automated synthesis via the automated analysis to the final evaluation of a bioassay. In order to minimize time losses incurred by the conversion of files, file formats and software programs must be fully compatible. Data sets should contain complete information about each single compound (structure, HPLC/MS analysis, biological activity, etc.).

As an example of such a program, Figure 17.7 shows the status page of an autosampler, which informs the operator about the current status of the system. The rack number, vial and grid are reported. At the same time, the microtiter plate indicates the hits and misses of the expected molecular masses by different coloring of the rack positions. During the analysis, new sample lists can be created, files imported from other sources, methods checked and results verified. Mass spectra can be acquired (Fig. 17.8) by browsing the schematically displayed microtiter plate with the mouse. In order to obtain comprehen-



**Figure 17.7.** Status page of an autosampler-MS guiding software program. (Reproduced with permission from Gilson-Micromass.)



**Figure 17.8.** Data page for browsing all mass spectra of an autosampler run. (Reproduced with permission from Gilson-Micromass.)



sive mass spectrometric information and to avoid failures, all samples are measured in the negative and positive scan modes. A value for the intensity threshold is pre-set so as to ensure detection of all compounds. In order to overcome the necessity of spectra interpretation, the results can then be presented as ‘hit’ and ‘non-hit’-columns reporting the synthetic hits. The data can also be exported to other software programs.

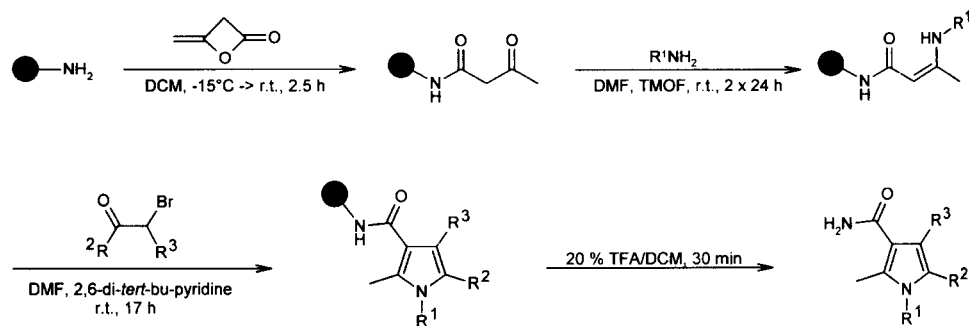
### 17.3.2 ES-MS-Analysis of Non-Oligomeric Libraries

In the following section we focus our attention on library analysis, and especially on libraries which are not related to oligomers. To demonstrate the possibilities and limits of this analysis, two typical compound libraries were chosen. The first group of libraries contains an aromatic scaffold, pyrroles, which were synthesized by the Hantzsch pyrrole synthesis. The second class of compounds are heterocyclic isoxazolines synthesized via a 1,3-dipolar cyclo-addition. In both cases the reaction conditions were first established on single compounds. Supporting mass spectrometric data are presented in Section 17.7 (Appendix).

#### 17.3.2.1 ES-MS-Analysis of a Pyrrole-Library

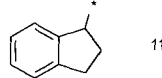



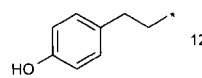

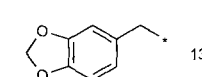
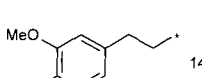
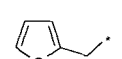
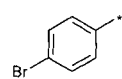
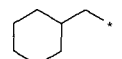
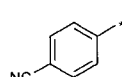
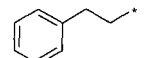
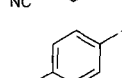
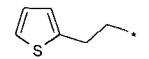
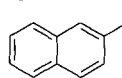

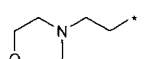
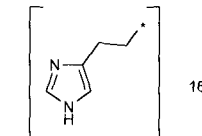
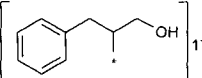
##### 17.3.2.1.1 Synthesis Aspects

A Hantzsch pyrrole synthesis was performed on PS Rink Amide resin. In the first step, the resin was acetonacetylated with diketene. The resulting polymer-bound form of acetoacetamide was converted into the corresponding enaminones by treatment with primary amines. These underwent a Hantzsch reaction with  $\alpha$ -bromoketones to form pyrroles [58]. After cleavage from the resin with 20 % TFA/DCM, the substituted pyrrole-3-carboxamides were obtained (Fig. 17.9). The synthesis of a library using the split-mix protocol was then accomplished by using 15 primary amines and 10  $\alpha$ -bromoketones. The building blocks used are shown in Table 17.3.



**Figure 17.9.** Hantzsch pyrrole synthesis on the solid phase.

**Table 17.3.** Building blocks for the pyrrole library synthesis.

R <sup>1</sup>	R <sup>2</sup>	R <sup>3</sup>	Name of sublibraries
H—*		H—*	PYL-1
		H—*	PYL-2
		H—*	PYL-3
		H—*	PYL-4
HO—CH <sub>2</sub> —CH <sub>2</sub> —CH <sub>2</sub> —*		H—*	PYL-5
		H—*	PYL-6
		H—*	PYL-7
		H—*	PYL-8
			PYL-9
		H—*	PYL-10
			

## 17.3.2.1.2 Direct ES-Investigation

The interpretation of mass spectra of compound libraries obtained from direct injection measurements is problematic for two main reasons: (i) The yield of ionization of the various components of a library can vary significantly depending upon the structure of the side chain and poorly ionized members are therefore easily overlooked; and (ii) even with 'soft' ionization techniques, fragmentation is possible and the additional unexpected peaks impede the interpretation of the mass spectrum.

This is exemplified on the directly measured library PYL-1 (Fig. 17.10), where the signals of the compounds Py-6 and Py-16 are absent, presumably due to poor ionization yields. Also, compounds Py-13 and Py-17 have the same nominal molecular mass and therefore cannot be differentiated by MS. The fragmentations  $[M + H - 17]^+$  and  $[M + H - 43]^+$  of the library molecules lead to a multiplicity of signals complicating the transparency of the mass spectrum. As an example, the signal at  $m/z$  249 can be assigned either to compound

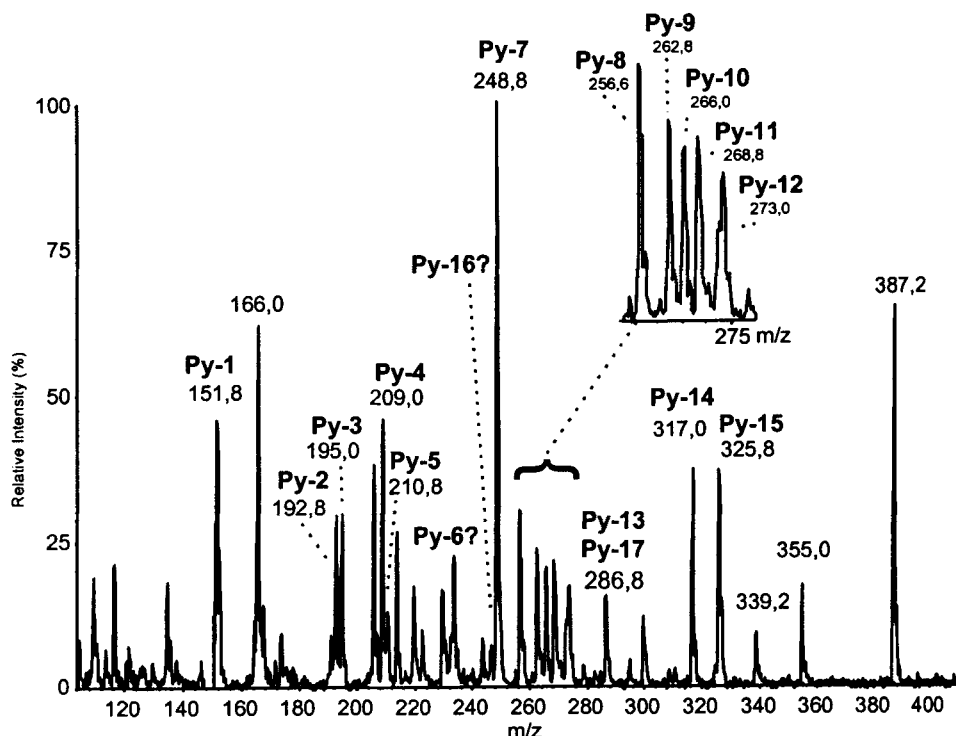


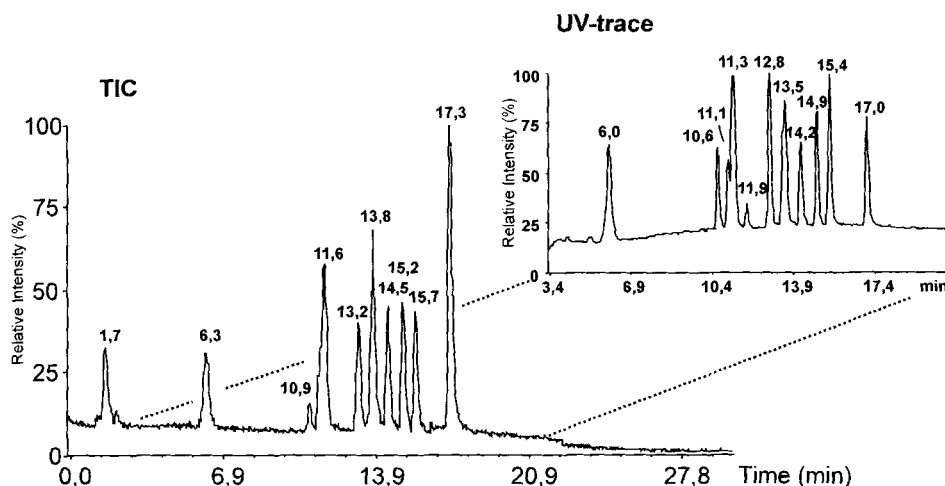
Figure 17.10. Direct ES-MS-spectrum of the crude PYL-1-library.

Py-7 ( $[M+H]^+ = 249$ ) or to the fragment of Py-10 arising from the characteristic loss of ammonia  $[M+H-17]^+$ . Some major signals, *e.g.*,  $m/z$  387 cannot be assigned at all.

For these reasons it is not advisable to analyze small-compound libraries by direct injection of the library. On the other hand, the mass spectra of libraries containing several hundreds to thousands of compounds tend to show the expected gaussian-like mass distribution around the average molecular mass. This kind of 'fingerprinting', which was advanced especially for the analysis of large-peptide libraries, may be accepted for synthesis control of a library, but the evaluation of mass spectra yields only qualitative information. Exact assignment of the compounds is impossible.

### 17.3.2.1.3 HPLC-ES-MS-Analysis

To attain maximum resolution of the components of a mixture by HPLC, the gradient of eluent must be effectively optimized. This procedure can be tedious and time consuming. The rate of sample generation in combinatorial chemistry makes it impossible to find time to optimize the chromatographic conditions for each sample. Fortunately, detection by mass spectrometry alleviates somewhat the need for all compounds to be resolved, since unresolved samples can usually be differentiated on the basis of their molecular

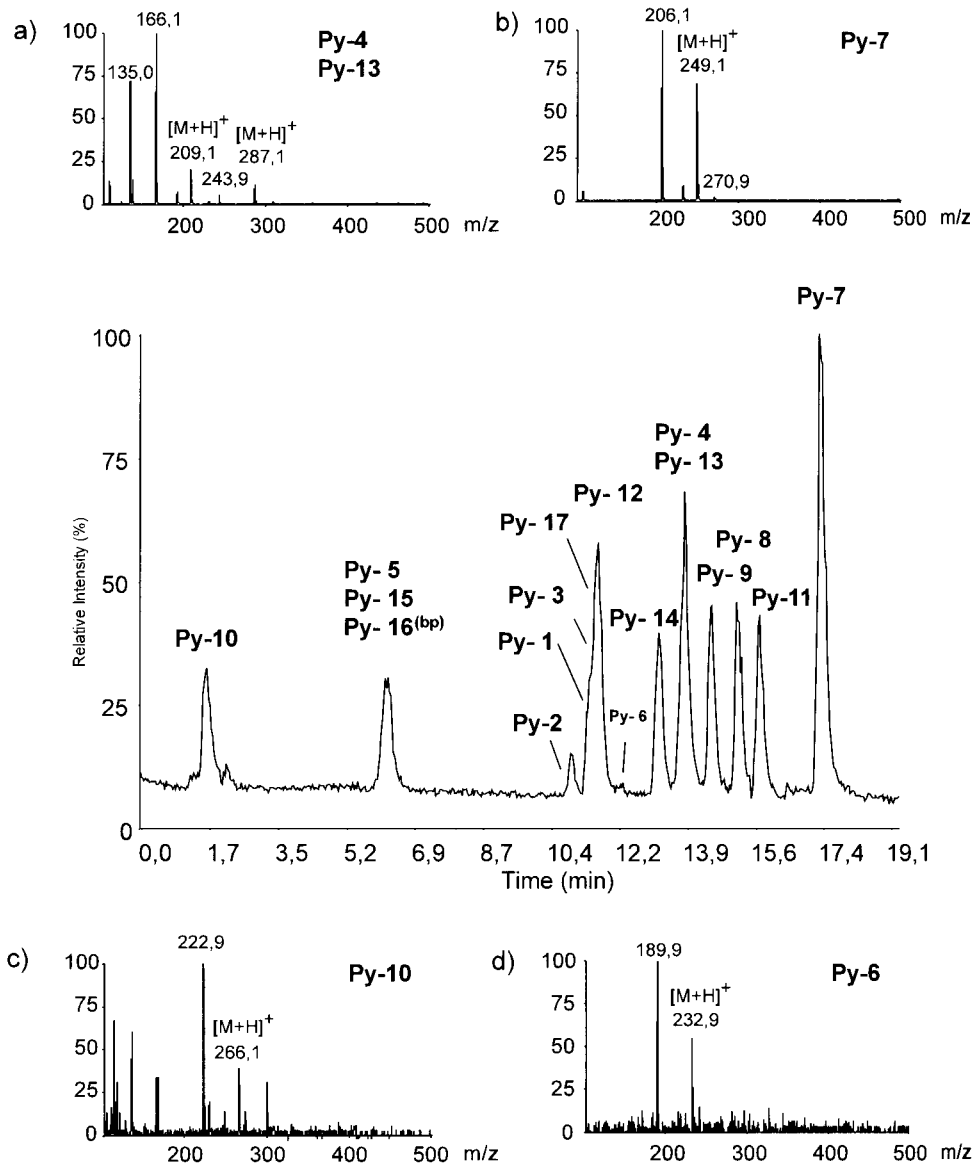


**Figure 17.11.** HPLC-ES-MS of PYL-1. (a) Reconstructed total ion current (TIC); (b) UV-trace.

mass. For these reasons, a standard gradient was used for all HPLC-MS reported here. Figure 17.11 shows the total ion current (TIC) of the pyrrole library (PYL-1). Ten, rather than the expected 17, peaks appear, indicating either peak overlap, poor ionization or absence of some compounds. The substances are seen to elute over a wide range of eluent polarity, with the most polar compound eluting at 1.7 min and the least polar compound at 17.3 min. Ionization yields of the molecules differ widely.

As expected, the UV-trace and the TIC differ greatly with respect to the signal intensity of the individual compounds on account of the different principles of detection. This is apparent for example for the signal at 10.6 min (UV-detection) and the analogous signal in the TIC at 10.9 min. (The dead time of the coupling between UV detector and the MS accounts for the constant shift in retention time of 0.3 min.) Diffusive and convective effects also lead to peak broadening: The two separated peaks in the UV-trace (11.1 min/11.3 min) overlap significantly in the TIC (11.6 min). Upon retrieval of the mass spectra from the TIC, 16 of the expected 17 compounds can be assigned on the basis of their molecular masses (Fig. 17.12).

Although some components co-elute, their assignment was nonetheless possible as a result of the formation of characteristic fragments  $[M + H - 17]^+$  by the loss of ammonia and  $[M + H - 43]^+$  by the loss of NCHO (spectra of Fig. 17.12(a) and (b)). Residues with amino groups or heterocycles (e.g., pyridine) often lead to short retention times, as the spectrum in Figure 17.12(c) of compound Py-10 shows. Compound Py-10 elutes almost together with the injection peak. Because the ionization yield of compound Py-6 is very poor, the peak is easily overseen in the TIC-trace. The complementary information of the UV-chromatogram (11.9 min) however indicated where to search for the corresponding weak mass spectrum, which indeed is found (spectrum of Fig. 17.12(d)). Experience indicates that compounds bearing a furane residue, such as Py-6, always show poor ionization yields in ES-MS. Compound Py-16 was not found in either chromatogram.



**Figure 17.12.** HPLC-ES-MS-chromatogram of the library PYL-1 and the assignment of library compounds. Spectrum a) Co-elution of Py-4 and Py-13. Spectrum b) Py-7 (highest ionization yield). Spectrum c) Py-10 (co-elution with the injection peak). Spectrum d) Py-6 (lowest ionization yield).

Compounds with isobaric nominal masses can pose a problem in the analysis of libraries. For example, compounds Py-13 and Py-17 of library PYL-1 (see Section 17.7, Appendix) have the same nominal molecular mass. Using a low-resolution quadrupole mass spectrometer, there is no chance of differentiating between these compounds in the

chromatogram. A deconvolution of the library to the single compound was performed. At least Py-13 shows an characteristic fragment of  $m/z$  135 which allows an unambiguous assignment of these pyrroles in the chromatogram. Problems with an assignment of the compounds with the same molecular mass can be overcome by design of the library prior to synthesis.

#### 17.3.2.1.4 Detection of Side Reactions

As mentioned in Section 17.3.2.1.3, compound Py-16 was not found in the chromatogram. Instead, an additional compound was observed at 6.2 min which co-eluted with Py-5 and Py-15. As the yield of the reaction is in general high, it is reasonable to suspect that Py-16 is converted to another compound in a side reaction. Synthesis of Py-16 as a single compound confirmed the occurrence of a side reaction (Fig. 17.14). The nucleophilicity of the aromatic imidazole-nitrogens promoted their complete alkylation. Figure 17.13 shows the proposed mechanism for this side reaction which yields an organic salt.

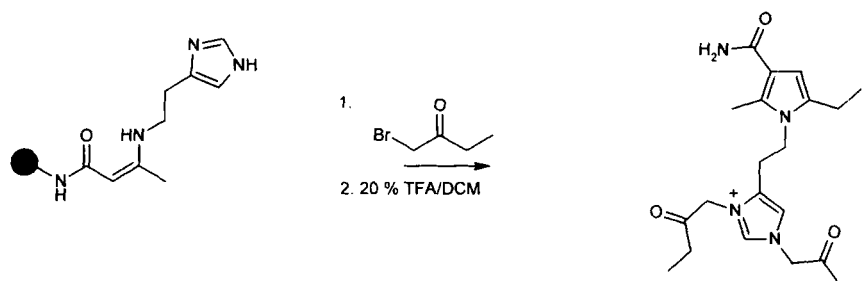


Figure 17.13. Mechanism of the side reaction.

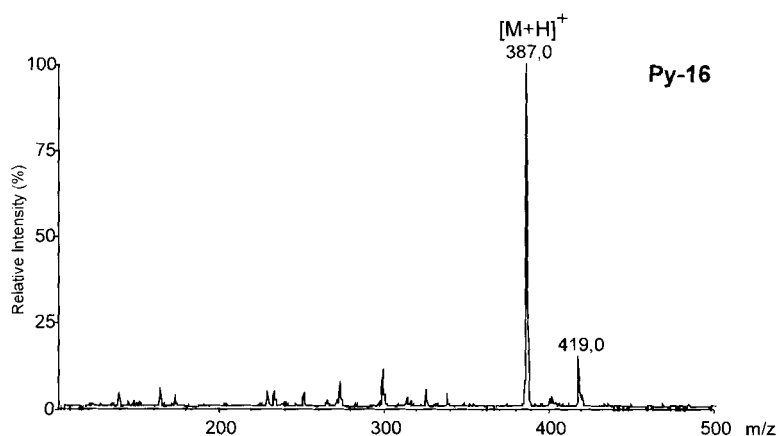
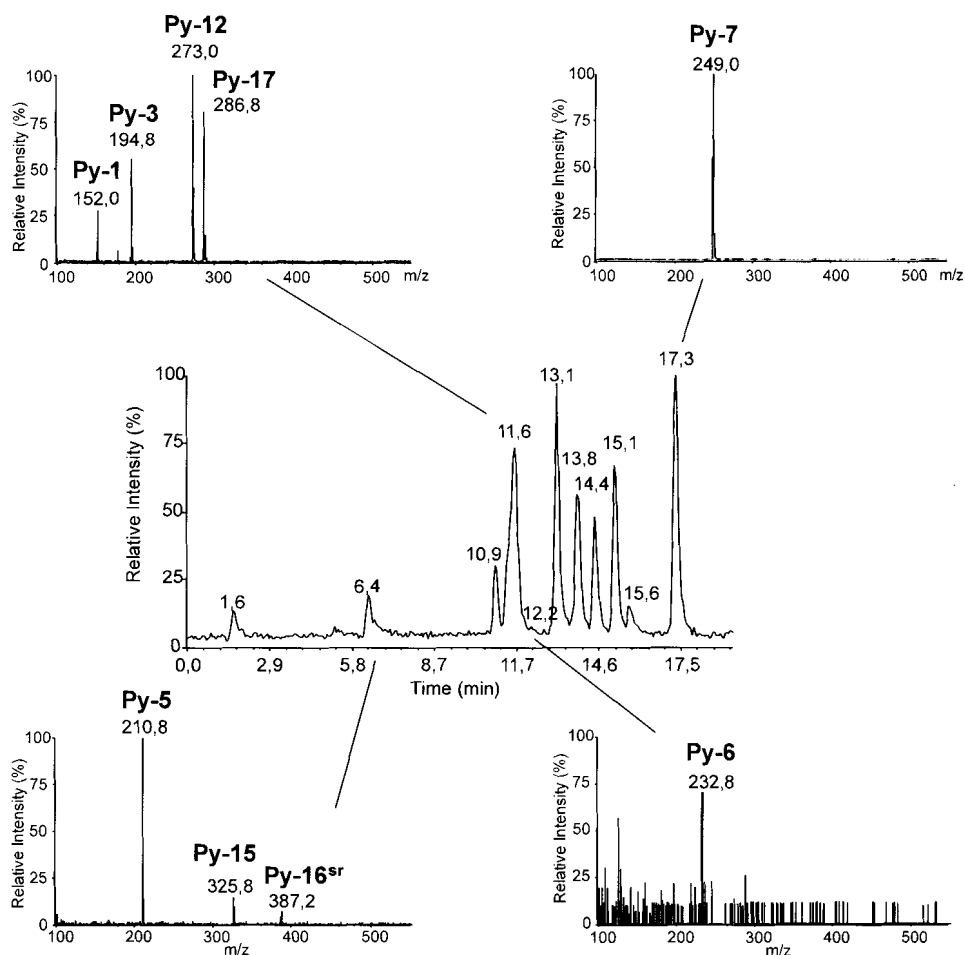


Figure 17.14. Mass spectrum of Py-16 after deconvolution of the library PYL-1.

The occurrence of this side reaction was observed with other educt components (*e.g.*, aminomethylenepyridine), and for this reason all analogous building blocks were subsequently excluded from synthesis.

### 17.3.2.1.5 Tandem Experiments

Use of a triple quadrupole mass spectrometer with ES-source opens up additional possibilities for the characterization of libraries. The **parent ion** scan can be used if all compounds of a library form the same fragment ion. In addition, the **neutral loss** experiment allows the detection of all molecules undergoing the same mass loss. In both cases the adequate fragmentation of library compounds must be ensured by the appropriate choice of



**Figure 17.15.** Characterization of library PYL-1 by the neutral loss experiment ( $\Delta m = -43$  amu).

collision energy, even of poorly ionized molecules. With this technique, a very high sensitivity, even for molecules which are scarcely ionized, can be achieved. Because all components of the pyrrole library form the characteristic fragments  $[M + H - 17]^+$  and  $[M + H - 43]^+$ , the neutral loss experiment (mass loss of 43 amu) was used for library characterization (Fig. 17.15).

#### 17.3.2.1.6 GC-EI-MS-Analysis

The volatility of a compound, governed primarily by its molecular mass and its polarity, is the main limiting factor for analysis by gas chromatography (GC). The fact that synthesis of combinatorial libraries on solid support very often leads to highly polar compounds after cleavage from the resin makes the use of GC appear unattractive, at least at a first glance. The advances in deactivation of the capillaries used in GC and the possibility of derivatization of the polar functional groups however render analysis by GC-MS feasible.

The underivatized pyrrole library PYL-1 was investigated by GC-MS using electron impact ionization. The chromatogram (Fig. 17.16) shows near-baseline separation for all compounds. Apart from Py-1 and the by-product of Py-16, all other 15 compounds could be assigned. In all cases the molecular ion as well as fragments suitable for structural analysis were detected. Compound Py-1 was not detected, probably due to its polarity or perhaps decomposition under the chosen column conditions. As discussed in Section 17.3.2.1.4, the by-product of Py-16 is an organic salt and is therefore not accessible for GC analysis. Compounds with alcoholic groups (Py-5 and Py-17) and phenolic groups (Py-12) show only poor ionization yields.

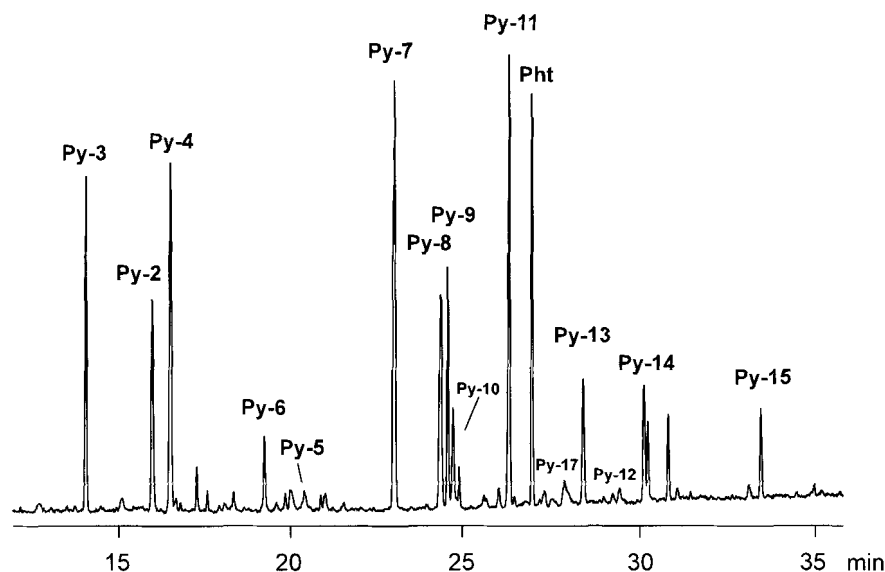


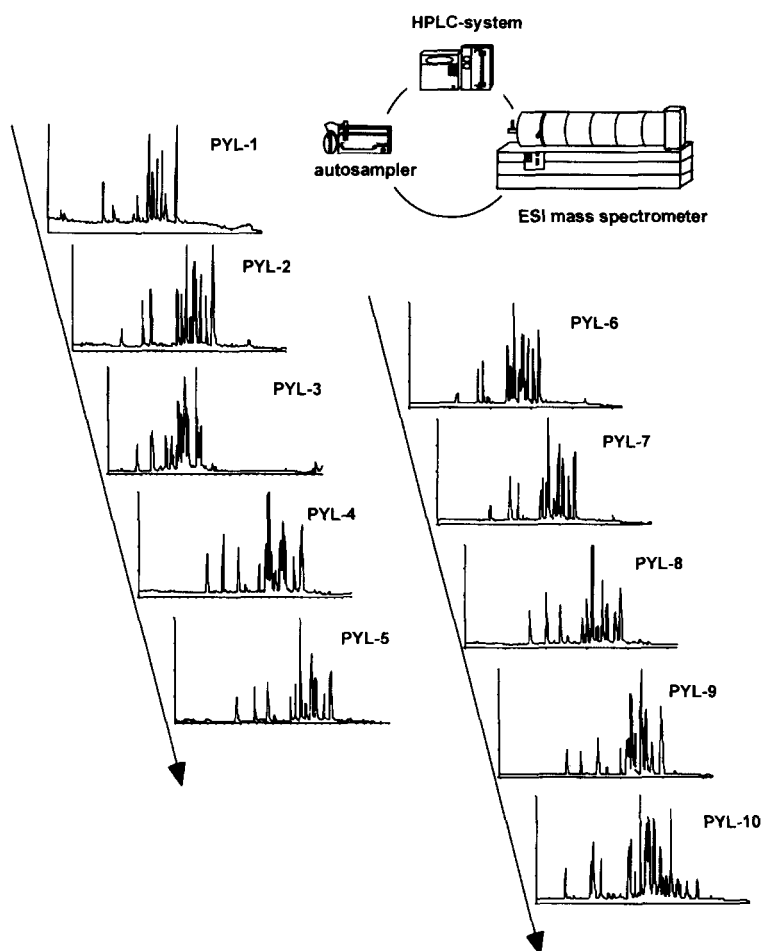
Figure 17.16. GC-EI-MS chromatogram of PYL-1.



Trimethylsilylation of the library with BSTFA (data not shown) leads to the corresponding trimethylsilylethers of Py-5, Py-12 and Py-17, the ionization yields of which are significantly higher than the underivatized parent compounds. Despite derivatization it was not possible to detect compound Py-1. Diisooctylphthalate (Pht), a common impurity in reagents and solvents stored in contact with plastic materials is also detected.

#### 17.3.2.1.7 Automated HPLC-MS-Analysis

Automation of such library analyses is possible, with suitable equipment existing for GC, HPLC and for CZE separation methods. However, the vast number of samples generated with this new synthetic principle require faster analysis. A high throughput for



**Figure 17.17.** Automated HPLC-MS-analysis of ten pyrrole libraries (PYL-1–PYL-10).

coupling of separation techniques with mass spectrometry is achieved through speeding up of separation methods as there are, *e.g.*, short HPLC columns, which allows sample analysis within a few minutes. Figure 17.17 shows the automated HPLC-MS library analysis in 30-minute intervals using a narrow-bore column. Depending on the polarity of the side chain introduced in R<sup>2</sup> and R<sup>3</sup>, the signal entities of compounds shift to a more polar gradient (PYL-3) or to a non-polar gradient (PYL-9). Evaluation of the TICs show, that despite library PYL-10 all ten libraries contain the expected compounds in high purity.

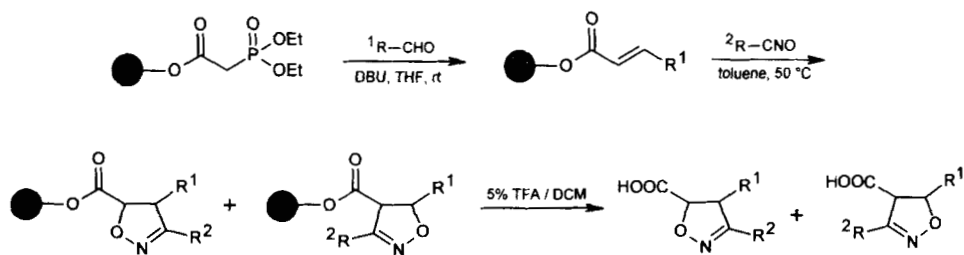
### 17.3.2.2 MS-Analysis of an Isoxazoline Library

#### 17.3.2.2.1 Synthesis

Synthesis of the isoxazoline library [59] was performed on 2-chlorotritylchloride resin. The resin was loaded with diethylphosphonoacetic acid and the polymer-bound phosphonate reacted with aldehydes to yield substituted E-cinnamic esters or substituted E-acrylic esters. This was followed by a 1,3-dipolar cyclo-addition with nitrileoxides, synthesized via the method of Mukaiyama and Hoshinoc [60] (Fig. 17.18). In this reaction, two regioisomers are formed, each of which exist in two enantiomeric forms. Table 17.4 shows the building blocks used in the library synthesis.

**Table 17.4.** Building blocks for the isoxazoline library synthesis.

R <sup>1</sup>	R <sup>2</sup>	Name of sublibraries
1.            5.            9.            13.		ISO-1
2.            6.            10.            14.		ISO-2
3.            7.            11.            15.		ISO-3
4.            8.            12.		ISO-4

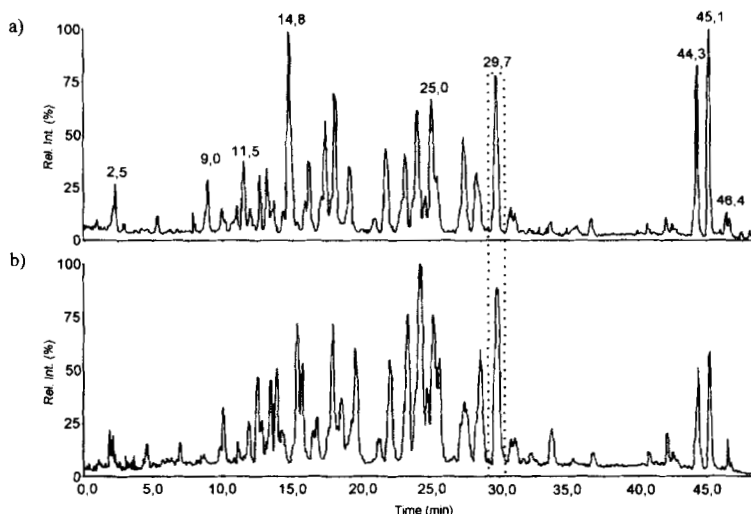


**Figure 17.18.** Isoxazoline synthesis via a 1,3-dipolar cyclo-addition.

### 17.3.2.2.2 HPLC-Analysis in Positive and Negative Ion Scan Mode

For the analysis of both single compounds and compound libraries, it is advisable, as mentioned above, to employ both positive and negative scan modes. For the case of HPLC-MS-analysis this requires that TFA, widely used as a solvent additive, be replaced by another acid, *e.g.*, formic acid. Experience has shown that formic acid is suitable for the majority of separations. In some cases, minor peak broadening is observed in comparison to TFA.

The TICs of the compound library ISO-1 in both positive and negative scan modes are shown in Figure 17.19. The presence of an ester and a carboxylic acid group as functional groups in all components of this library make them accessible to detection in both positive and negative scan mode. To facilitate the assignment of the compounds, the gradient



**Figure 17.19.** Reconstructed total ion currents of library ISO-4. (a) Positive scan mode; (b) negative scan mode.

was optimized (see Section 17.7; Appendix). Because of the possibility of stereoisomers, at least 32 compounds are expected to be present in the library.

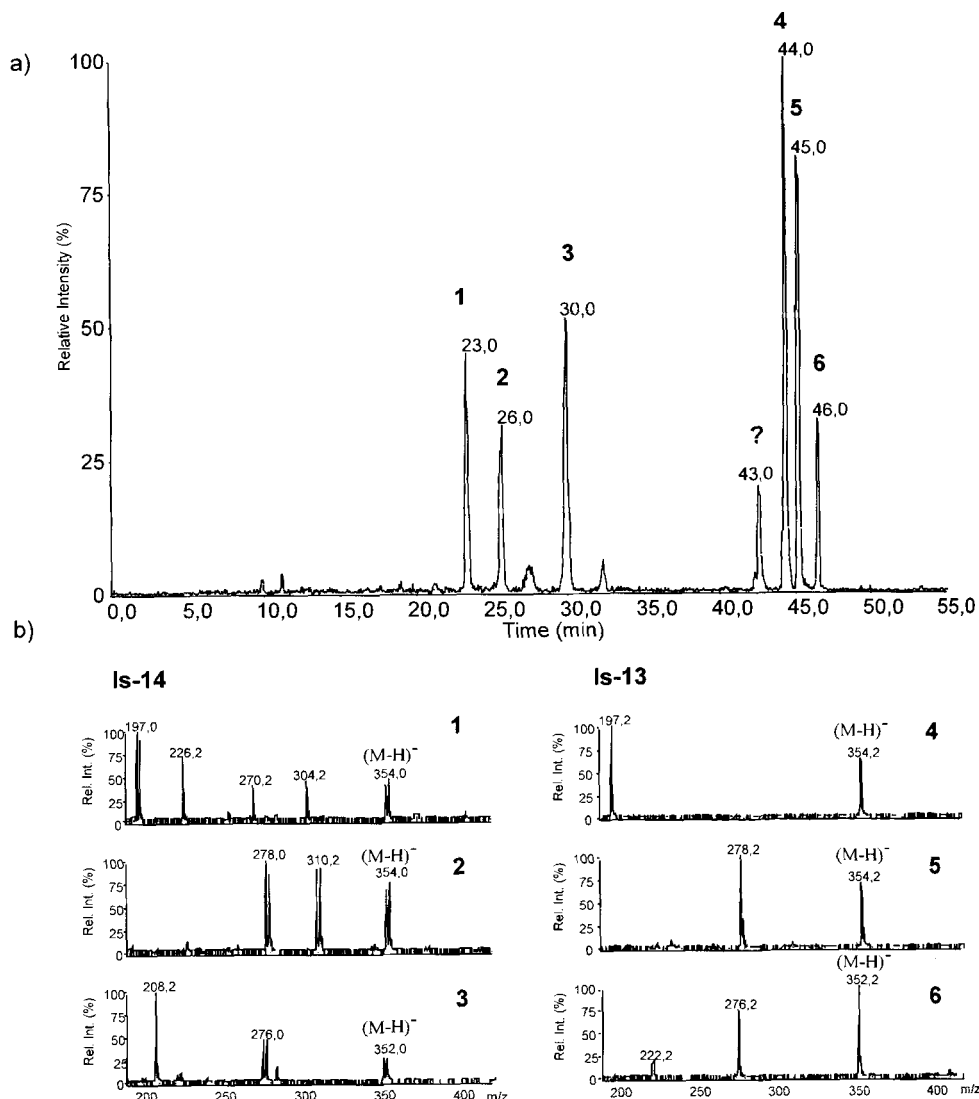
The mass spectrometric evaluation of the TIC revealed the presence of by-products. These became particularly apparent when a sample which had been stored for a week at 4°C was reinvestigated. Compounds with molecular masses corresponding to the loss of 2 amu from the original library components were detected, presumably by the loss of hydrogen from the isoxazolines to form the aromatic isoxazoles with their enlarged conjugated  $\pi$ -electron system. This observation of the formation of by-products upon storage indicates that the conditions of storage should be carefully considered, and that a second analysis should be undertaken before compounds which have been stored are subjected to a bioassay.

#### 17.3.2.2.3 Evaluation Methods for Library Analysis

The manual investigation and interpretation of such complicated chromatograms is a very time-consuming enterprise. To facilitate the evaluation of HPLC-chromatograms there are several software-aided methods available, the two main methods being the 'extract-ion'-function and the 'landscape'-plot.

##### Extract-Ion-Function

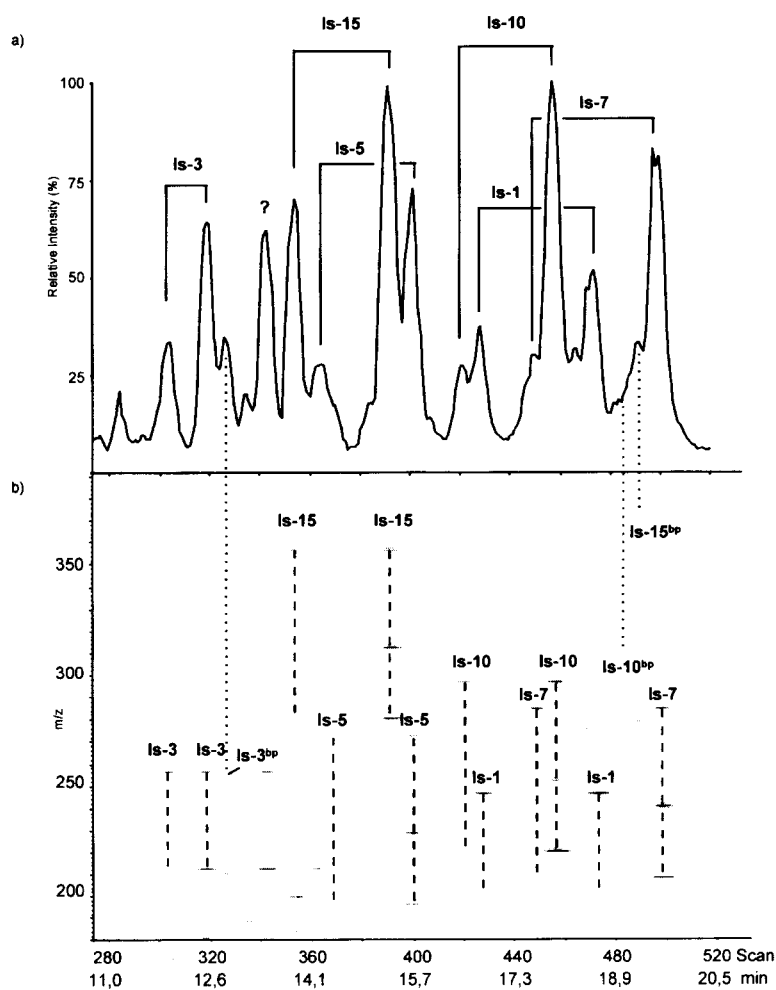
With the extract-ion-function, an ion-specific chromatogram is constructed by plotting the intensity of a single selected ion or a group of ions, extracted from the series of mass spectra, against time. The search for the expected masses is much faster than the evaluation of each signal in the TIC. In addition to the molecular ions, characteristic fragments can be used to further check the presence of a molecule in a library. Figure 17.20(a) shows the mass trace for  $m/z$  352–355 extracted from the TIC in the negative ion mode of Figure 17.19. The two groups of signals are indicative of the two library compounds Is-13 and Is-14. On account of their isotopic pattern, the mass spectra 1 and 2 can be assigned to the brominated isoxazoline Is-14, spectra 4 and 5 to the isoxazoline Is-13 (Fig. 17.20(b)). Additional signals in spectrum 1 are due to peak overlapping with other library compounds. Spectra 3 and 6 represent the oxidized isoxazoles, which elute later than the isoxazolines. This increase in retention time of the isoxazoles relative to their corresponding isoxazolines was observed for all library compounds. Characteristic differences in fragmentation can be observed in the spectra for the different isomers. At the same time, the limitation of mass spectrometry becomes apparent. Because for the oxidized by-products there is only one peak in the HPLC-profile, it could not be decided whether they are oxidation products of only one regioisomer (derived from the compounds out of spectra 1 and 4 or 2 and 5 respectively) or co-eluting oxidation products of both regioisomers. These problems require further NMR-studies for unambiguous structure elucidation.



**Figure 17.20.** (a) Mass trace of  $m/z$  352–355 of library ISO-4 using the ‘extract-ion’-function in negative ion mode. (b) Spectra and compound assignment of Is-14 and Is-13.

### Landscape-Plot

A major disadvantage of the ‘extracted’ ion traces is the possibility of overlooking unknown by-products since only the expected masses are monitored. An excellent overview of all masses is given by the 2-dimensional landscape-plot with the retention times on the  $x$ -axis, and the molecular masses on the  $y$ -axis. Figure 17.21(a) and (b) show the TIC and the related landscape-plot respectively of a part of Figure 17.19(b). The common problem



**Figure 17.21.** Part of the HPLC-MS of library ISO-4 in negative-ion mode. (a) Reconstructed total-ion-chromatogram. (b) Two-dimensional contour-plot and compound assignment (vertical dashed lines).

in HPLC-analysis of peak overlapping can be visualized more transparently in the landscape-plot. The assignment of compounds is indicated by paired vertical lines with the corresponding regioisomers. Different fragmentations can be observed for each regioisomer. Moreover, the isoxazole by-products (bp) are shown for Is-3, Is-10 and Is-15. The analysis of the landscape-plot enables the analysis of all library compounds and their assignment in the TIC. Additional signals also arise from unidentified by-products.

## 17.4 Conclusions

The developments in combinatorial chemistry represent a challenge for the traditional mass spectrometric analytical techniques. For the majority of single compounds, ES-ionization and APC-ionization are the methods of choice. The analytical power of mass spectrometry is supported by the use of autosampling systems. The routinely performed detection in positive and negative scan mode provides complementary information on the substances analyzed. Adduct formation with alkali-ions and fragmentation of molecules can provide additional information to ensure correct assignment of molecular mass and molecular structure. Software programs are available for the automatic evaluation of spectra on a 'hit' and 'non-hit' basis. Nevertheless, in order to replace the analyst, some care must be taken in the estimation and the universality of these programs.

It is not advisable to rely solely on direct injection measurements of libraries, and a separation should be included before the mass spectrometric analysis. HPLC-MS-coupling is generally the method of choice for the analysis of small-compound libraries. To avoid time-consuming optimization, the use of a standard linear gradient is usually adequate. An UV-detector in series can provide complementary information. The library analysis further showed that ES-MS is found to be suitable for the low molecular mass range, although increasing baseline noise limits the mass range to a minimum of about  $m/z$  100. As with single compound analysis, sample throughput can be enhanced by automation.

For libraries with an upper mass limit of approximately 500 amu, GC-MS can prove advantageous. Capillary-GC on fused silica capillaries is characterized by high separation power and the ability to analyze relatively polar substances. In many cases, the problems of tailing, thermal instability and volatility associated with excessive analyte polarity can be overcome by the use of derivatization techniques.

Library analysis becomes more complicated when the formation of stereo- or regioisomers is possible, as the size of the library is increased by the number of possible isomers. In order to allow a rapid and uncomplicated evaluation, it is therefore important to design the library before synthesis such that the library size is restricted to a maximum of 15 to 20 possible compounds. It must be stressed that the MS investigations presented here have not fully explored the potency of either the mass spectrometric or the chromatographic techniques. The authors are conscious of the wealth of other experiments and possibilities to analyze fully a specific combinatorial library with optimized analytical techniques. Therefore, it is important to realize that in library analysis, compromises have to be made according to the priority of full sample characterization or high throughput.

## 17.5 Material and Methods

### 17.5.1 Electrospray-Mass Spectrometry

Mass spectra were recorded on an API III TAGA SCIEX (Perkin Elmer, Thornhill, Ontario, Canada) equipped with an electrospray ion source. Samples were diluted in ace-

tonitrile/water (1:1) at a concentration of 0.1 mg/ml. For single measurements, the sample was introduced into the ion source via a constant eluent flow (80  $\mu\text{l}/\text{min}$ ) of a solvent delivery system (Model S 1020, Sycam, Gilching, Germany) at 2-minute intervals (Gilson Abimed, Model 131). The orifice voltage was set to 80 V in positive and to  $-80$  V in negative ion scan mode.

For LC-MS couplings, a double syringe pump (Applied Biosystems, Model 140A) with a flow rate of 200  $\mu\text{l}/\text{min}$  was used. Separations were performed on a Nucleosil C-18  $2\times 100$  column (Grom, Herrenberg, Germany). An UV-detector (UVIS 204, Linear, Reno, Nevada, USA) was connected in series with the mass spectrometer and the flow rate to the ESI source was reduced to 80  $\mu\text{l}/\text{min}$  by stream splitting. Argon was used as the collision gas for neutral loss experiments. For high-throughput HPLC-MS analyses, the injection valve of the autosampler was connected to the HPLC-column. For the HPLC-analysis of the pyrrole libraries, a linear gradient 10% B to 100% B in 25 min was used. (Solvent A: 0.1% formic acid in water; solvent B: 0.1% formic acid in acetonitrile). Isoxazoline libraries were analyzed using 15% B to 50% B in 35 min and 100% B / 55 min as a gradient.

### 17.5.2 GC-EI Mass Spectrometry

GC-MS analyses were performed on a Varian MAT 112S (Bremen, Germany) connected to a Carlo Erba Fractovap 2900 gas chromatograph. Library compounds were separated on a Hewlett Packard Ultra-2 fused silica capillary (12 m $\times$ 0.2 mm;  $d_r = 0.33$   $\mu$ ). The column temperature was programmed as follows: 50°C for 2 min; 25°C/min to 150°C, 5°C/min to 300°C. The column head pressure was 50 kPa  $\text{H}_2$ .

Approximately 0.1 mg of library PYL-1 was trimethylsilylated by heating at 60°C for 30 min with a mixture of pyridine and BSTFA (1:1; 100  $\mu\text{l}$ ).

## 17.6 Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 323) and the Bundesministerium für Bildung, Forschung und Technologie BMBF project 03D0037A7. The authors thank T. Erlenmeyer and K. Bonk (Micromass, Idstein, Germany) for support with information material, and Dr. J. Zelter for helpful discussions.

## 17.7 Appendix

Mass spectrometric data of the analyzed libraries PYL-1 and ISO-4 are presented in this summary.



**Appendix 17.1.** Mass spectrometric data of the pyrrole library PYL-1.

Compound	Molecular mass (monoisotopic)	Found mass (M + H) <sup>+</sup>		
		Direct ES	LC-ES-MS	GC-MS
Py-1	152.2	151.8	152.0	–
Py-2	192.3	192.8	192.9	192.0
Py-3	194.3	195.0	195.1	194.0
Py-4	208.3	209.0	209.1	208.0
Py-5	210.3	210.8	210.9	(210.0); 282.0 (d)
Py-6	232.3	–	(232.9)	232.0
Py-7	248.4	248.8	249.1	248.0
Py-8	256.4	256.6	257.1	256.0
Py-9	262.4	262.8	262.9	262.0
Py-10	265.4	266.0	266.1	265.0
Py-11	268.4	268.8	268.9	268.0
Py-12	272.4	273.0	272.9	(272.0); 344.0 (d)
Py-13	286.3	286.8 (i)	287.1	286.0
Py-14	316.4	317.0	317.1	316.0
Py-15	325.5	325.8	325.8	325.0
Py-16	246.3	387.2 (bp)	387.2 (bp)	– (bp)
Py-17	286.4	286.8 (i)	287.1	(286.0); 358.0 (d)

– not found

i isobaric molecular weights

bp by-product

( ) weak signal intensity

d derivatized with trimethylsilylchloride

**Appendix 17.2.** Mass spectrometric data of the isoxazoline-library ISO-4.

Compound	Molecular mass (monoisotopic)	Found mass (M-H) <sup>-</sup>	
		Isomers	By-product
Is-1	257.3	256.2	254.2
Is-2	267.2	–	–
Is-3	267.2	266.2	264.2
Is-4	278.3	277.2	275.0
Is-5	283.3	282.2	280.2
Is-6	285.3	284.2	282.2
Is-7	295.3	294.2	292.2
Is-8	305.3	304.2 (i)	302.2 (i)
Is-9	305.3	304.2 (i)	302.2 (i)
Is-10	307.3	306.2	304.2
Is-11	311.7	310.2	308.2
Is-12	345.2	344.0	342.0
Is-13	355.5	354.2	352.2
Is-14	355.2	354.0	352.0
Is-15	367.4	366.2	364.2

– not found

i isobaric molecular weights

bp by-product

( ) weak signal intensity

## 17.8 References

- [1] J.A. Loo, *Europ. J. Mass Spectrom.* **1997**, 3, 93.
- [2] R.D. Süßmuth, G. Jung, *J. Chromat. B* **1999**, in press.
- [3] D. J. Ecker, S. T. Crooke, *Biotechnology* **1995**, 13, 351.
- [4] S. Stefanovic, K.-H. Wiesmüller, J. Metzger, A. G. Beck-Sickinger, G. Jung, *Bioorg. Med. Chem. Lett.* **1993**, 3, 431.
- [5] J. W. Metzger, C. Kempter, K.-H. Wiesmüller, G. Jung, *Anal. Biochem.* **1994**, 219, 261.
- [6] T. S. Kuhn (Ed.), "The structures of scientific revolutions", University of Chicago, Chicago, **1962**.
- [7] R. A. Zambias, D. A. Boulton, P. R. Griffin, *Tetrahedron Lett.* **1994**, 35, 4283.
- [8] E. De Hoffmann, J. Charlette, V. Stroobant (Eds.), "Mass Spectrometry", Wiley, Chichester, **1996**.
- [9] A. G. Harrison (Ed.), "Chemical Ionization Mass Spectrometry", CRC Press, Boca Raton, Fla., **1983**.
- [10] B. A. Baldwin, F. W. McLafferty, *Org. Mass Spectrom.* **1973**, 7, 1353.
- [11] M. Barber, R. S. Bordoli, G. J. Elliot, R. D. Sedgwick, A. N. Tyler, *Anal. Chem.* **1982**, 54, 645A.
- [12] R. M. Caprioli, T. Fan, J. S. Cottrell, *Anal. Chem.* **1986**, 58, 2949.
- [13] M. Karas, F. H. Hillenkamp, *Anal. Chem.* **1988**, 60, 2229.
- [14] B. A. Thomson, J. V. Irbaine, P. J. Dziedzic, *Anal. Chem.* **1982**, 54, 2219.
- [15] E. C. Huang, T. Wachs, J. J. Conboy, J. D. Henion, *Anal. Chem.* **1990**, 62, 713A.
- [16] W. Paul, H. P. Reinhard U. v. Zahn, *Z. Phys.* **1958**, 152, 143.
- [17] J. E. Campana, *Int. J. Mass Spec. Ion Phys.* **1980**, 33, 101.
- [18] R. J. Cotter, *Anal. Chem.* **1992**, 64, 1027A.
- [19] M. V. Buchanan, R. L. Hettich, *Anal. Chem.* **1993**, 65, 245A.
- [20] J.W. Metzger, K.-H. Wiesmüller, V. Gnau, J. Brünjes, G. Jung, *Angew. Chem. Int. Ed. Engl.* **1993**, 32, 894.
- [21] S. Kienle, K.-H. Wiesmüller, J. Brünjes, J.W. Metzger, G. Jung, *Fresenius J. Anal. Chem.* **1997**, 359, 10.
- [22] P.-H. Lambert, J.A. Boutin, S. Bertin, J.-L. Fauchere, J.-P. Volland, *Rapid Commun. Mass Spectrom.* **1997**, 11, 1971.
- [23] B.E. Winger, J.E. Campana, *Rapid. Commun. Mass Spectrom.* **1996**, 10, 1811.
- [24] J.P. Nawrocki, M. Wigger, C.H. Watson, T.W. Hayes, M.W. Senko, S.A. Benner, J.R. Eyler, *Rapid Commun. Mass Spectrom.* **1996**, 10, 1860.
- [25] B.J. Egnor, G.J. Langley, M. Bradley, *J. Org. Chem.* **1995**, 60, 2652.
- [26] B.J. Egnor, M. Cardno, M. Bradley, *J. Chem. Soc. Chem. Commun.* **1995**, 2163.
- [27] R.A. Zambias, D.A. Boulton, P.R. Griffin, *Tetrahedron Lett.* **1994**, 35, 4283.
- [28] B.B. Brown, D.S. Wagner, H.M. Geysen, *Mol. Div.* **1995**, 1, 4.
- [29] M.C. Fitzgerald, K. Harris, C.G. Shevlin, G. Siuzdak, *Bioorg. Med. Chem. Lett.* **1996**, 6, 979.
- [30] T. Carell, E.A. Wintner, A.J. Sutherland, J. Rebek Jr., Y.M. Dunayevskiy, P. Vouros, *Chem. Biol.* **1995**, 2, 171.
- [31] Y. Dunayevskiy, P. Vouros, T. Carell, E.A. Wintner, J. Rebek Jr., *Anal. Chem.* **1995**, 67, 2906.
- [32] A.S. Fang, P. Vouros, C.C. Stacey, G.H. Kruppa, F.H. Laukien, T. Carell, E.A. Wintner, J. Rebek Jr., *Proc. 44th ASMS Conf. on Mass Spectrom. and Allied Topics*, Portland, OR, 12-16 May **1996**, 1047.
- [33] K. Beimann, in J.A. McCloskey (Editor), *Methods in Enzymology*, Academic Press, San Diego, CA, **1990**, Vol 193, p. 295.
- [34] Q. Wu, *Anal. Chem.* **1998**, 70, 865.
- [35] N.J. Haskins, D.J. Hunter, A.J. Organ, S.S. Rahman, C. Thom, *Rapid Commun. Mass Spectrom.* **1995**, 9, 1437.
- [36] W.J. Haap, J.W. Metzger, C. Kempter, G. Jung, *Mol. Div.* **1997**, 3, 29.
- [37] W.J. Haap, D. Kaiser, T.B. Walk, G. Jung, *Tetrahedron* **1998**, 54, 3705.
- [38] Y.M. Dunayevskiy, P. Vouros, E.A. Wintner, G.W. Shipps, T. Carell, J. Rebek Jr., *Proc. Natl. Acad. Sci. USA* **1996**, 93, 6152.
- [39] C. Chen, L.A. Ahlberg Randall, R.B. Miller, A.D. Jones, M.J. Kurth, *J. Am. Chem. Soc.* **1994**, 116, 2661.
- [40] C. Chen, L.A. Ahlberg Randall, R.B. Miller, A.D. Jones, M.J. Kurth, *Tetrahedron* **1997**, 53, 6595.
- [41] B. Yan, J. B. Fell, G. Kumaravel, *J. Org. Chem.* **1996**, 61, 7467.

- [42] M. F. Gordeev, D.V. Patel, E. M. Gordon, *J. Org. Chem.* **1996**, *61*, 924.
- [43] R.S. Garigipati, J.L. Adams, S.K. Sarkar, *J. Org. Chem.* **1996**, *61*, 2911.
- [44] M.C. Fitzgerald, K. Harris, C.G. Shelvin, G. Siuzdak, *Bioorg. Med. Chem. Lett.* **1996**, *6*, 979.
- [45] B.J. Egner, J. Langley, M. Bradley, *J. Org. Chem.* **1995**, *60*, 2652.
- [46] N.J. Haskins, D.J. Hunter, A.J. Organ, S.S. Rahman, C. Thom, *Rapid Commun. Mass Spectrom.* **1995**, *9*, 1437.
- [47] D.C. Stahl, K.M. Swiderek, M.T. Davis, T.D. Lee, *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 532.
- [48] J. Wu, K. Chatman, K. Harris, G. Siuzdak, *Anal. Chem.* **1997**, *69*, 3767.
- [49] D.P. Little, T.J. Cornish, M.J. O'Donell, A. Braun, R.J. Cotter, H. Köstner, *Anal. Chem.* **1997**, *69*, 4540.
- [50] D.C. Tutko, K.D. Henry, B.E. Winger, H. Stout, M. Hemling, *Rapid Commun. Mass Spectrom.* **1998**, *12*, 335.
- [51] B.E. Winger, J.A. Castoro, D.C. Tutko, J.E. Campana, *Proc. 44th ASMS Conf. on Mass Spectrom. and Allied Topics*, Portland, OR, 12-16 May **1996**, 1134.
- [52] H.N. Weller, M.G. Young, S.J. Michalczyk, G.H. Reitnauer, R.S. Cooley, P.C. Rahn, D.J. Loyd, D. Fiore, S.J. Fischman, *Mol. Div.* **1997**, *3*, 61.
- [53] C.E. Kibbey, *Lab. Robot. Automat.* **1997**, *9*, 309.
- [54] L. Zeng, L. Burton, K. Yung, B. Shushan, D.B. Kassel, *J. Chromat. A* **1998**, *794*, 3.
- [55] G. Hegy, E. Görlach, R. Richmond, F. Bitsch, *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1894.
- [56] L.C.E. Taylor, R.L. Johnson, R. Raso, *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 387.
- [57] A.A. MacDonald, S.H. Dewitt, S. Ghosh, E.M. Hogan, L. Kieras, A.W. Czarnik, R. Ramage, *Mol. Div.* **1995**, *1*, 183.
- [58] A. Trautwein, R. D. Süßmuth, G. Jung, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2381.
- [59] H. Richter, *unpublished results*.
- [60] T. Mukaiyama, T. Hoshino, *J. Am. Chem. Soc.* **1960**, *82*, 5339.

# 18 High-Resolution Magic Angle Spinning (MAS) NMR Spectroscopy for On-Bead Analysis of Solid-Phase Synthesis

*Ralf Warrass and Guy Lippens*

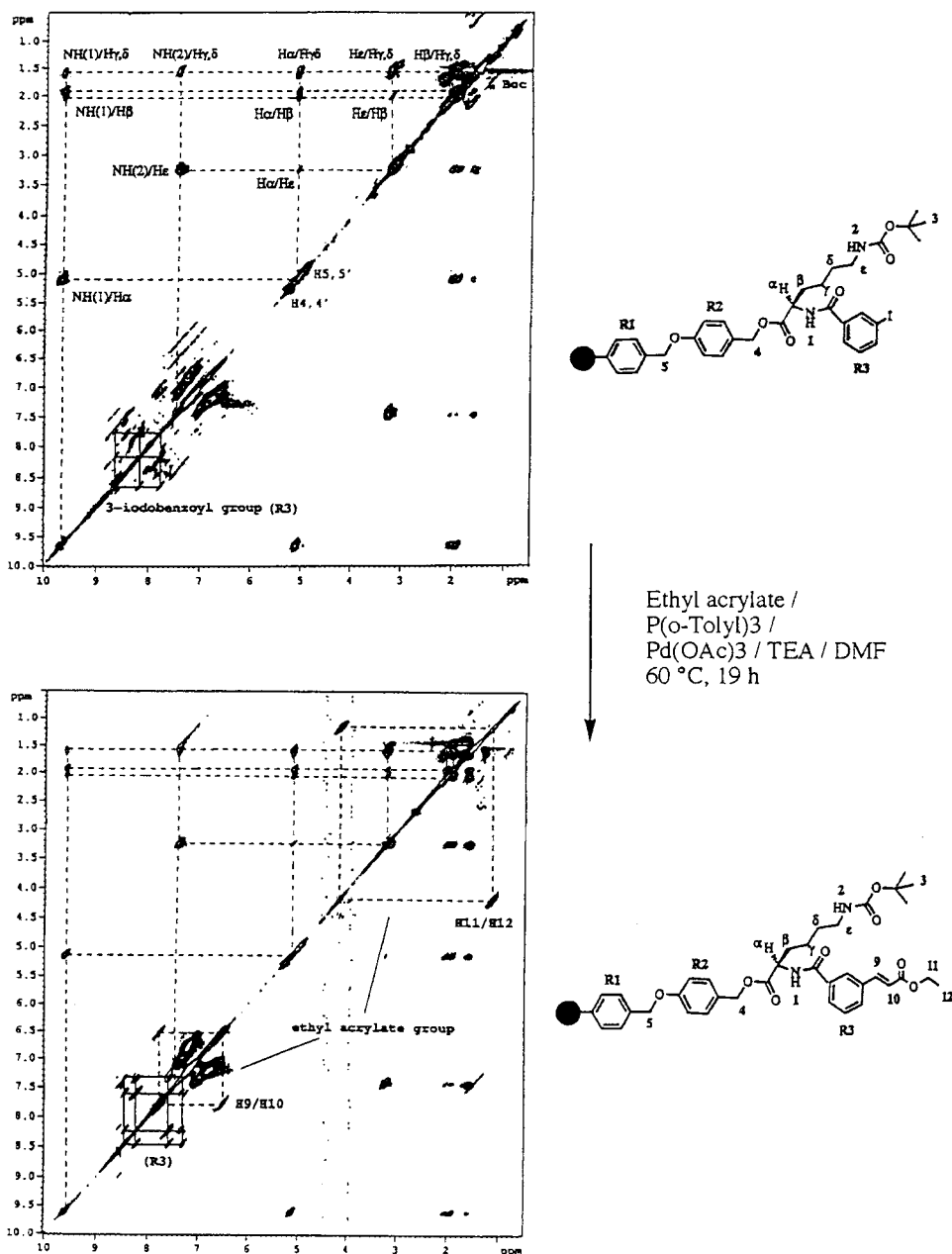
## 18.1 Introduction

Nuclear magnetic resonance (NMR), which was first discovered in the field of physics, has developed into one of the analytical tools of choice for the identification of organic molecules in solution, as it allows the unambiguous assignment of every proton and carbon nucleus. Especially since the advent of the two-dimensional (2D) techniques, where correlation of nuclear spins became feasible [1], NMR has developed into a routine technique in most laboratories. Its major drawbacks, however, remain its inherent lack of sensitivity (Planck's constant being involved in the energy distribution of the spins!), and its extreme sensitivity to magnetic field inhomogeneities. Major technical advancements have led to stable superconducting magnets that generate magnetic fields close to 20 Tesla, or a corresponding 800 MHz proton Larmor frequency, with field inhomogeneities that can be smaller than 1 Hz. Whereas this impressive achievement diminishes the inherent limitations of the technique, the field inhomogeneity causes major problems when one wants to study truly inhomogeneous samples, such as resin-bound molecules encountered in solid-phase organic synthesis. Even when 'mobility is injected' in such systems by appropriate swelling conditions, the inhomogeneous nature of the sample, with its insoluble matrix, attached molecules, swelling and interstitial solvent, leads to diverse magnetic fields that all tend to destroy the  $B_0$  homogeneity, and therefore to proton lines that are several hundreds of Hz wide.

After a short overview of existing methods, we will highlight in this chapter on several practical aspects of HR MAS NMR. We will underline the efficiency of HR MAS NMR studies by introducing experiments to quantitatively monitor solid-phase supported reactions and describing a method that allows the use of protonated solvent. Recently published reviews covering NMR methods, analytical tools in combinatorial chemistry and in the drug discovery process may be used to have an easy access to this exciting field [2–4].

## 18.2 Methods Used for On-Bead NMR Analysis

$^{13}\text{C}$  gel-phase NMR spectroscopy, because of smaller line widths and better chemical shift dispersion of the  $^{13}\text{C}$  lines, is a reliable technique to determine the success or failure of chemical transformations. However, as a consequence of the low sensitivity of the



**Figure 18.1.** Phase-sensitive two-dimensional (2D) MAS TOCSY (total correlation spectroscopy) NMR spectra to monitor the performance of a Heck reaction on a Wang resin. The sample was suspended in *d*<sub>6</sub>-pyridine and spun at 2000 Hz in 7-mm rotors using a conventional solid-state NMR probe of a 300 MHz. A 70-ms MLEV-17 spin-lock was introduced and the data were acquired with 16 scans of each 256 t1-increments [26].

method (1% natural abundance of  $^{13}\text{C}$ ), acquisition of a spectra with a suitable signal-to-noise ratio can take several hours. To circumvent the problem of low sensitivity,  $^{13}\text{C}$ -enriched samples [5, 6] or  $^{19}\text{F}$  gel-phase NMR [7, 8] have been used to analyze support-bound structures with standard NMR equipment. The disadvantage is that labels have to be introduced in the bound molecules, thus limiting its general use. Proton NMR of resin-bound structures became feasible once it was realized that the application of magic angle spinning (MAS) [9, 10] can be used to reduce in a spectacular way the field inhomogeneities due to the granular nature of the sample [11], and the technique has been applied successfully by a number of groups in solid-phase organic chemistry (SPOC) [12–25].

One-dimensional (1D) spectra form the basis for further 2D spectra, and basically the whole battery of liquid-phase NMR experiments is available for the further identification of the compounds attached to the solid support. A nice example is presented by Pop et al. [26] in monitoring the carbon-carbon-forming Heck reaction on Wang resin using 1D and 2D experiments (Fig. 18.1). In another study,  $^1\text{H}/^{15}\text{N}$ -HMQC and NOE spectroscopy was used for the determination of conformational changes of different hexapeptides attached to a solid support [21]. Since the determination of coupling constants ( $J$ ) is often desirable for structural analysis, a multiple-quantum filtered COSY (E. COSY) experiment for the reduction of inhomogeneous line broadening was developed [27].

Further efforts to minimize the amount of resin to be investigated by NMR spectroscopy has led to the development of single bead measurements. Sarkar et al. [17] presented the detection of  $^{13}\text{C}$ -labeled 2,4-dimethoxybenzoic acid bound to a single resin bead, which had a particle size of 100  $\mu\text{m}$  corresponding to 800 pmol of sample. This work shows how sensitivity in NMR spectroscopy has been improved within the last few years, although 4 hours of acquisition time were needed for an  $^1\text{H}$  MAS NMR spectra. Another disadvantage is the  $^{13}\text{C}$  labeling of the structures required to obtain overnight  $^1\text{H}$ - $^{13}\text{C}$  inverse correlated (HMQC) NMR spectra. Simultaneous to Sarkar et al. [17], Pursch et al. [28] have performed single bead measurements on TentaGel resins in order to monitor the progress of a hydantoin synthesis sequence by  $^1\text{H}$  MAS NMR spectroscopy. For these experiments, macro beads with mean particle sizes of 700  $\mu\text{m}$  (corresponding to 10–65 nmol of polymer-bound structures) were swollen in  $d_6$ -DMSO or  $\text{CDCl}_3$  and placed into a 4-mm rotor with insert. Another interesting application of HR MAS is the structure determination of molecules attached to multipin crowns [23].

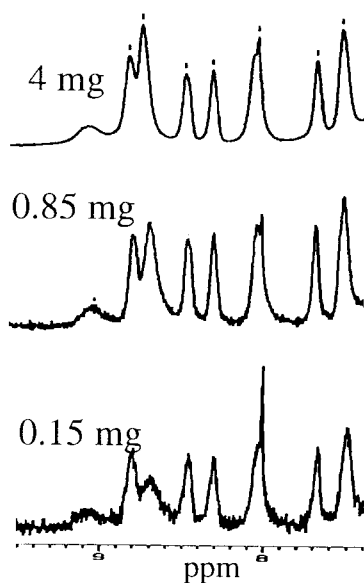
### 18.3 Sample Preparations, Observations and Quantifications by HR MAS NMR

We have addressed in our laboratory several aspects of the more rapid and robust identification of resin-bound entities. Firstly, related to the question of sensitivity and sample preparation, how much of a resin do we really need to obtain workable spectra in a given time lapse of 1 minute, and what is the optimal sample preparation procedure? Whereas we and others in the early stages prepared the sample by swelling the resin in a separate recipient, and transferring part of the slurry into the rotor by a specially machined spatula, we later found out that swelling the resin beads directly into the rotor led in a faster and more economic way (less resin and deuterated solvent is needed) to more reprodu-

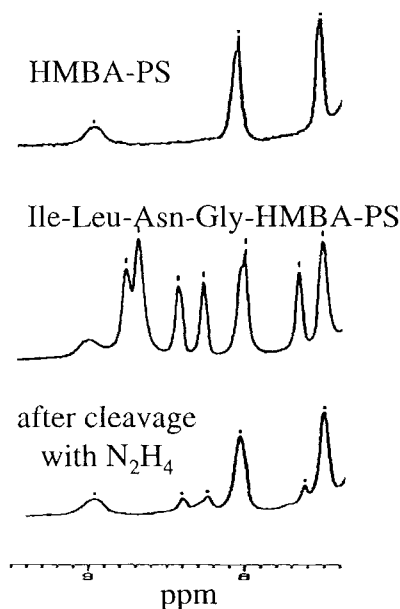
cible samples. The influence of the solvent, essential for correct swelling of the beads, has been discussed at length, but without doubt, we found that the best evaluation criterion consisted of monitoring the diameter of the individual swollen bead under a graduated microscope [29]. This, rather than the macroscopic volume of a given quantity of beads, indicates the amount of solvent that actually has penetrated into the bead and is solvating the attached molecules. To generalize this idea, a solvent that has a good swelling capacity of a polymer is also useful for the corresponding NMR studies (DMF, DCM,  $\text{CDCl}_3$ , pyridine, etc. are good solvents for polystyrene; water can be used as solvent if PEGA resins are employed). By swelling the resin directly in the rotor with an appropriate deuterated organic solvent, we obtained reasonable spectra in 1 minute on sub-milligram quantities of resin (Fig. 18.2). The ease of the methods allows the fast and reliable control of both synthesis and cleaving conditions, as is demonstrated in Figure 18.3, where the  $^1\text{H}$  spectra of the pure resin, the product-carrying resin and the resin following the treatment with a cleavage cocktail (in this case hydrazine in THF/MeOH) are shown. In the depicted example it becomes clear that a small amount of product resisted the cleaving conditions.

Another interesting phenomenon observed by HR MAS NMR concerns the role of residual water inside the resin, retained especially in PEG resins even after prolonged drying under reduced pressure. First, we never observed  $\text{NH}_2$  signals, during analysis of N-terminal free peptides (in Fig. 18.4, H-Leu-Lys(Boc)-Gly-Glu(tBu) bound to TentaGel resin). That this is due to rapid exchange with protons of residual water molecules is further confirmed by the spectra shown in Figure 18.4. Presaturation of the water signal at

### Ile-Leu-Asn-Gly-HMBA-PS



**Figure 18.2.**  $^1\text{H}$  spectra (region of amide protons) of different amounts of a polystyrene/1% divinylbenzene resin (0.5 mmol/g; in  $d_7$ -DMF) carrying the tetrapeptide Gly-Asn-Leu-Ile through an hydroxymethyl benzoic acid (HMBA) linker. Acquisition was performed in 1 minute (16 scans).

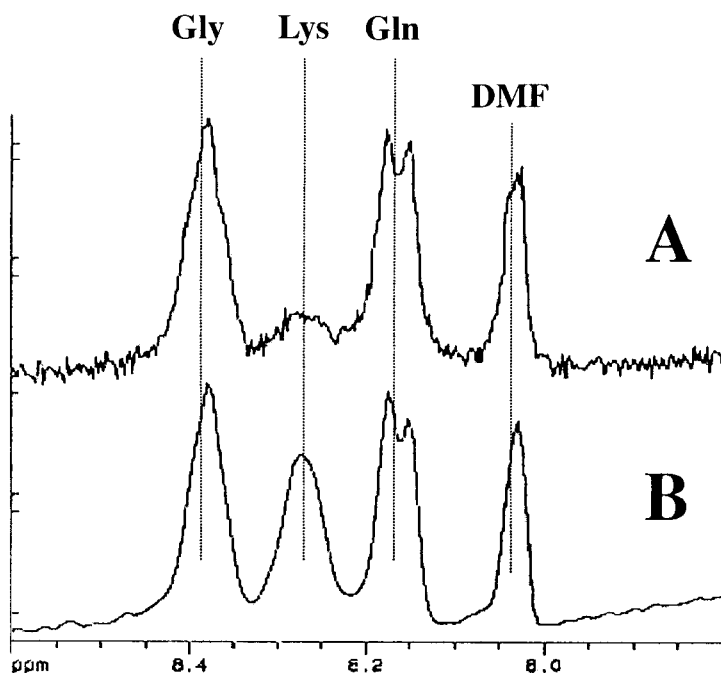


**Figure 18.3.**  $^1\text{H}$  spectra (region of amide protons) of a polystyrene/1% divinylbenzene resin (0.5 mmol/g; in  $d_7$ -DMF) allow fast and reliable controls of synthesis and cleavage conditions. Top:  $^1\text{H}$  spectra of the pure HMPA-PS resin; center:  $^1\text{H}$  spectra after the synthesis of a tetrapeptide Gly-Asn-Leu-Ile; bottom:  $^1\text{H}$  spectra after the cleavage of the tetrapeptide with hydrazine. The latter spectrum shows some small amounts of product still attached to the solid support.

3.75 ppm led to a loss of signal intensity only for the amide proton of the second-last amino acid (8.26 ppm). Water seems therefore sequestered near the free N terminus, while still communicating with the last two amino acids. We even observed an exchange cross-peak between water and the second-last amino acid amide proton in NOESY spectra. This is especially interesting since amide bond formation in peptide synthesis is done by activated esters and anhydrides, which are sensitive to moisture, and illustrates the gain of information by the analysis of polymer-bound structures by HR MAS NMR.

A second problem we have addressed is the one of quantification of the molecular entities [30]. Quantification is the first step in reaction monitoring, without which no reaction optimization is possible. After cleavage, NMR has been used successfully to quantify the products of a SPOC reaction, but our effort was directed to quantifying the molecules directly on the resin. To know the exact amount of resin-bound moieties is an often underestimated problem. This becomes clear, if a precise quantity of resin-bound product has to be separated from a large batch in order to further elaborate its structure. Since the resin charge (loading of the resin) was determined several reaction steps before (often by the UV quantification of cleaved Fmoc derivatives), it most probably altered by the simple inclusion of moisture or solvent, by premature cleavage and by the additional weight of the synthesized structure (which can be a major part of the resin weight). In a first experiment to monitor quantitatively a solid-phase reaction, we performed a Horner–Emmons transformation [31] directly in the HR MAS NMR rotor (4 mg of resin carrying a phosphono diester and 40  $\mu\text{l}$  of  $d_7$ -DMF containing LiBr,  $\text{Et}_3\text{N}$  and an aldehyde), and quantified the polymer-bound reaction products by comparing its integrated signal to that of the soluble, exactly quantified, starting material. With this technique, reaction ki-





**Figure 18.4.**  $^1\text{H}$  spectra (region of amide protons) of different amounts of a polystyrene/1% divinylbenzene resin (0.27 mmol/g) carrying the tetrapeptide Glu(tBu)-Gly-Lys(Boc)-Leu through a Wang linker. (A) Without presaturation; (B) with presaturation of the water signal. The presaturation of the water signal at 3.75 leads to a loss of signal intensity of the second-last amino acid amide proton (8.26 ppm). This phenomenon can be explained by the presence of water in the region of the peptide N terminus.

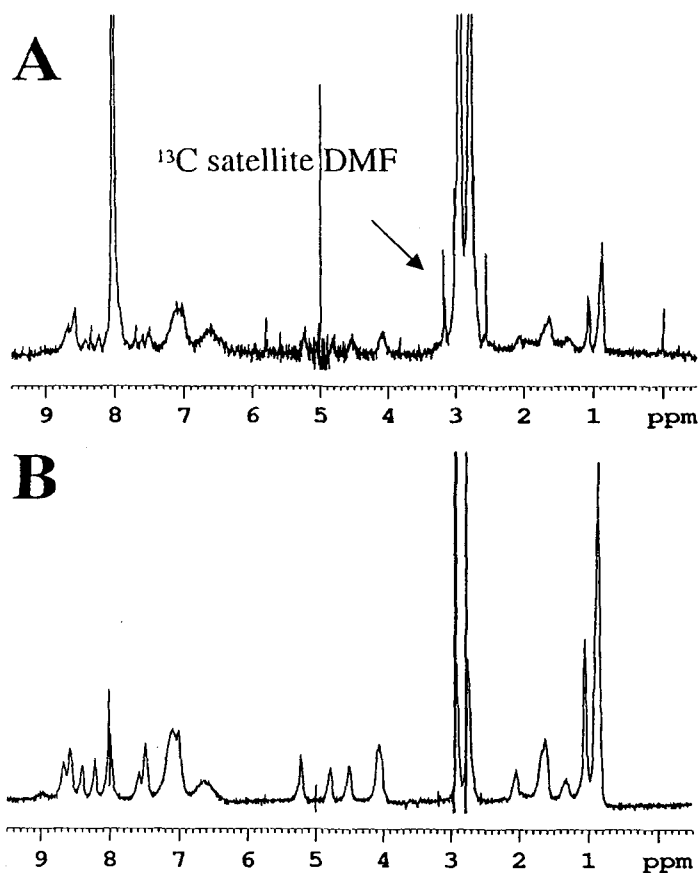
netics can be followed under different reaction conditions in order to optimize yields and purity of the resin-attached product. Since we do not pass by the time-consuming cleave and analysis procedure, possible modifications of the analyzed molecules by cleaving conditions are avoided.

Despite the straightforward implementation of the reaction in the HR MAS NMR rotor, one major problem of this technique is that the reaction conditions are very different from the standard ones in standard reaction vessels, due to the small volume of the sample and the high rotation speeds during NMR analysis. A second experiment therefore consists of sampling a small amount of resin (a few milligrams) at different points in time, and preparing them into a NMR sample by working steps of washing, drying in a volatile solvent, and reswelling in a deuterated solvent in the NMR rotor. In order to quantify the amount of molecules on the resin, a given amount of soluble standard with a well-distinguishable NMR signal is added, and its integral is compared to that of the resin-bound molecule. Despite its general use, special care should be used for the recycle delay with this external quantification technique, as the soluble molecule can have appreciable longer  $T_1$  relaxation times than the resin-bound molecules.

During our HR MAS NMR studies we observed that for certain soluble molecules such as the tetramethylsilane (TMS, used as a reference), two signals appear in the presence of a polystyrene resin, corresponding to the molecules inside the beads and those in the interstitial solvent. As the most recent HR MAS probes are equipped with pulsed field gradients [16], one can use the sequences developed to study the diffusion behavior of molecules [32] in the supermolecular system that is the swollen resin. Identification of internal and external substances is therefore possible, as we expect and observe lower diffusion coefficients for molecules inside the polystyrene matrix. The precise origin of this differential chemical shift value is not only a bulk effect provoked by the abundant aromatic rings, as we observed similar differences in non aromatic resins such as the PEGA resin. Further research into the influence of the nature of the resin on the NMR signals and, more importantly, on the chemical reactivity, is currently in progress in our group.

## 18.4 Use of Protonated Solvent in HR MAS NMR

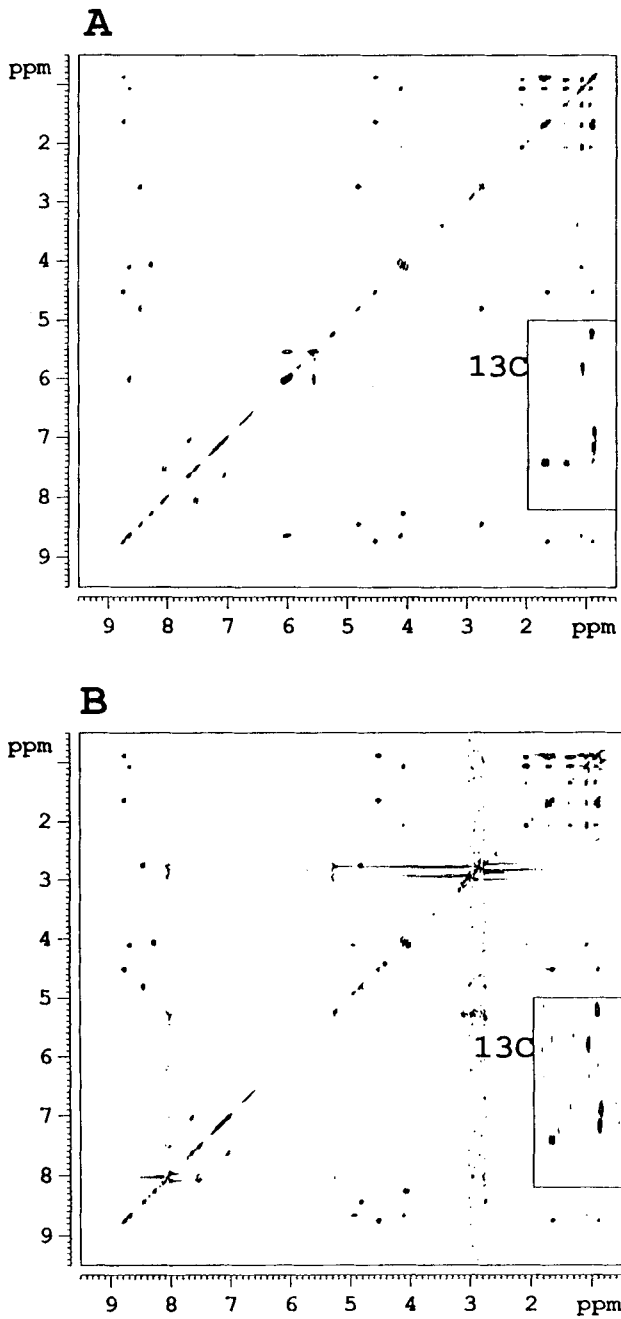
While studying the diffusion of the molecules, we realized that this could be a very effective technique to suppress the solvent signal. A major problem with the technique of HR MAS remains indeed the need for the above-described working steps of washing, drying and reswelling in a deuterated solvent prior to NMR analysis. Despite the fact that this procedure is considerably faster than most cleave-and-analysis techniques, it is still too slow to allow a rapid monitoring of the many samples that might generate a typical combinatorial chemistry program, where the use of deuterated solvents throughout the synthesis is far too cost-demanding. We therefore used a LED pulse sequence [32–34], originally devised as a sequence for measurement of the diffusion coefficient, to eliminate all freely diffusing molecules, including the protonated solvent [35]. As can be seen in Figure 18.5, this allows the use of a higher receiver gain, and only the reacted species, which is bound to the insoluble matrix, retains its NMR signals. The method is the first general analytical technique that requires no working steps for analysis of resin-bound structures under standard reaction conditions (protonated solvent, excess of reagent), except for the transfer of a small amount of resin from the reaction vessel to the NMR rotor. This method is not restricted to 1D NMR experiments, and the LED pulse sequence can be included in standard TOCSY, HSQC and NOESY experiments (for the first two, see Fig. 18.6). In order to transfer reproducibly the approximately same amount of resin into the rotor, we first shake vigorously the vessel, sample 1 ml of suspension into a pipette, allow it to rest for a short interval during which the beads sink into the tip of the pipette (true e.g. for DMF), and then transfer this to the rotor. Returning to our need to quantify structures attached to the polymer, we note that the addition of an external standard no longer works, as the diffusion filter eliminates all signals of soluble molecules. Quantification can only be carried out by comparing signals to other signals of protons covalently attached to the resin support.



**Figure 18.5.**  $^1\text{H}$  spectra of 0.15 mg of polystyrene/1% divinylbenzene resin (0.27 mmol/g) carrying the tetrapeptide Gly-Asn-Leu-Ile through an HMBA linker; (A) in  $h_7$ -DMF, (B) in  $h_7$ -DMF with diffusion filter (10%  $d_7$ -DMF is added as lock signal).

## 18.5 Summary and Perspectives

HR MAS NMR greatly simplifies the optimization procedure of solid-phase organic synthesis, as it represents an easy, reliable and very effective way of analyzing solid support-bound structures at the atomic level (all carbons and protons can be used). In the future, the aim of our and other groups will lead to investigations of larger molecules and their secondary structure on solid support and the rapid and automated characterization of resin-bound combinatorial libraries. We believe that MAS NMR methods will enter solid-phase chemistry laboratories in as powerful a manner as NMR spectroscopy once entered classical organic chemistry laboratories. The analytical observation of molecules attached to a solid support is one of the major problems in solid-phase chemistry, but with new developments in NMR and new resins, the high standard of liquid-state analysis will



**Figure 18.6.** 2D NMR spectra (TOCSY and HSQC (insert)) of 2.5 mg of polystyrene/1% divinylbenzene resin (0.38 mmol/g) carrying the deprotected tetrapeptide Gly-Asn-Leu-Ile through an HMBA linker. (A) In  $d_7$ -DMF; (B) in  $h_7$ -DMF with diffusion filter.

be achieved. Furthermore, the use of the diffusion experiments described above leads to possibilities that liquid-state NMR does not have, resulting in another advantage of solid-phase organic chemistry in modern organic chemistry.

## 18.6 References

- [1] J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, *J. Chem. Phys.* **71** (1979), 4546–4553.
- [2] P. A. Keifer, *Curr. Opin. Biotech.* **10** (1999), 34–41.
- [3] M. J. Shapiro, J. R. Wareing, *Curr. Opin. Chem. Biol.* **2** (1998), 372–375.
- [4] P. A. Keifer, *Drug Discovery Today* **2** (1998), 468–478.
- [5] T. Kanemitsu, O. Kanie, C.-H. Wong, *Angew. Chem.* **110** (1998), 3574–3577; *Angew. Chem. Int. Ed. Engl.* **38** (1998), 3415–3418.
- [6] G. C. Look, C. P. Holmes, J. P. Chinn, M. A. Gallop, *J. Org. Chem.* **59** (1994), 7588–7590.
- [7] D. Stones, D. J. Miller, M. W. Beaton, T. J. Rutherford, D. Gani, *Tetrahedron Lett.* **39** (1998), 4875–4878.
- [8] A. Svensson, K.-E. Bergquist, T. Fex, J. Kihlberg, *Tetrahedron Lett.* **39** (1998), 7193–7196.
- [9] E. R. Andrew, A. Bradbury, R. G. Eades, *Nature* **183** (1959), 1802–1803.
- [10] I. J. Lowe, *J. Phys. Rev. Lett.* **2** (1959), 285–287.
- [11] L. Vanderhart, in D. M. Grant, R. K. Harris (Eds.): *Encyclopedia of Nuclear Magnetic Resonance, Vol. 5*, Wiley & Sons, New-York 1996, p. 2938–2946.
- [12] W. L. Fitch, G. Detre, C. P. Holmes, J. N. Shoolery, P. A. Keifer, *J. Org. Chem.* **59** (1994), 7955–7956.
- [13] R. C. Anderson, J. P. Stokes, M. J. Shapiro, *Tetrahedron Lett.* **36** (1995), 5311–5314.
- [14] R. C. Anderson, M. A. Jarema, M. J. Shapiro, J. P. Stokes, M. Ziliox, *J. Org. Chem.* **60** (1995), 2650–2651.
- [15] R. S. Garigipati, B. Adams, J. L. Adams, S. K. Sarkar, *J. Org. Chem.* **61** (1996), 2911–2914.
- [16] W. E. Maas, F. H. Laukien, D. G. Cory, *J. Am. Chem. Soc.* **118** (1996), 13085–13086.
- [17] S. K. Sarkar, R. S. Garigipati, J. L. Adams, P. A. Keifer, *J. Am. Chem. Soc.* **118** (1996), 2305–2306.
- [18] P. A. Keifer, L. Baltusis, D. M. Rice, A. A. Tymiak, J. N. Shoolery, *J. Magn. Res.* **119** (1996), 65–75.
- [19] T. Wehler, J. Westman, *Tetrahedron Lett.* **37** (1996), 4771–4774.
- [20] M. J. Shapiro, J.-Chin, R. E. Marti, M. A. Jarosinski, *Tetrahedron Lett.* **38** (1997), 1333–1336.
- [21] R. Jelinek, A. P. Valente, K. G. Valentine, S. J. Opella, *J. Magn. Res.* **125** (1997), 185–187.
- [22] P. A. Keifer, *J. Org. Chem.* **61** (1996), 1558–1559.
- [23] J. Chin, B. Fell, M. J. Shapiro, J. Tomesch, J. R. Wareing, A. M. Bray, *J. Org. Chem.* **62** (1997), 538–539.
- [24] J. Chin, B. Fell, S. Pochapsky, M. J. Shapiro, J. R. Wareing, *J. Org. Chem.* **63** (1998), 1309–1311.
- [25] R. Riedl, R. Tappe, A. Berkessel, *J. Am. Chem. Soc.* **120** (1998), 8994–9000.
- [26] I. Pop, C. Dhalluin, B. Déprez, P. Melnyk, G. Lippens, A. Tartar, *Tetrahedron* **52** (1996), 12209–12222.
- [27] A. Meissner, P. Bloch, E. Humpfer, M. Spraul, O. W. Sorensen, *J. Am. Chem. Soc.* **119** (1997), 1787–1788.
- [28] M. Pursch, G. Schlotterbeck, L.-H. Tseng, K. Albert, W. Rapp, *Angew. Chem.* **108** (1996), 3034–3036; *Angew. Chem. Int. Ed. Engl.* **35** (1996), 2867–2869.
- [29] C. Dhalluin, C. Boutillon, A. Tartar, G. Lippens, *J. Am. Chem. Soc.* **119** (1997), 10494–10500.
- [30] R. Warrass, G. Lippens,
- [31] C. R. Johnson, B. Zhang, *Tetrahedron Lett.* **36** (1995), 9253–9256.
- [32] S. J. Gibbs, C. S. J. Johnson, *J. Magn. Res.* **93** (1991), 395–402.
- [33] A. J. Dingley, J. P. Mackay, B. E. Chapman, M. B. Morris, P. W. Kuchel, B. D. Hambly, G. F. King, *J. Biomol. NMR* **6** (1995), 321–328.
- [34] A. S. Altieri, D. P. Hinton, R. A. Byrd, *J. Am. Chem. Soc.* **117** (1995), 7566–7567.
- [35] R. Warrass, J.-M. Wieruszkeski, G. Lippens, *J. Am. Chem. Soc.* **121** (1999), 3787–3788.

# 19 Automated Combinatorial Chemistry

*Martin Winter, Karl-Heinz Wiesmüller and Günther Jung*

## 19.1 Introduction

The generation of organic molecules is traditionally a manual process, involving the chemist in many hours of work to plan and organize, to dissolve, to mix, react, purify and analyze compounds. Using traditional methods, the chemist usually creates up to several hundred new compounds per year. Fully automated organic chemistry with highly developed hardware and software accelerates the synthesis process to thousand-fold productivity.

Parallel organic synthesis has recently become a preferred mode in the practice of combinatorial chemistry due to the necessity to produce numerous discrete compounds in milligram quantities (see Chapter 1; see also [1–11]). Automated chemical synthesis is a necessity in order to keep pace with highly automated screening laboratories which often have the capacity to screen up to hundred thousand compounds per day (see Chapter 14). High-throughput modular synthesis systems are often based around a unique reaction vessel design. This allows direct link from combinatorial synthesis arrays to arrays of traditional high-throughput screening (HTS) of compound collections for the discovery of new lead compounds [12]. Automation includes both solid- and solution-phase chemistry, and is also capable of tackling the tedious purification processes and sample preparation for high-throughput analytical characterization. Flexible systems allow the performance of a wide range of tasks, including the highly parallel generation of compounds for primary screening, the synthesis of intermediates, synthons, derivatives of lead compounds, optimization by parallel modification, and iterative determination of process parameters.

The requirements for automated chemistry are:

- Preparation of multi-milligram quantities of compounds as well as microscale synthesis
- Cheap, adaptable and easily transportable and cleanable reaction vessels
- Modularity of equipment for a range of options
- Option for many diverse chemical reaction conditions, *e.g.*, inert atmosphere, corrosive reagents, temperature control
- Process and monitoring for flexible work-up procedures, including evaporation and extraction methods
- Products transfer to standard 96-well format compatible with HTS technologies
- Software instruction by computer-generated reaction instructions; no need for special training or programming knowledge

- Plausibility control for newly written chemfiles and for the synthesis procedure
- Excluding cross-contamination during reagents transfer and during synthesis and work-up procedures
- Delivery of resin particles and solid chemicals
- Rapid delivery of any reagent to the reaction vessel for complex reactions and synthesis schemes.
- Rapid wash cycles, highly parallel for hundreds of reactors
- Auditability for safety considerations and to allow automatic recording of procedures, synthesis history and results.

## 19.2 Parallel Synthesis of Individual Compounds

Manufacturing a complex diversity of substances today forms part of the standard methodological repertoire of combinatorial chemistry – to the extent, that is, that we have in mind substance mixtures. Often, however, reciprocities between the individual components of such mixtures impact negatively on the identification of individual compounds active in a biological or biochemical system. The iterative procedure ('deconvolution') involving repeated resynthesis and testing of mixtures of lower complexity which leads finally to the individual compounds can give rise to problems, particularly with non-peptide substance libraries. A massive extension of simultaneous synthesis of individual compounds has therefore increasingly replaced the synthesis of substance mixtures. Parallel syntheses in the required amounts only make sense with automated procedures, which in recent times have been the focus of intense developmental efforts and have now reached the market in the form of new gadgetry. The automated pipetting device supporting the synthesis steps – which is already in widespread use – is only one instance of the multiplicity of technical options that today's researchers are able to deploy.

Only by means of modern information processing it is possible to manage the flood of data generated by high throughput, ranging from synthesis to the results of active-substance/agent screening.

The selection of building blocks to manufacture thousands of compounds hardly makes sense any more using classical methods, no more than does routine perusal of individual test data from HTS. Coding procedures, *e.g.*, bar codes on 96-well plates, have replaced the classical designation of individual vessels in synthesis and test compounds. This coding is retained until the results of bioassays are analyzed. The already tested compounds receive, in the compound database, a feature enhanced by the classical chemical data and descriptors, indicating their location in the space of biological effects. The stored data permit the search for new lead structures to be rationalized, since a computer-assisted preselection of highly promising compounds from the general substance archive can be made for screening purposes. Thus, only the highly promising so-called 'focused test libraries', which are oriented to the biological target molecule, are fed into HTS.

In what follows, we shall be looking at automated synthesis techniques developed to manufacture individual compounds. Although an automated synthesis of compound mixtures along the same lines is possible in theory, for the reasons stated above it plays an insignificant role.

## 19.3 Productivity in the Research Laboratory

Productivity – a notion defined with the manufacture of defined uniform products in mind – has been assuming ever greater importance (and nowhere more so than in research and development) since the advent of new technologies in the research laboratories of the natural sciences.

Whereas assembly line methods have, for many decades now, been at the heart of industrial production (because of their capacity efficiently to combine distinct work steps), automation technology is increasingly replacing even these mostly manually performed work steps. Usually the two aspects go hand in hand with efficiency improvements, which is undoubtedly the strongest force driving technological innovation.

In research and development to date, however, other aspects have often loomed larger: creativity, intuition, even the complex sensory co-ordination that researchers need to perform and evaluate experiments, have all impeded increased recourse to technological solutions. In addition, laboratory researchers mainly face such a diversity of tasks that, other than in industrial production, the rationalizing of work steps by gadgetry and machines hardly makes sense. For both reasons, automation technology usually plays a subordinated role here.

Which factors, therefore, are driving the current technological revolution in research? A systematic division of labor for uniform work steps in larger research laboratories, for example in the pharmaceutical industry, is feasible to only a limited extent, and in all likelihood has already passed its high-water mark. With the advent of such systematic approaches as combinatorial chemistry and HTS, however, parallel methods have been rapidly emerging. These require the use of efficient technical gadgetry and machines capable of coping with high sample volumes – thus saving considerable time and money per processed sample. Then too, such technologies as solid-phase synthesis significantly facilitate the handling of multiple parallel reactions, thus moving parallel synthesis within the functional potential of automation. Technical advances in automation technology further expand this functional potential, permitting the gap between the desirable and the possible to be bridged in ever-more laboratory applications. An overview of what is currently technically feasible with automation is provided in the section below.

## 19.4 Automation Concepts: How Multicomponent Systems are Organized

### 19.4.1 General Considerations

Laboratory automation commences with the use of simple laboratory gadgets like horizontal shakers, as even these take over certain manipulations from laboratory workers (in this case, the time-consuming task of shaking flasks during liquid–liquid extraction). This enables technicians to concentrate more on other laboratory activities that require, for instance, constant visual or sensory monitoring. But these too have increasingly fallen, thanks to technological advances, within the orbit of laboratory automation, which has



partly been accompanied by major throughput jumps. A case in point is pipetting robots, which can perform comprehensive pipetting sequences largely on their own, even when deployed for liquid level detection.

Robot development to automate parallel organic synthesis has often evolved from pre-existing methods for peptide and nucleic acid synthesis. For example, the teabag method [13] of peptide synthesis has been automated [14] and improved by introducing new vessels and tagging technique, as well as the multipin peptide synthesis method [15] which has been commercialized (Chiron) and automated [16]. These approaches involve either moving a solid supported reactant to a matrix of different reagents, which must be assembled manually or fully automated [17], or employing commercially available liquid handlers to deliver reagents to the desired wells [18]. The open environment of syringe pump-driven, liquid-handling robots has to be modified such that they can deliver reagents to an environment having temperature and inert atmosphere control. Instruments employing these features include those commercially available from companies such as Advanced Chem Tech (Louisville, KY, USA), Bohadan Automation (Mundelein, IL, USA), MultiSynTech (Bochum, Germany) and Tecan (Research Triangle Park, NC, USA) (Table 19.1). Closed systems that utilize valve blocks to control the flow of reagents to reaction vessels via a common line have been elaborated to manage a larger number of reagents and reaction vessels. This approach is exemplified by several instruments now available (Table 19.1).

The equipment used today for technical support of combinatorial chemistry laboratories range from most simple devices up to highly sophisticated work stations.

About 10 years ago, Boehringer Ingelheim KG [19] and Chiron [20] began to develop proprietary robotic combinatorial synthesizers based on systems developed for the parallel processing of bioassays. Today, all drug discovery companies use commercially available third- or fourth-generation units for automated chemistry that support manual synthesis procedures. Several combinatorial chemistry automation systems are designed for the production of more than 1000 small organic drug-like molecules per day by parallel synthesis. It is obvious however, that large numbers of compounds request simple chemistry and few step reaction sequences. Mostly, the partially automated systems include re-

**Table 19.1.** Suppliers of laboratory automation components and systems

Abimed Analysentechnik GmbH	Martin Christ GmbH
Accelab GmbH	Merck KGaA
Advanced ChemTech	Mettler-Toledo
Argonaut Technologies	Myriad Synthesiser
Bohdan Automation	MultiSynTech GmbH
Chemspeed Ltd	OPAL Jena
H + P Labortechnik GmbH	RAPP POLYMERE
Gebr. HAAKE GmbH	Sartorius GmbH
Hamilton Deutschland GmbH	SYKAM GmbH
Huber GmbH	TECAN AG
Johanna Otto GmbH	Waters-Micromass
Julabo GmbH	Zinsser Analytik
Labsystems GmbH	Zymark GmbH

action blocks containing 96 reaction vessels to allow direct transfer to standard 96-well or multiply-96-well (384, 1152) plates ready for HTS.

Traditional firms for laboratory robotics and several start-up companies, including suppliers of periphery equipment such as vacuo-concentrators, entered the market for professional equipment designed according to the needs of automated chemistry. Reaction blocks for handling of solid-phase as well as solution-phase chemistry have been designed.

To carry out different complex synthesis procedures, reaction blocks that accommodate a wide variety of organic solvents for synthesis, agitation, washing steps and cleavage are necessary. Synthesis procedures must be carried out under temperature control (mainly  $-30^{\circ}\text{C}$  to  $+130^{\circ}\text{C}$ ) and inert gas atmosphere preferably in glass-made filter tubes (solid-phase chemistry) or standard glass vessels (solution-phase chemistry).

Bench-top synthesizers are mostly equipped with one- or two-arm liquid handling high-throughput and combinatorial systems with easy to use Microsoft Windows-based software. Most systems feature simultaneous control of two XYZ arms. This allows the systems to deliver solvents and reagents concurrently, thus exceeding the speed of fluid delivery over similarly equipped single-arm robots. These arms are fitted with septum-piercing needles for fluid handling within sealed vessels. Each needle is connected to a liquid-handling system consisting of standard resolution digital syringe pumps and a valve system for connecting to solvent reservoirs.

Such systems are characterized by the number of manipulating XYZ arms, including septum-piercing devices, the number of syringe pumps and volume of syringes, the number of valves and their way actuation, the number of independent system fluids and the size and numbers of docking platforms for reaction blocks.

In addition to reaction chambers and delivery systems, a number of supervising and sensor systems are of utmost importance for control and safety reasons. Sensors in automated workstations include measurement of temperature (thermocouple, thermistor, semiconductor), pressure, liquid flow and gas or liquid level. To monitor the presence or absence of vessels or devices, systems like capacitance, inductivity, ultrasonic monitors, magnetic sensors or optical sensors (reflective, beam interruption, color) can be integrated in automated workstations.

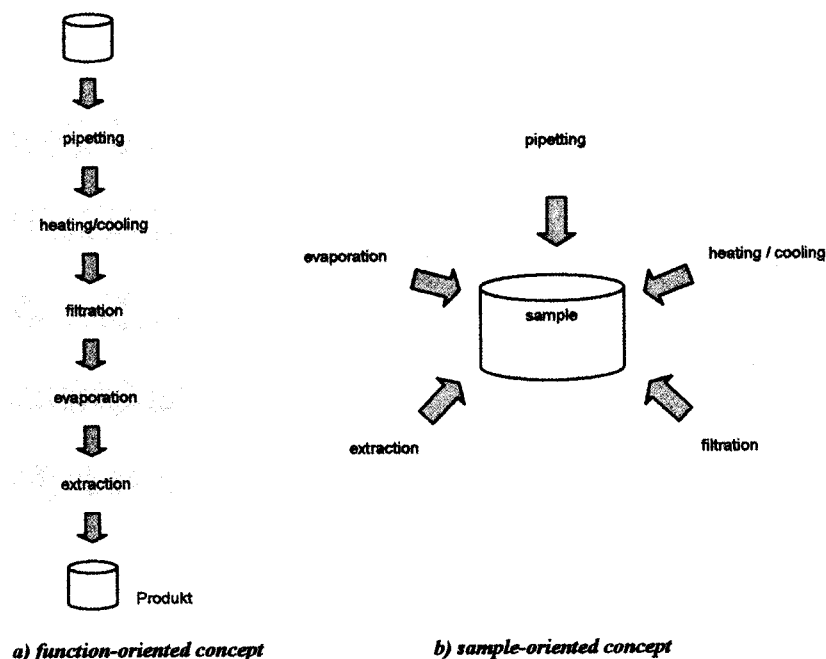
Although all of these different strategies for automation have succeeded to varying degrees, there is a gap between the chemist's needs and the inherent limitations of the technologies upon which automation is based. For example, systems based upon either traditional robotic liquid handlers or valve block systems both require multiple reagents and/or solvents to travel a common path to a reaction vessel, as each reagent does not have its own dedicated delivery system. To reduce the possibility of cross-contamination, extensive washing of needles and lines is utilized. This consumes both excess solvent as well as time, which becomes a substantial factor when attempting to perform chemical reactions simultaneously. Delivering a relatively large number of chemical reagents to an array of, e.g., 96 reaction vessels under an inert environment in a controllable order rapidly, quantitatively, and without cross-contamination is a fundamentally new need created by combinatorial chemistry.

In particular, experts are aware of problems addressed to the existing technological approaches of reagent delivery, work-up procedures, synthesis monitoring, supervising systems including solvent detection.

Thus, the automation of modern laboratories has in part already happened, even if gaps outstanding in the automation continue to create bottlenecks that restrict overall productivity. Plugging these gaps – and so diminishing the problems frequently encountered (e.g., in the sample logistics) when HTS is performed in a series of manual steps – is the paramount goal of integrated automation.

At this point, conceptual considerations gain in importance. Whether users are able to organize discrete specialized gadgets and automated machines into a well-balanced, de-centrally controlled equipment complex amenable to their purposes, or whether such gadgets and machines are best left to be managed and controlled from a central computer, is a question that must be decided in light of expenditure and anticipated benefits. But the fundamental factor holding back full automation in the chemical synthesis laboratory is rather the following: the highly diverse operations performed in chemical synthesis require quite exceptional flexibility, such as is beyond the reach of sequentially operating automated gadgetry.

However, automation systems that can handle the product and process diversity specially required by research laboratories on the basis of sample-oriented automation concepts (Fig. 19.1) – and with high productivity to boot – have recently begun to appear on the market. Equipment pertaining to laboratory automation may therefore, from the perspective of organization, break down into three kinds of systems: decentralized, function-oriented, and sample-oriented.



**Figure 19.1.** Classical function-oriented automation concept compared with central-automated sample-oriented systems.

### 19.4.2 Decentralized Automation Systems

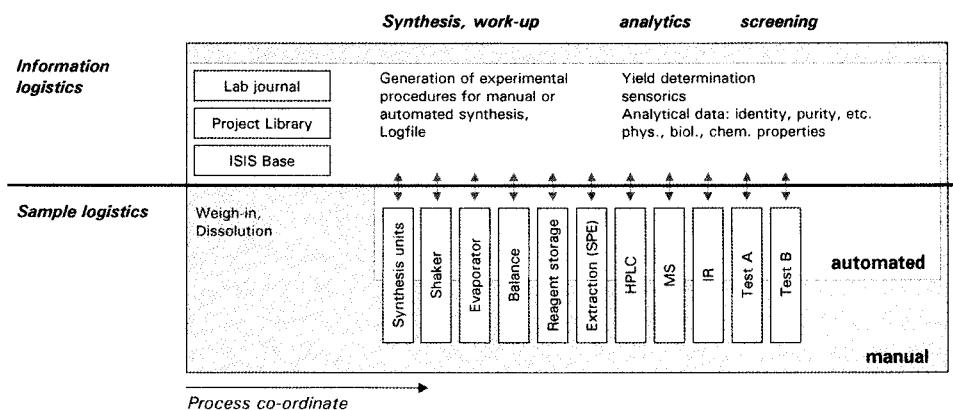
Decentralized systems usually result incrementally as gadgetry which is introduced on a piece-meal basis to automate simple laboratory manipulations. At a later stage, this may come to include automated machines discharging more complex functions (such as pipetting sequences). Typical systems of this kind are laboratories featuring such equipment as the following:

- Pipetting robots to distribute synthesis building blocks, reagents, and solvents
- Incubation stations (mostly shakers, heating and cooling blocks, etc.)
- Evaporators (vacuum centrifuges, vacuum shakers)
- Automated machines for solid-phase extraction
- Autosampler-assisted analytical devices (MS, HPLC, LC-MS, etc.)

When automation along these lines reaches a certain critical mass, employment of gadgetry to capacity no longer grows at the same pace, since the organization of laboratory-internal operations proves increasingly difficult. It thus becomes the task of ‘scheduling’, as it is called, to organize the logical operations for each individual sample or sample batch in such a way that the sequence of work steps is adhered to and the pressure on time and resources minimized. In this procedure, too, documentation of operations and adherence to defined work protocols are two aspects of great significance.

The flexibility of decentralized automated systems, on the other hand, confers an important advantage over and against many integrated systems. The malfunctioning of one of the component gadgets does not bring down the whole system; instead the others operating independently can go on being run, thus raising statistical availability. Another point is that the sources of malfunctioning can by and large be readily located; and they are also easier to repair, as only small units of gadgetry are involved.

The state of play described above under ‘decentralized systems’ is what we encounter in many laboratories today. Not least among the stresses imposed on such laboratories by a high sample volume, however, are organizational bottlenecks, from which result systemic inefficiencies in terms of information flow and employment of technical resources to capacity. Many of these problems can be solved by deploying computer software and automation techniques. Here, ‘data world’ and ‘sample world’ move in parallel over broad stretches, and in their interactions beyond the ‘point of divergence’ lie fertile potentials for overcoming these systemic inefficiencies (Fig. 19.2).



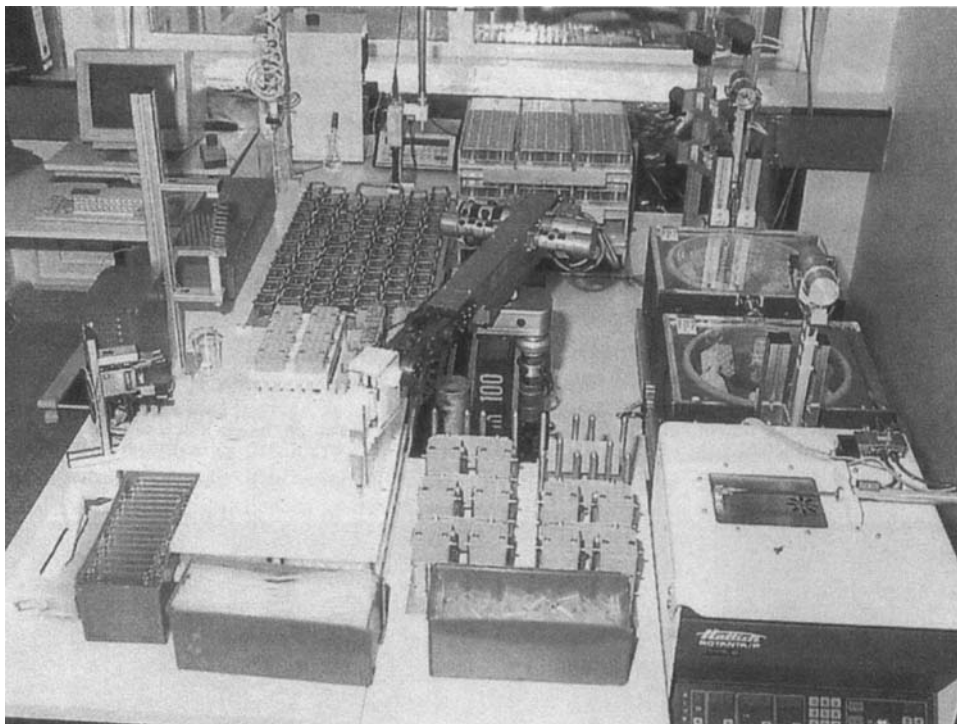
**Figure 19.2.** Resource exploitation in automated laboratories: The different resources (gray boxes) can be used by either automated systems (yellow area) or laboratory personnel (blue area). The process of laboratory work-procedures can be managed by the means of computer technology (information logistics) as well as automation technology (sample logistics).

### 19.4.3 Central-controlled, Function-oriented Multicomponent Systems

The task of taking the decentralized automation system (Fig. 19.3) consists of several mutually independent gadgets and organizing them into a centrally monitored and/or controlled system. However, this task is aggravated by the fact that hooking up the disparate gadgets electronically often requires completely different interface protocols, while suitable software for compartmentalizing discrete processes – a process called ‘scheduling’, as we have seen – is not always available as a matter of course. Without such software, however, employing the system to capacity or better can prove an elusive goal.

The outcome is that typical laboratory automation systems with hierarchical controls, *i.e.*, hooked up to a central master computer, mostly consist of a number of specialized gadgets that are supplied with samples by a central robotic arm. The so-called peripheral gadgets are, for example, standard laboratory equipment which has been subsequently adapted, where necessary, for use in the system in question; or else they have been specially built for that system. Here, automated functionality has largely been devolved to these peripheral gadgets, while the central robot only handles the transporting of samples between stations. This method is taken from HTS, with the sample bearer role mostly devolved to microtiter plates (MTP), these being transported by, *e.g.*, pipetting stations to incubators and ELISA readers. In cases of routine repetition of uniform sequential steps and simultaneous processing of several samples (96 or 384 per MTP), this approach is perfectly rational. New samples can, without undue effort, be introduced into the operation, since only a few parameters alter from run to run.

However, the variable operations encountered in combinatorial synthesis, particularly when we consider purification and solution-chemistry work steps, place raised demands on the flexibility of such a system, since a multiplicity of sample-specific processes and pa-

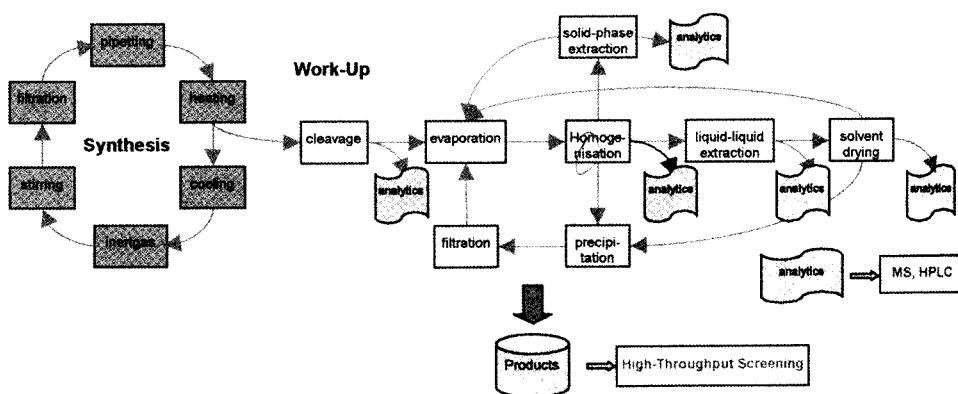


**Figure 19.3.** A typical function-oriented multicomponent system. A central computer system controls all the components, including the robot arm, which performs transport tasks to carry samples among the other components. (Illustration courtesy of ISRA Systemtechnik GmbH.)

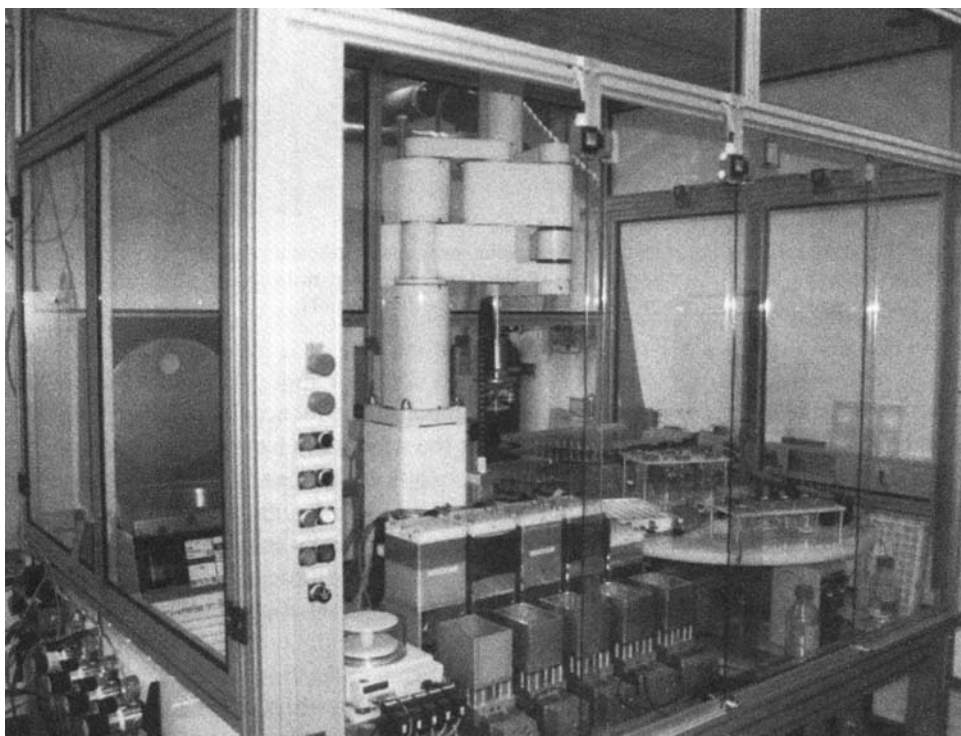
rameters must be secured. Thus, for example, pipetting sequences at various peripheral stations are indispensable and must, so long as the robot only handles sample transport, be suitably automated by several pipetting installations. Unless each peripheral gadget with a pipetting function can be additionally equipped with a manipulator with its own electronic controls, the central robotic arm has to individually transport samples to each pipetting station. Here, the advantage of multiple sample processing gives way before slower single-sample processing.

#### 19.4.4 Central-automated, Sample-oriented Multicomponent Systems

The classical function-oriented automation concept forfeits efficiency the instant a diversity of process and products are targeted. On the other hand, by centralizing automation, *i.e.*, by devolving automation tasks away from the periphery to a powerful central robot, highly variable processes can be tackled, *e.g.*, in combinatorial synthesis. One feature of these so-called sample-oriented systems is that the pipetting function is centralized.



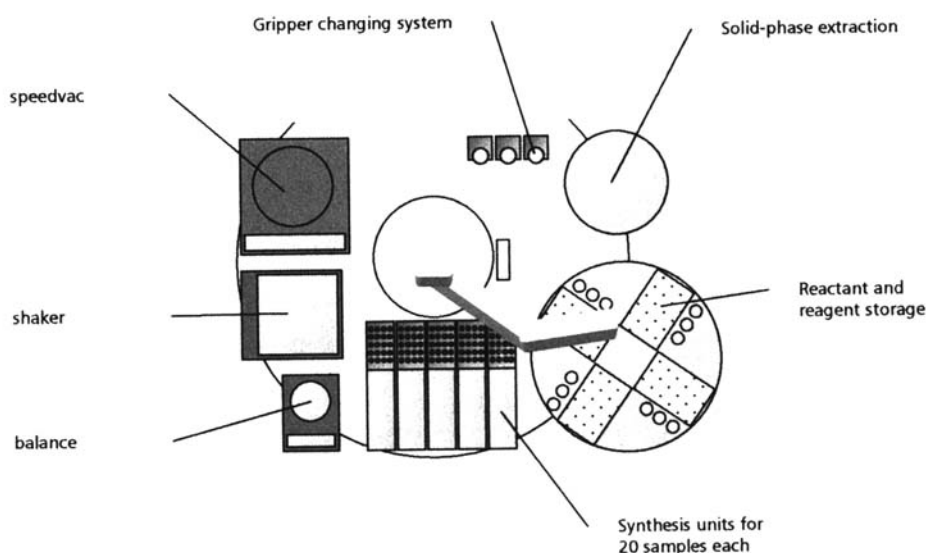
**Figure 19.4.** Individual functional networking for single samples or sample batches in a sample-oriented system with a view to combinatorial synthesis.



**Figure 19.5.** In the course of developing the robotic platform ARCoSyn for fully automated synthesis and purification of compound arrays, a sample-oriented concept has been realized that subordinates sample functionality, thus avoiding complex transport processes between spatially separated individual functionalities. The central component is an industrial robot, which – in respect of flexibility (the gripper changing system for several centralized functions), work space (optimal utilization of available surface area, no need for a translation axis), precision, and loading capacity (option of using modules for both miniaturization and upscaling) – is adapted to the requirements of the laboratory automation concept.

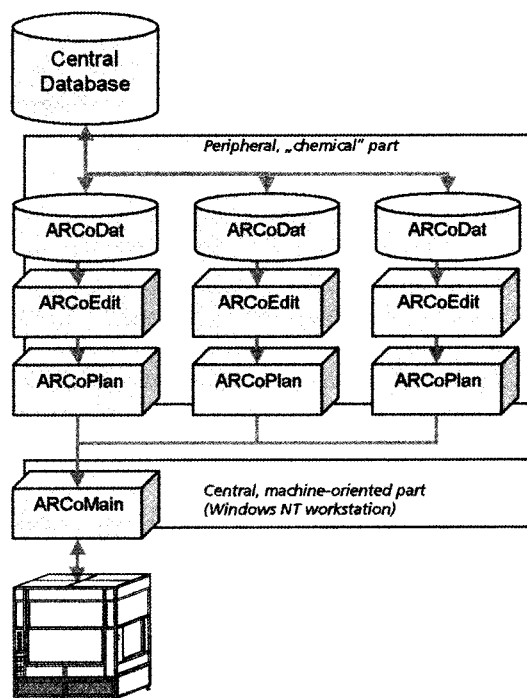
What additionally typifies this concept is that it permits close networking between automated and manual utilization of resources – users and automation system can, as a rule, use the integrated laboratory gadgetry at one and the same time, which results in a highly open and flexible structural set-up. The individual functional networking for single samples or sample batches in a sample-oriented system with a view to combinatorial synthesis is shown in Figure 19.4.

To address these needs directly, we have developed a novel automation concept, ARCoSyn, which is based on a central robot with high accuracy and weight capacity. This technology includes also newly developed reactor blocks and vent systems. A series of different grippers are designed for replacing also complex otherwise manually performed steps (Fig. 19.5). With this technology as its core, we have designed and built a highly flexible instrument which is applicable not only for combinatorial solid-phase organic synthesis, but also for solution-phase synthesis, as well as biotechnological applications such as parallel production and purification of recombinant proteins or other fermentation processes. The newly developed supervising and chemistry software (Figs. 19.6–19.10) takes advantage of the unique machine properties. The basic features of the two-dimensional parallel array synthesizer and associated software control systems are described.



**Figure 19.6.** Further components spun off from automation and laboratory technologies were combined to form an automation platform with a surface area of  $2.5\text{ m} \times 1.8\text{ m}$ , into which modules were integrated for parallel synthesis and purification of 100 compounds. As far as possible, standard laboratory gadgets were deployed (in part slightly modified). Modules such as the synthesis units used for solid-phase synthesis by heating/cooling/stirring/inert gas on a 50-mg scale are, like the other components, interchangeable with other systems. Thus, the underlying automation platform takes over the functions of a large part of the gadgetry line-up of a conventional laboratory.





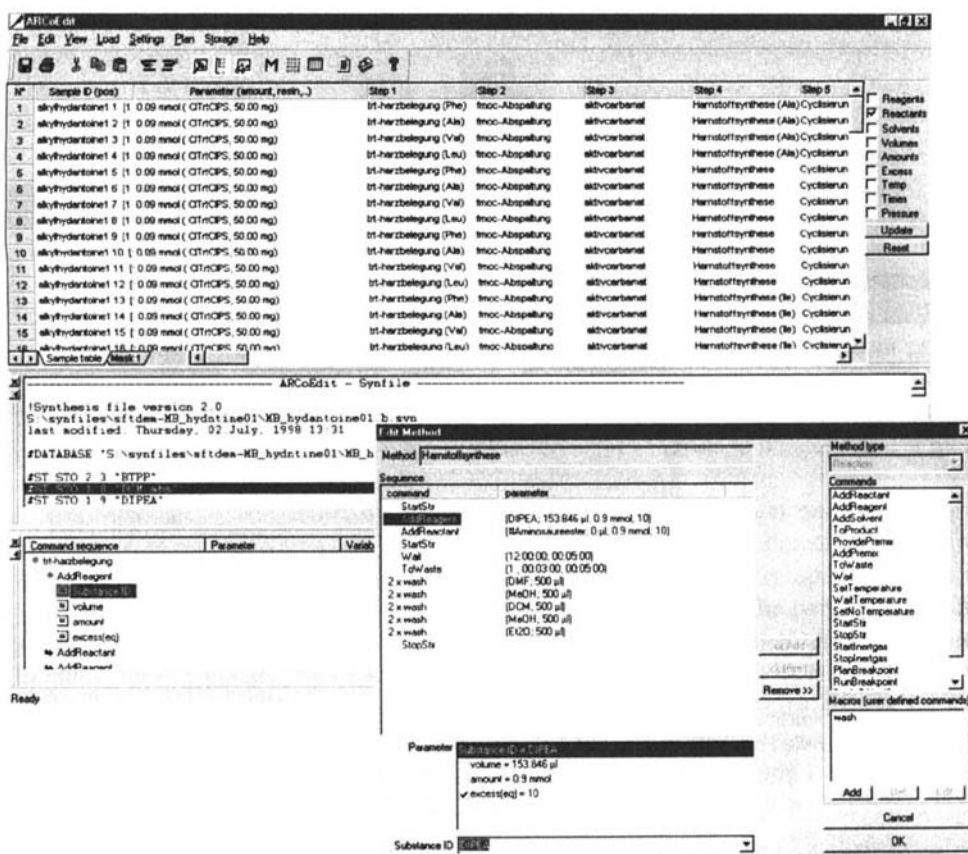
**Figure 19.7.** Subjecting the hardware/software interface to robotic control proceeded according to the criteria of multitasking capacity, facilitation of user interventions, and the number of interfaces and digital I/O paths. The control and planning software also reflects the sample-oriented concept of the workstation, inasmuch as the user can dispense with knowing the technical details of the systems software and simply program their operations in the manner of typical laboratory protocols. The central building block is a sample table in which are stored all operations and sample-specific parameters. To use the off-line building blocks of the software package, no more than a basic grasp of systems configuration is required. The building blocks of entry, planning, and administration of syntheses can also be installed decentrally, thus permitting user-specific synthesis management. The control and monitoring module, by means of which programmed operations and individual gadgetry in the system are controlled and deployed, forms the central machine-oriented component, which is itself run on the master PC that is directly hooked up to the system.

The screenshot shows the **MD\_hydantoin01.asy - DataEditor** software interface. The main window displays a table of substances with columns for name, ID, concentration, solvent, and temp. range. A detailed property window is open for the substance **4-Fluorophenylisocyanat** (ID: 119545).

name	ID	concentration	solvent	temp. range
1,4-Diisocyanato-butan	38215	7.95 mol/l	none	1..230 °C
2,4-Difluorphenyl-isocyanat	36952	8.38 mol/l	none	1..42 °C
2,6-Diisopropylphenyl-isocyanat	38437	4.97 mol/l	none	1..117 °C
2-Chloroethyl-isocyanat	23040	11.88 mol/l	none	1..136 °C
3,4-Dichlorphenyl-isocyanat	36221	not defined	none	1..41 °C
3,5-Dinitrophenyl-isocyanat	42220	not defined	none	1..82 °C
3-(Trifluormethyl)-phenylisocyanat	91			
3-Chlor-phenylisocyanat	21			
3-Isocyanatopropyl-triethoxysilan	51			
3-Methoxy-phenylisocyanat	61			
4-(6-Methyl-2-1-Diisocyanato-1,3,5-triazin-2-yl)isocyanat	61			
4-(Trifluormethyl)-phenylisocyanat	91			
4-Chlor-phenyl				
4-Fluor-phenyl				
4-Methoxy-phenyl				
4-Nitrophenyl				
Allyl-isocyanat				
Benzoisulfonyl				
Benzylisocyanat				
Chlorisulfonyl				
Cyclohexylisocyanat				
Ethyl-isocyanat				

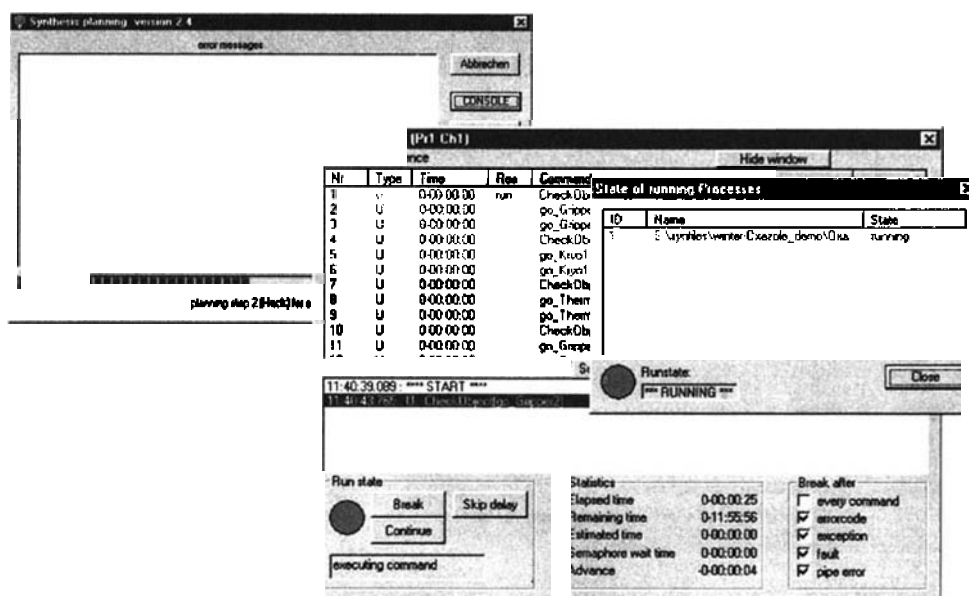
The detailed property window for **4-Fluorophenylisocyanat** shows the following details:

- ID:** 119545
- assignment:** DEHYDRATING\_AGENTS, AA\_METHYL\_ESTERS, ABSOLUTE\_SOLVENTS, ALKYL\_HALIDES, BASES, ISOCYANATES
- min. temp.:** 50 °C
- max. temp.:** 50 °C
- molecular weight:** 137.11 g/mol
- concentration:** 8.90 mol/l
- density:** 1.22 g/cm<sup>3</sup>
- solvent:** none
- protection liquid:** system liquid
- waste allowed:** Waste 1, Waste 2, Waste 3, Waste 4, Waste 5, Waste 6, Waste 7
- options:** use inertgas, moisture-sensitive, Dätor slow



**Figure 19.9.** The ARCoEdit module. Operations are sample-specifically stored in a sample table, meaning that in each sample each operational step can be tracked and edited by the user. Such a step comprises what is called a 'method', containing the requisite experimental procedures for a synthesis or purification step. This permits methods to be defined whose parameters can still be varied from sample to sample as a source of diversity. The pipetting volumes are calculated by the ARCoEdit module from the operant stoichiometries, as well as from the substance data stored in the substance database.

- ◀ **Figure 19.8.** With the ARCoDat module, the users can build themselves synthesis-specific data sets for end-stage compounds. Database interfaces permit substance data to be imported for stoichiometric calculations and plausibility/control. Search functions allow building block ensembles to be imported from, e.g., in-house central databases such as Oracle and MS-Access or files like SD-Files or Excel spreadsheets.



## 19.5 Collateral Technologies for High-Throughput Purification and Analysis

High-throughput organic synthesis created new bottlenecks in work-up procedures and analytical characterization of organic compound collections. Additionally, the connection of extremely high volumes of data from analytics with structure information, compound storage information and screening data will become a throughput-limiting step in hit finding and validation and lead discovery and optimization. The analysis of compound arrays has been achieved by extended automated instruments, reduction of separation or detection time.

Data from HPLC, MS, NMR and IR have to be stored in standard analytical file formats and transferred to a generic data viewer connected to standard databases that are valuable in drug discovery.

The information management also affords special and extensive efforts for high-throughput production of large numbers of different scaffolds and compounds. The registration of a compound collection before synthesis helps to generate and organize data necessary for the synthesis process. A series of reaction steps, reaction conditions and appropriate building blocks are registered in the primary database. Clusters of data are listed for the compounds to be generated, including the proprietary ID number.

Reagent management (for individual reagents or cocktails of reagents!) includes reagent names and physico-chemical parameters, as well as synthetic procedures. The synthetic protocol is usually identical for all members of a collection or, dependent on the automation equipment, subcollections are specifically addressable.

The purification and characterization of the individual compounds produced by automated combinatorial chemistry – which earlier on was pushed into the background by rational new synthesis techniques – have in recent years been regaining importance. The reason lies in the growing frequency of samples declared false-positive, which after primary biological screening were fed into the time-consuming processes of secondary screening and lead structure optimization. Unequivocally characterized, highly pure compounds without unknown by-products enhance the indicativeness of hits from primary screening.

Today it is possible, using modern separation and analysis techniques (Chapters 17 and 20), for the desired compound at high sample throughput levels to be identified simultaneously and filtered out of a compound mixture.

High-throughput preparative HPLC coupled to electrospray ionization mass spectrometry (Chapter 17), which disposes upon a signal for collecting detected compounds of the defined molecular mass, is one of the highly promising new developments in this area. Such systems can be incorporated for synthesis purposes into the periphery of automated multicomponent systems, thus making a valuable contribution to the rationalization and quality enhancement of combinatorial synthesis processes. The combination of automated synthesis, purification and on-line instrumental identification (NMR, IR, MS) will become feasible in the near future, and as a matter of routine operation. Analytic methods of structure elucidation will then also be able to be combined with automated combinatorial chemistry.

## 19.6 Summary and Prospects

Automated combinatorial chemistry has, over the past ten years, firmly established itself as an autonomous chemical discipline. Ensembles of highly diverse compounds in the substance libraries of the chemical industry constitute a valuable reserve for developing new active substances.

However, the required high throughput in screening for new biologically active compounds, catalysts and materials as favored by combinatorial chemistry and modern assay developments throws up new questions. Only by incorporating modern information processing, automation, miniaturization and microfluid research, (not to mention modern separation, analysis and detection methods) will it be possible to master the future challenges posed by the search for new active substances. Expanding the techniques of combinatorial chemistry to areas other than drug finding enlarges the repertoire of diverse compounds and materials requiring characterization.

Combinatorial technologies have, in recent years, been implicated pivotally in tabling novel strategies for identifying new materials or catalysts – strategies that are also deployable in the search for biopolymers with outlets in therapy or large-scale technology.

Parallel generation, processing, and characterization of large numbers of discrete compounds (*i.e.*, ranging from many thousands to hundreds of thousands) now appears a goal attainable in the not too distant future. Required above all is efficient networking between the implicated factors from synthesis, analysis, and the biological sciences. And such networking must (in abstract terms) be articulated in an overarching organizational concept. This involves information processing on the one hand, whose task it is to integrate theoretical methods into the development process; and automation on the other, which will play an analogous role on the technical level, this by making available an efficient sample logistics plus technical methods capable of high throughput.

Chemistry, the biological and material sciences, information technology and automation technology – these are the fields that will supply the technological resources to drive this development. Moreover, it is to this pool of knowledge we may look, in coming years, for more innovations such as are already achieving expression today in the form of new ideas, for instance on miniaturized synthesis and test methods or the realization of efficient – and at the same time highly variable – integrated automation systems.

## 19.7 Acknowledgements

The development of the workstation ARCOSyn has been supported by Boehringer Mannheim GmbH, Boehringer Ingelheim KG, Hoechst AG, Merck KGaA, Schering AG and the Bundesministerium für Forschung und Technologie ('Automated combinatorial chemistry', Grant 03 D 0037; Coordinators: G. Jung, K.-H. Wiesmüller). In addition, we are especially grateful for the continuous cooperation of the following scientists: Prof. Dr. Austel, Dr. Buchstaller, Dr. Devant, Prof. Dr. Gruhler, Dr. Heck, Dr. Henke, Dr. Köppen, Dr. Künzer, Dr. Mertens, Dr. Peters, Dr. Prien, Dr. Raddatz, Dr. Rölfing, Dr. Sauer, Dr. Schöllkopf, Dr. Schnorrenberg, Dr. Stadtmüller, Dr. Stähle, Dr. Tiebes, Dr. von Roedern, Prof. Dr. Zell, and Dr. Zimmermann.

## 19.8 References

- [1] Combinatorial libraries, synthesis, screening and application potential (1995) Cortese R (ed), de Gruyter, Berlin, Germany, ISBN 3-11-014395-X.
- [2] Combinatorial peptide and nonpeptide libraries – a handbook for the search of lead structures (1996) Jung G (ed), VCH-Wiley-Weinheim, Germany, ISBN 3527-2938-9.
- [3] Molecular diversity and combinatorial chemistry libraries and drug discovery (1996) Chaiken IM, Janda KM (eds); American Chemical Society, Washington, DC, USA, ISBN 0-8412-3450-7.
- [4] Combinatorial chemistry, synthesis and application (1997) Wilson SR, Czarnik AW (eds); VCH-Wiley, Weinheim, ISBN 0-471-12687-X.
- [5] Annual reports in combinatorial chemistry and molecular diversity, vol. 1 (1997) Moos WH, Pavia MR, Ellington AD, Kay BK (eds); ESCOM Science, NL-Leiden, ISBN 90-72199-23-5.
- [6] Methods in molecular biology, vol. 87 (1998) Combinatorial peptide library protocols, Cabilly S; Humana Press, Totowa, NJ, USW, ISBN 0-89603-392-9.
- [7] The combinatorial index (1998) Bunin BA; Academic Press, ISBN 0-12-141340-3.
- [8] Combinatorial chemistry and molecular diversity in drug discovery (1998) Gordon EM, Kerwin JF Jr. (eds), Wiley-Liss, Weinheim, Germany, ISBN-0-471-15518-7.

- [9] Gordon EM, Barrett RW, Dower WJ, Fodor SPA, Gallop MA (1994) Applications of combinatorial technologies to drug discovery, 2: Combinatorial organic synthesis, library screening strategies, and future directions, *J Med Chem* 37: 1385–1401.
- [10] Früchtel JS, Jung G (1996) Organic chemistry on solid supports, *Angew Chem* 108:19-46, *Angew Chem Int Ed Engl* 35: 17–42.
- [11] Thompson LA, Ellman JA (1996) Synthesis and applications of small molecule libraries, *Chem Rev* 96: 555–600.
- [12] Oldenburg KR, Zhang J-H, Tongming C, Maffia A, Blom KF, Combs AP, Chung TDY (1998) Assay miniaturization for ultra-high throughput screening of combinatorial and discrete compound libraries: a 9600-well (0.2 microliter) assay system, *J Biomolecular Screening* 3: 55–62.
- [13] Houghten RA (1985) General methods for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids, *Proc Natl Acad Sci USA* 82: 5131–5135
- [14] Beck-Sickinger AG, Dürr H, Jung G (1991) Semiautomated T-bag peptide synthesis using 9-fluorenyl-methoxycarbonyl-strategy and benzotriazol-1-yl-tetramethyl uronium tetrafluoroborate activation, *Peptide Res* 4: 88–94.
- [15] Maeji NJ, Bray AM, Geysen HM (1990) Multi-pin peptide synthesis strategy for T cell determinant analysis, *J Immunol Methods* 134: 23–33.
- [16] Gombert FO (1998) Semi-automated HTP preparation of libraries: single compounds using pins bearing radiofrequency tags, IBC conference on high throughput compound characterization and purification & integrating combinatorial chemistry into the discovery pipeline, Sept. 14–15, Arlington, VA, USA.
- [17] Wiesmüller KH, Treffer U, Spohn R, Jung G (1993) Fully automated synthesis of peptides on polyethylene pins, In: Eberle AN, Schneider CH (eds) *Peptides 1992*: 308–309. Escom, NL-Leiden.
- [18] Gausepohl H, Boulin C, Kraft M, Frank RW (1992) Automated multiple peptide synthesis, *Pept Res* 6: 315–320.
- [19] Schnorrenberg G, Gerhardt H (1989) Fully automatic simultaneous multiple peptide synthesis in micromolar scale – Rapid synthesis of series of peptides for screening in biological assays, *Tetrahedron* 45: 7759–7762.
- [20] Zuckermann RN, Kerr JM, Siani MA, Banville SC (1992) Design, construction and application of a fully automated equimolar peptide mixture synthesizer, *Int J Pept Protein Res* 6: 497–506.

This Page Intentionally Left Blank

# 20 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS) for Analysis of Compound Libraries

*Tilman B. Walk, Axel W. Trautwein, Holger Bandel and Günther Jung*

## 20.1 Introduction

Combinatorial solid phase organic chemistry [1–4] is an established key technology in the lead finding process in the pharma research. The obvious advantage is the higher throughput compared to conventional methods, because automation assisted procedures [5] can be used without workup of intermediates. The resulting products of parallel synthesis can consist of collections of defined single compounds or defined mixtures of compounds (libraries). The size of a library ( $A^nB^n$ ) is determined by the number ( $n$ ) of reactants and the number of reaction steps. The exponential relation between the size of the library and the number of reaction steps (e.g. 4) leads even with low numbers of educts (e.g. 10 per step) to a large number of endproducts (10000).

Assays for testing these compound libraries are automated to a high degree and optimized for reproducible recognition of *in vitro* activities at the lowest possible sample consumption. High-throughput screening of many thousands of compounds per day is state of the art [6]. These developments have forced the development of high-throughput technologies for analysis. Higher efficiency has been achieved in the “off-bead” analysis by ES/MALDI-TOF/quadrupole-mass spectrometric systems [7–9], RP-HPLC [10] and in the “on-bead” analysis by NMR [11] and FT-IR [12] spectroscopy. Previously less utilized methods such as FT-IR microscopy even allow exact and simultaneous “on-bead” analyses of polymer-bound compound collections consisting of hundreds of resin beads with one single measurement [13].

The preferred use of MS methods is due to the low sample consumption, fast sample preparation, and relatively simple automation. MS is the method of choice for analyses of libraries.

As shown in the following ES-FT-ICR mass spectrometry sets new standards with respect to effectiveness, informative value, and sample consumption [15].

The principles of ion cyclotron resonance (ICR, 1950) [16], its correlation with the mass of the ion (1951) [17], the prerequisites of a measuring cell (1970) [18], and the application of Fourier transformation for data acquisition (FT-ICR, 1974) [19, 20] are long known. The first FT-ICR mass spectrometer was shown by Comisarow and Marshall in 1974 [20]. Stronger superconducting magnets [21], faster computers and the introduction of ES and MALDI ion sources [22] gave rise to the youngest generation of commercial FT-ICR mass spectrometers (Companies: IonSpec; ThermoQuest/Finnigan; Bruker Daltonik). In 1998 the FT-ICR-MS equipped with ES, Nanospray, MALDI and EI ion sour-



ces and on-line HPLC or CE was installed in the Institute of Organic Chemistry at the University of Tübingen. The computer equipments allow the registration of high resolution mass spectra in a mass range of 4000 Da within less than 200 ms. In combination with the modern ion sources the coupling with separation techniques is possible (HPLC, CE, CZE, etc.) [23, 24]. With respect to mass accuracy and resolution FT-ICR-MS is by far better than other mass detectors (TOF, quadrupol, sector field).

First an introduction to the physical parameters of the FT-ICR method is presented, which has been used seldomly for routine analysis. This shows the possibilities but also the limits of the mass detection method. Then the immense potentials of ICR-MS for analysis in combinatorial chemistry will be demonstrated by examples of measurements of synthetic compound mixtures and simple compound collections. The application of the coupling of micro-HPLC and ES-FT-ICR mass spectrometer underlines the universal applicability of this powerful detection method for single bead analysis.

## 20.2 Physical Basis of FT-ICR Mass Spectrometry

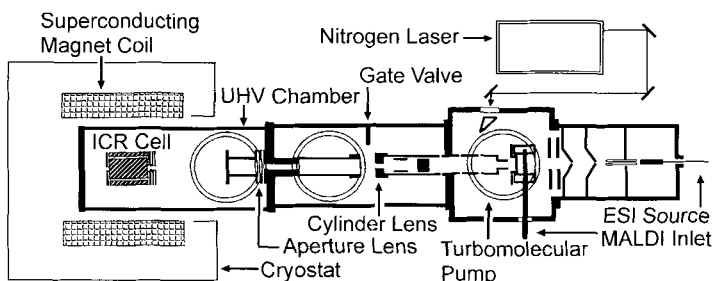
### 20.2.1 Instrument

The instrumental set-up consists of units for ion generation, ion transfer and the analyzer cell. All ionization methods (ES, FAB, FD, MALDI, EI, etc.) can be combined with FT-ICR mass detection (Fig. 20.1).

#### 20.2.1.1 Superconducting Magnet

The ICR cell is surrounded by a superconducting magnet similar to those common in NMR spectrometers. The magnet systems of commercially available FT-ICR mass spectrometers have magnets with field strengths of 3.0, 4.7, 7.0 or 9.4 Tesla.

Since the B-field must be parallel to the ion beam, the cryomagnet system requires a horizontal setup in contrast to that of NMR spectrometers. The magnetic fields of all



**Fig. 20.1.** Scheme of the instrumental setup for the Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (FT-ICR-MS)

cryomagnets used in ICR-MS are efficiently shielded because of the necessity of closely positioned auxiliary equipments such as pumping units of the high vacuum system and HPLC. Shielding can be done passively by iron plates or actively by a second cryomagnet compensating the magnetic field of the first one. Outside the 5 Gauss line one can work safely with magnetizable materials (1 Gauss = 0.1 milli Tesla).

### 20.2.1.2 Vacuum System

The difference in pressure between ion generation (atmospheric pressure, approx. 1013 mbar) and ICR cell ( $5 \times 10^{-10}$  mbar) requires a three-stage differential vacuum system (1. stage:  $10^{-5}$  mbar; 2. stage:  $10^{-8}$  mbar; 3. stage:  $10^{-10}$  mbar). The continuity of the pressure level during the measurement is very important because the quality of the measurement depends on a carefully controlled vacuum.

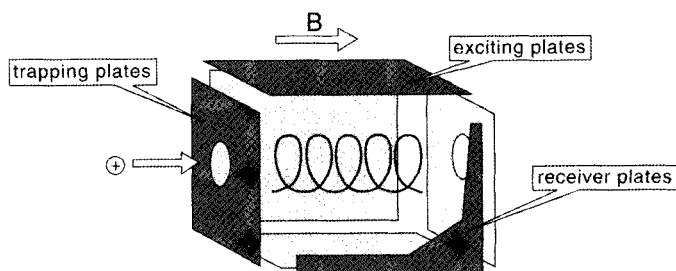
Pumping is performed by cryo or turbomolecular pumps (1200 l/s). For collision induced dissociation (CID) experiments several gases (Ar, CO<sub>2</sub>) can be introduced via pulsed magnetic valves, which may also be used for introducing reactive gases. Thus, reactions of ions trapped in the analyzer cell with gases can be analysed by ICR mass spectrometry including the determination of reaction kinetics [25, 26].

High vacuum locks with separately pumped prevacuum allow the introduction of samples for MALDI or laser desorption (LD) ionisation experiments.

The complete vacuum unit can be baked out at 200 °C in order to remove the remainder of the analytical samples introduced. An efficient water cooling system is necessary to have optimal temperature conditions for the turbomolecular pumps (90000 rpm) and the analyzer cell within the cryomagnet.

### 20.2.1.3 Analyzer Cell

The analyzer cell of the ICR mass spectrometer can store and analyze the generated and transferred molecular ions. A number of different cell geometries (cubic, cylindrical, hyperbolic, ...) have been developed to maximize the quadrupole character of the cell to achieve higher repetition rates for remeasurements, better resolution and mass accuracy



**Fig. 20.2.** Cubic analyzer cell of FT-ICR mass spectrometer

and a reduction of axial ion loss in B-field direction [27]. Today cubic analyzer cells are used in FT-ICR mass spectrometers, which are built up from six plates. The plate pair perpendicular to the B-field function as trapping plates while the remaining two plate pairs parallel to the B-field function as excitation and detection plates receptively. Holes in the trapping plates allow the entrance of ions into the cell while the trapping potential is reduced. Upon application of a radio frequency voltage (5 kHz to 5 MHz) to the excitation plates ion packages of uniform mass are formed. The cyclotron frequency of each single ion package can be measured via the receiver plates.

## 20.2.2 Detection Principle

### 20.2.2.1 Cyclotron Motion

ICR-MS [17] is based on the phenomenon of cyclotron motion of the ions in the analyzer cell. This motion is caused by a combination of magnetic (B-) and electric (E-) fields (trapping plates). The ions are forced to circulate as a result of the Lorentz force (equation 1: m, v, q are mass, velocity and charge of the ions).

$$F_{Lorentz} \equiv m \frac{dv}{dt} = qE + q(v \otimes B) \tag{1}$$

Supposing that the B-field extends in z-direction of the cartesian coordinate system, that the influence of the E-field is negligible and considering the angular acceleration  $dv/dt = v_{xy}^2/r$  then equation 2 results from 1 as follows:

$$\frac{mv_{xy}^2}{r} = qv_{xy} B_0 \tag{2}$$

The angular velocity  $w = v_{xy}/r$  is defined by the ratio of the ion velocity in the xy plane ( $v_{xy}$ ) and the radius of the cyclotron motion (r). Therefore from 2 follows 3 and 4:

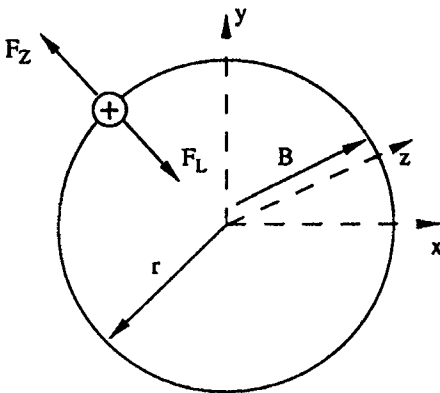


Fig. 20.3. Effect of the Lorentz force and the zentripedal force on ions in the analyzer cell

$$m\omega^2 r = qB_0 \omega t \quad 3$$

or

$$\omega_c = \frac{qB_0}{m} \quad 4$$

Equation 4 is the cyclotron equation,  $\omega_c$  corresponding to the cyclotron frequency (ICR frequency). The ICR frequencies are between a few kHz and several MHz. Equation 4 indicates that all ions with the same ratio of ion charge to ion mass possess the same ion cyclotron frequency. In contrast to other mass analyzers (magnetic/electric sector field, time-of-flight (TOF), quadrupol), the ion velocity has no direct influence on the relation between the measured value and the ion mass (ICR: frequencies  $\omega_c$  in equation 4; sector field; radius of deflection  $r$  for magnetic selection in equation 5; TOF: flight time  $t$ ).

$$r = \frac{mv}{qB} = \frac{1}{B(2mU/q)^{1/2}} \quad 5$$

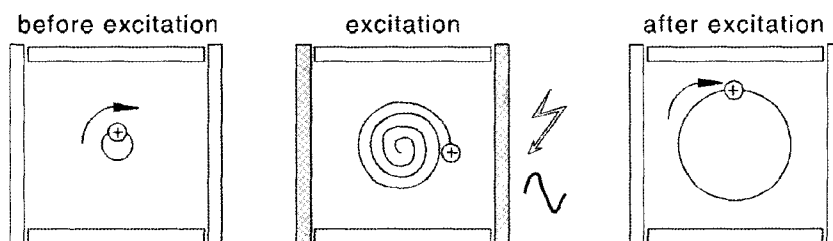
For exact results it is not necessary to equalize the ion velocities during ICR measurements. Ion velocities have no influence on the results of the measurements, this is the basis for the higher performance of FT-ICR-MS compared to other detection methods.

#### 20.2.2.2 Signal Generation

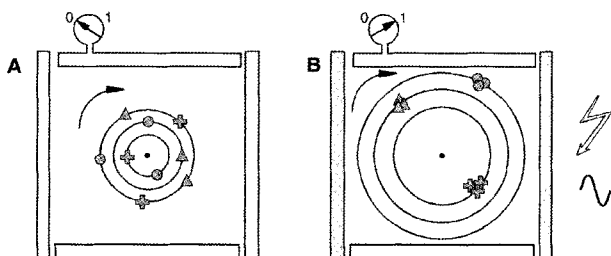
Ions in the analyzer cell are distributed statistically. Because of the broad band irradiation of high frequency oscillating fields via the exciting plates, all ions are excited in resonance [28]. The absorbed energy results in an increased orbital radius (Fig. 20.1).

As a result of the complex relationship between left and right circulating electric field vector of the oscillating electric field of the exciting plates and the B-field ion packages are formed within the ICR cell. These packages consist of ions of the same  $m/z$  ration (Fig. 20.2).

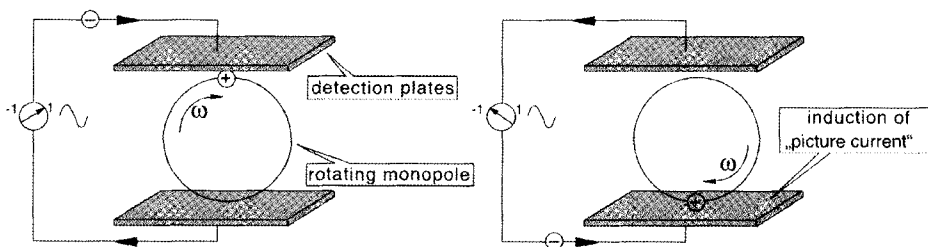
The formation of ion packages induces alternating currents called picture current. The frequencies of the alternating potentials correspond to the cyclotron frequencies of the ions in the ICR cell. Using equation 6 the corresponding ion masses can be calculated.



**Fig. 20.4.** Scheme of the cyclotron movement of single ions, before, during and after exciting by a high frequency changing field. (The magnetic B-field flows vertically to the drawing layer.)



**Fig. 20.5.** After the excitation (**B**) ions with the same  $m/z$  ratio form ion packages. The velocity of the ions and the cyclotron radius are changed, but the cyclotron frequency  $\omega$  is unmodified.



**Fig. 20.6.** Induction of the picture current on the detector plates in the analyzer cell released by rotating ion packages.

With time, the coherence of the ions diminishes due to ion/ion and ion/neutral molecule interactions. The resolution  $m/\Delta m$  is improved proportionally to the prolongation of the signal fade constant  $\tau$  ( $B$ : magnetic field strength;  $q$ : ion charge;  $m$ : ion mass).

$$\frac{m}{\Delta m} = \frac{B\tau q}{2m} \quad 6$$

The intermolecular interaction can be minimized by high vacuum. That means the gain in  $\tau$  prolongation directly correlates with the results of ICR measurements. The longer the coherence of the ions (the higher  $\tau$ ), the longer the ion cyclotron frequencies can be measured and the better is the accuracy and resolution of the measurement.

The intensity of the amplitudes of the exciting frequencies ( $A_{rf}$ ) and the duration of the irradiation ( $t$ ) determine the orbital radii ( $r$ ) occupied by the ion packages (equation 7). High exciting power can lead to discharge of the ions on the detection or excitation plates of the ICR cell.

$$r = \frac{A_{rf}t}{B} \quad 7$$

### 20.2.2.3 Mass Accuracy

The mass accuracy is defined as the ratio of the absolute value of the difference between the measured mass and the true mass to the true mass, multiplied by  $10^6$  (equation 7). The mass accuracy is given in ppm.

$$\text{mass accuracy} \equiv \frac{m_{\text{found}} - m_{\text{calcd}}}{m_{\text{calcd}}} \times 10^6 \quad 8$$

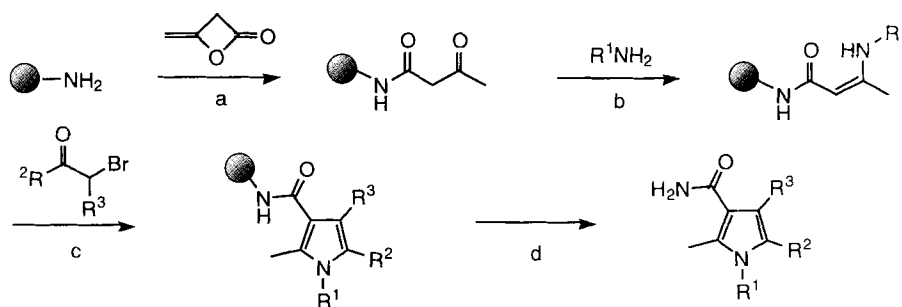
### 20.2.2.4 MS<sup>n</sup>-Experiment

Trapped ions can be stored for between minutes and several hours within the analyzer cell of ICR mass spectrometers. Unwanted ions can be selectively removed by irradiation with resonant alternating fields, thus leaving only the desired parent ions in the cell. These parent ions can be dissociated by collision with inert gases (CID: collision induced dissociation, or CAD: collision activated dissociation) [29]. The dissociation can also be achieved via UV or infrared light (UVDD, ultraviolet photodissociation [30]; IRMPD, multiphon infrared photodissociation [31]) The resulting daughter ions are either measured, or can be selected for further fragmentation. This cycle of selection-fragmentation-measurements can be carried out up to four times (MS<sup>4</sup>) [32].

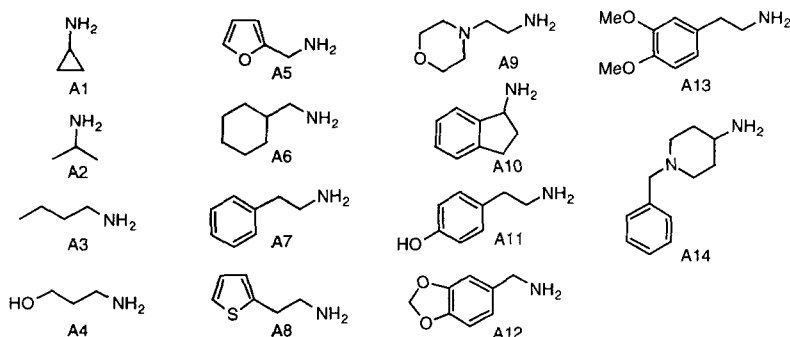
## 20.3 FT-ICR-MS Analysis

### 20.3.1 Combinatorial Compound Collections

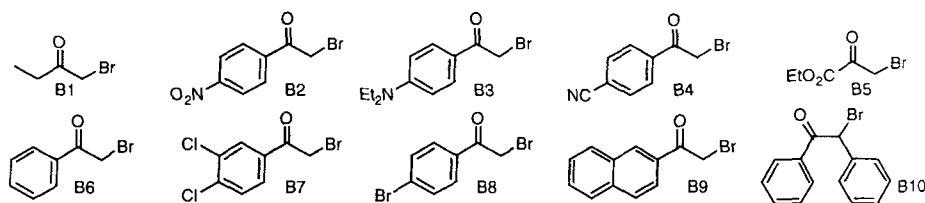
As pointed out in the introduction, test systems for the search of biologically active substances are largely automated and designed for reliable recognition of *in vitro* activities with minimal sample consumption (high-throughput screening).



**Fig. 20.7.** Solid-phase synthesis of pyrrole amides by the split-mix method: **a.** acetoacetylation with diene; **b.** separation of the resin into 14 aliquots, enaminone formation with 14 different primary amines; **c.** combination of the 14 aliquots, separation into 10 aliquots, reaction with 10 different  $\alpha$ -bromoketones; **d.** cleavage from the polymer support. [33]



**Fig. 20.8.** Structure of the amines A1–A14 ( $R^1$ ) used for the synthesis



**Fig. 20.9.** Structure of the  $\alpha$ -bromoketones B1–B10 ( $R^2$ ,  $R^3$ ) used for the synthesis

Compound collections for mass screening are prepared by the methods of combinatorial chemistry. Thus, hundreds or even thousands of compounds can be synthesized in parallel in a single day (high-throughput synthesis). The requirement for rapid and exact characterization of the compounds (high-throughput analysis) arises from this development. The aim is to achieve highest degrees of automation resulting in most effective and rapid mass spectrometric analysis. The measurements illustrated in the following were all done with the aid of automated samplers and software specifically designed for this automated procedures. The combinatorial synthesis of the pyrrole amide compound collection used for this demonstration was carried out on solid phase [33].

Briefly, the synthetic strategy and the reaction conditions used are summarized in the following.

### 20.3.1.1 Synthesis of Pyrrole Amide Collections

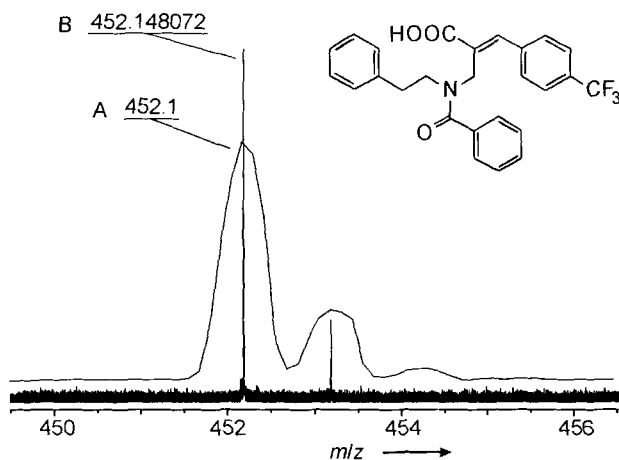
The library of 140 pyrrole-3-carboxamides was synthesized following the split-mix protocol with 14 primary amines and 10  $\alpha$ -bromo-ketones as variable building blocks. In the first step Rink-Amide-AM-PS resin was acetoacetylated at  $-15^\circ\text{C}$  by addition of diketene to the resin suspended in  $\text{CH}_2\text{Cl}_2$ . After 0.5 h at  $-15^\circ\text{C}$  and 2 h at room temperature the resin was washed and dried. A negative Kaiser test proved the complete conver-

sion of the amino functions to the polymer bound acetoacetic amide. The resin was then divided into 14 aliquots. Each of the aliquots was treated with one of the 14 primary amines shown in Fig. 20.8 resulting in 14 different immobilized enamines. This transformation was performed in DMF at room temperature for  $2 \times 24$  h in the presence of TMOF as dehydrating agent. The 14 resin aliquots were washed, dried, combined and divided into 10 aliquots. In the heterocyclization step each aliquot reacted with one of the  $\alpha$ -bromoketones (Fig. 20.9) for 17 h (3 h for bromopyruvic acid ethyl ester) in DMF at room temperature. 2,6-Di-*tert*-butyl-pyridine was used as base. In the last step the 10 sub-libraries were cleaved from the resin aliquotes with 20% TFA/ $\text{CH}_2\text{Cl}_2$ .

The results of the direct measurements and those of micro-HPLC/MS measurements of the sublibrary resulting from the combination of amines A1 to A14 ( $\text{R}^1$ ) and the  $\alpha$ -bromoketone. B6 ( $\text{R}^2$ ,  $\text{R}^3$ ) will be shown in the following section.

### 20.3.1.2 Automated Direct Measurement

In Figure 20.1 mass spectra from measurements obtained by ES quadrupole (A) and ES-FT-ICR mass spectrometry (B) of a compound prepared by solid phase organic synthesis are superimposed. The ES-FT-ICR measurement gives a mass accuracy which is greater by a factor of 550 (A:  $\delta = 221$ ; B:  $\delta = 0.4$ ). Accordingly, the difference between calculated and measured mass of the FT-ICR determination corresponds to 0.00018 Da in this example. The resolution is about a factor of 15 better than for comparable measurement with a quadrupole analyzer (A: 4 500; B: 70 000). The FT-ICR measurement was taken in the broad band mode at a scanned mass range of 1 400 Da. This measurement mode



**Fig. 20.10.** Superimposed measurement in negative mode of a compound obtained by solid phase synthesis with ESI-quadrupole (A) and ESI-FT-ICR mass spectrometry (B). Starting points for elementary analysis:  $[\text{M}-\text{H}]^-$  measured: 452.148072 Da, charge  $-1$ , min./max. double bond equivalents (DBE) 0.5/30, tolerance 0.0005 Da. Elementary analysis for  $M_r = 452.148072$ : C<sub>26</sub> O<sub>3</sub> N<sub>1</sub> S<sub>0</sub> H<sub>21</sub> F<sub>3</sub>;  $[\text{M}-\text{H}]^-$  calculated: 452.1478911, 15.5 DBE, deviation 0.4 ppm. The calculation of possible molecular formulae on the basis of the very precise ESI-FT-ICR measurement thus gave solely the correct composition.

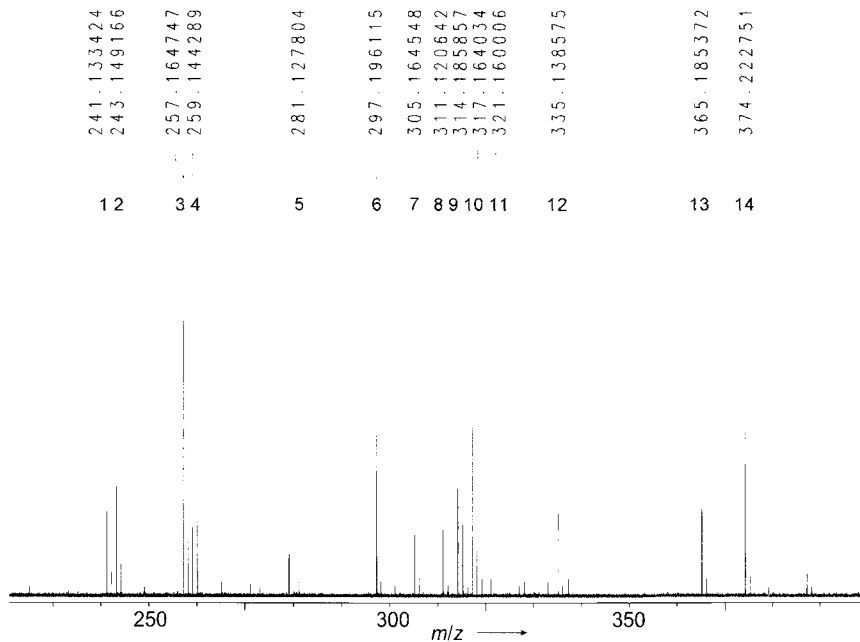


is not designed for the highest resolutions, but nevertheless a resolution of 70 000 was obtained. The corresponding measurement in high resolution mode at a scanned mass range of only 100 Da gave a resolution of 350 000 (data not shown).

The amount of sample required for an FT-ICR measurement is 10 times less than for the quadrupole measurement (A: 40 pmol; B: 4 pmol). The tolerance of the FT-ICR measurement resulting from the mean accuracy of the measurements is very small (0.0005 Da), for which reason the number of postulated molecular formulae is small. Owing to the high resolution, the isotope peaks are especially easy to recognize. When the exclusion rules are taken into consideration (nitrogen rule, double bond rule, valences), the calculation for this example gives solely the elementary composition which corresponds to the actual molecular formula. The FT-ICR measurement thus enables rapid elementary analysis of individual compounds in the pharmacologically interesting range of up to 500 Da.

This rapid molecular formula determination may also be applied to complex mixtures such as those which arise from combinatorial chemistry. To demonstrate the performance of this method in this respect, a split-mix compound collection of 140 different pyrrole amides consisting of 10 subcollections of 14 compounds each was synthesized on Rink-amide-AM-resin.

The results of the subcollection with 2-bromoacetophenone as building block ( $R^2 = C_6H_5$ ,  $R^3 = H$ ) are illustrated as an example. The total amount of sample of 20 pmol



**Fig. 20.11.** Automated routine ESI-FT-ICR measurement of the pyrrole amide library. [Extract from 220–400 m/z; acquisition time: 30 s]

During this acquisition the full range of 1400 m/z was scanned 16 times. Processing of the acquired data (smoothing parameters, number of data points, peaklabeling) was performed automatically.

**Table 20.1.** Result of the ESI-FT-ICR-MS from figure 20.11. The high mass accuracy (average 1.05 ppm) permits only one, two or three possible elementary constitutions for one compound.

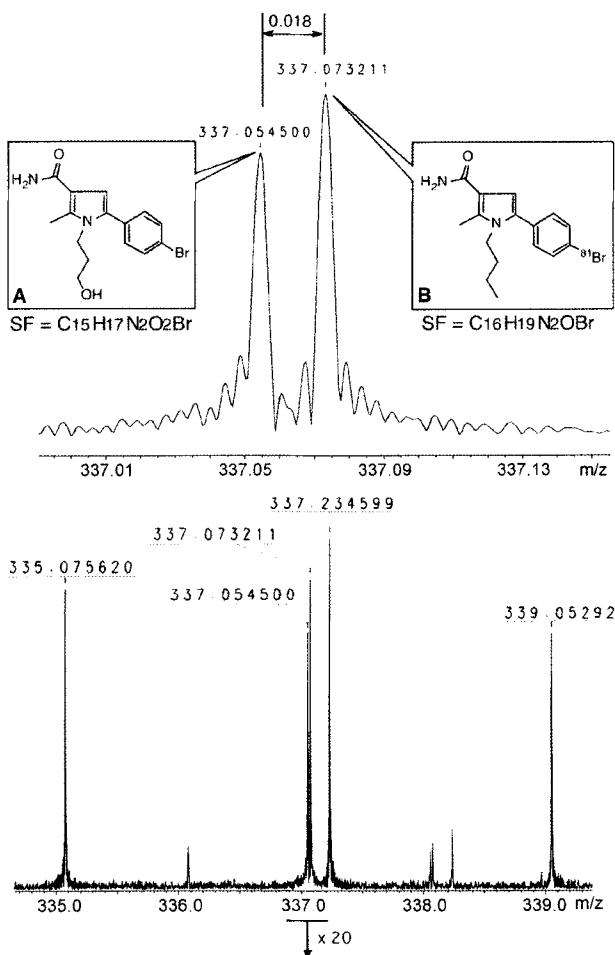
Compd	Mol. formula	Mass calcd [M + H] <sup>+</sup>	Mass found [M + H] <sup>+</sup>	$\Delta$ mass	$\delta$	Hits
1	C <sub>15</sub> H <sub>16</sub> N <sub>2</sub> O	241.133530	241.133424	0.000106	0.44	1
2	C <sub>15</sub> H <sub>18</sub> N <sub>2</sub> O	243.149179	243.149166	0.000013	0.05	1
3	C <sub>16</sub> H <sub>20</sub> N <sub>2</sub> O	257.164829	257.164747	0.000082	0.31	1
4	C <sub>15</sub> H <sub>18</sub> N <sub>2</sub> O <sub>2</sub>	259.144093	259.144289	0.000196	0.75	2
5	C <sub>17</sub> H <sub>16</sub> N <sub>2</sub> O <sub>2</sub>	281.128444	281.127804	0.000640	2.27	2
6	C <sub>19</sub> H <sub>24</sub> N <sub>2</sub> O	297.196127	297.196115	0.000012	0.04	1
7	C <sub>20</sub> H <sub>20</sub> N <sub>2</sub> O	305.164829	305.164548	0.000281	0.92	2
8	C <sub>18</sub> H <sub>18</sub> N <sub>2</sub> OS	311.121294	311.120642	0.000652	2.09	3
9	C <sub>18</sub> H <sub>23</sub> N <sub>3</sub> O <sub>2</sub>	314.186290	314.185857	0.000433	1.37	1
10	C <sub>21</sub> H <sub>20</sub> N <sub>2</sub> O	317.164829	317.164034	0.000795	2.50	2
11	C <sub>20</sub> H <sub>20</sub> N <sub>2</sub> O <sub>2</sub>	321.159743	321.160006	0.000263	0.81	2
12	C <sub>20</sub> H <sub>18</sub> N <sub>2</sub> O <sub>3</sub>	335.139008	335.138575	0.000433	1.29	2
13	C <sub>22</sub> H <sub>24</sub> N <sub>2</sub> O <sub>3</sub>	365.185955	365.185372	0.000583	1.59	2
14	C <sub>24</sub> H <sub>27</sub> N <sub>3</sub> O	374.222674	374.222751	0.000077	0.20	2
			mean value:	0.000326	1.05	1.57
			standard deviation:	0.000264	0.83	

gives a mean sample consumption of 1.4 pmol per compound for the routine measurement. All 14 expected products could be detected by mass spectrometry.

The resolution averaged over all 14 compounds ( $m/\Delta m_{50\%}$ ) is 38000. The mean mass accuracy  $\Delta m$  is 0.00032 Da (1.05 ppm) at a standard deviation of 0.00026 Da. The possible molecular formulae were calculated for all 14 masses of the measurement. The exact masses for the elements C, H, N, O, and S were considered by the program, and results with a nitrogen content of over 50 weight percent were rejected. The very high mass accuracy allows only one or two, at the most three, possible elementary compositions (hits) for each compound. The individual compounds are thus so precisely characterized that, with prior knowledge of the reaction conditions, the structural formulae may also be determined.

The very high accuracy and resolution of the measurements alone represent a considerable increase in the efficiency of mass spectrometric methods. The precise separation by mass of the direct measurement of a compound mixture frequently makes prior separation by HPLC unnecessary. As the result of direct determination, time-consuming and expensive tagging systems for compound mixtures of up to 100 components become unnecessary. The time saved even for low sample availability is considerable.

To meet the demands of high-throughput analysis, we automated the sample input. For this purpose an HPLC system was coupled directly to the FT-ICR mass spectrometer. The integrated automatic sample injector inputs a new sample for measurement into the FT-ICR mass spectrometer by flow injection every two minutes. Together with the synchronized start of the FT-ICR measurement, this allows a sample throughput of 300 samples per night. At the end of the measurement the raw data collected are processed automatically according to the wishes of the user (number of data points, choice of smoothing parameters, peak labeling, etc.). After the automatic sample injector is reset the mass spectrometer is ready for use.



**Fig. 20.12.** Completely resolved signal of compound **A** and **B** with a mass difference of solely 0.0018 Da (pyrrole amide compound library with  $R^1$  = amines A1 to A14 and  $R^2R^3$  = 2-bromo-4-phenylacetophenone)

Processing of the obtained data allows a comparison with the expected results. A summary of the results gives information about relative mass peak intensities, found masses expected masses and the relative error of measurement. The program also allows the search in the spectra for typical adduct peaks such as sodium or ammonium ions. The evaluation program can be edited by the user in order to allow user specific search criteria.

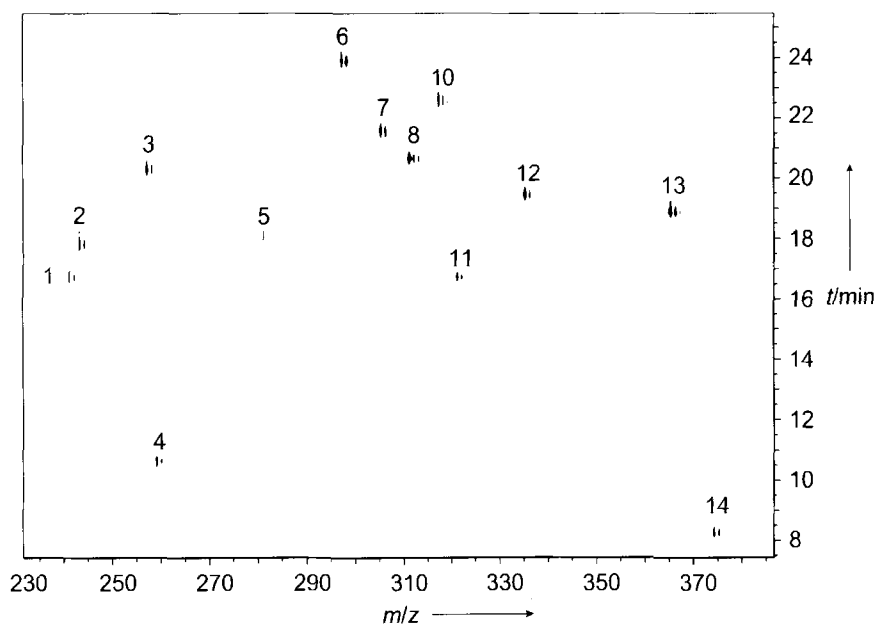
The 60° arrangement of the spray capillary in the ion source (Analytica of Bradford Inc.) avoids contamination of the transfer capillary thus allowing permanent measurements for days with the turbo pumped vacuum system. Therefore, the mass spectrometer can be immediately used after installation of the automatic sample injector.

The ESI-FT-ICR-MS and automatic sampler were coupled through a fused silica capillary (75  $\mu\text{m}$  internal diameter, 365  $\mu\text{m}$  external diameter). The samples were injected into the continuous solvent stream (flow rate 30  $\mu\text{L}/\text{min}$ , 60% acetonitrile, 40% water, each with 0.1% formic acid) by the automatic sample injector of the HPLC system (Hewlett Packard, HP-1100 series). Measurements with a known compound collection provided the data for three-point calibration of the mass spectrometer.

### 20.3.1.3 MicroLC/ES-FT-ICR-MS Coupling

The direct high resolution FT-ICR-MS measurements (i.e. without previous separation) give sufficient information to solve most problems in combinatorial chemistry. However, separations by HPLC, CE, etc. give additional information on isobaric compounds and also a correlation between the yield as indicated from the UV trace and the corresponding mass. New rapid algorithms for the Fourier transformation and fast computers for acquisition allow short scan times (250 ms) for a large mass range (4000 amu). Therefore the good time resolution allows the detection of compounds eluting in very short intervals.

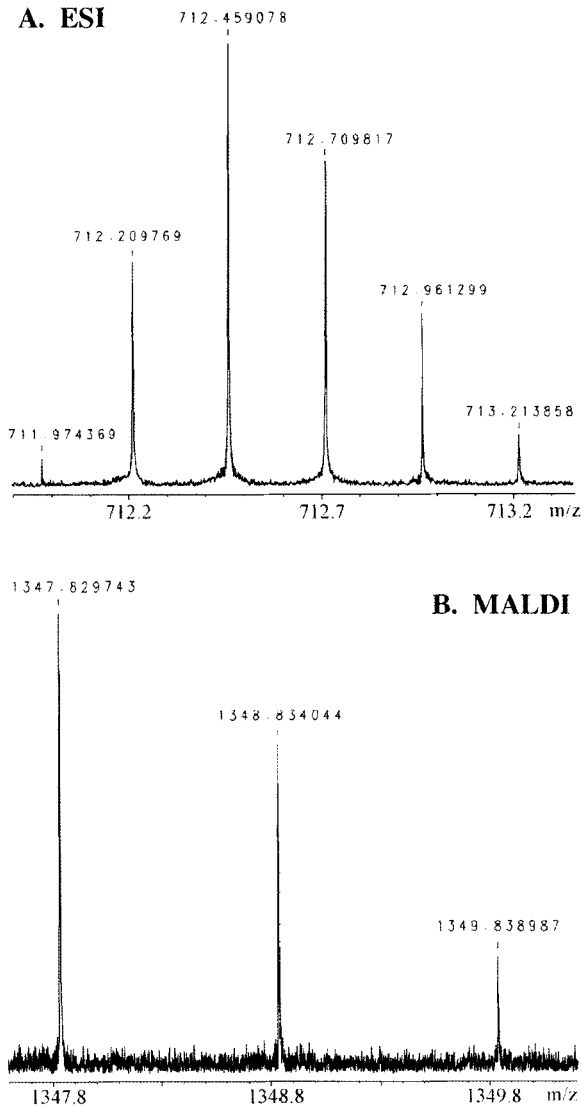
The results of such a microHPLC/ES-FT-ICR mass spectrometric analysis of the pyrrole amide subcollection is shown in Figure 20.13 as contour plot. The separation of the subcollection was carried out on an RP C18 column (GROM-Sil ODS7; 1  $\times$  6 mm). For



**Fig. 20.13.** Contour plot from the HPLC/FT-ICR-MS coupling of a pyrrole amide compound collection ( $R^2 = \text{C}_6\text{H}_5$ ,  $R^3 = \text{H}$ ). This method allows to prove the identity of isomers. The isotope resolution of the molecule ion peaks is recognizable.

better separation a gradient of 10-80% acetonitrile/water (25 min) and a flow rate of 50  $\mu\text{l}/\text{min}$  were used. The eluent and the column were kept at 45°C. The solvents contained 0.1% formic acid.

In the case of gradient elution the low flow rates required a considerable reduction of the retardation volume between static gradient mixer and column head and photodiode



**Fig. 20.14. above: A.** ESI-ultrahigh-resolution-measurement of Mellitin ( $[\text{M} + 4\text{H}]^+$ -Peak,  $[\text{M} + \text{H}]^+$  mono-isotope: 2843.87 Da); S/N = 65; resolution: 320 000.

**below: B.** MALDI-ultrahigh-resolution-measurement of compound P (polypeptide); single shot; sample consumption is in the range of some femto moles; S/N = 45; resolution 380 000.

array detector. This was achieved by removing pulsation absorbers, reducing the lengths and diameter of gradient capillaries and placement of the static gradient mixer directly in front of the injection valve. Furthermore a high pressure micro flow cell with small volume (1.7  $\mu\text{l}$ ) with long path length (6 mm) was used. The advantages for the use of such micro-HPLC systems are the low flow rates allowing the splitless coupling to the ES source and a reduced eluent consumption.

The measurement described utilizes fully all technical possibilities for improving the performance of mass spectrometric methods for use in high-throughput analysis. These are currently the best possible resolution and mass accuracy over a broad mass range, short measurement times, lowest sample consumptions and automated sampling and processing. This hyphenated method is thus an important aid in synthesis optimization and control of individual compounds and compound collections in combinatorial chemistry.

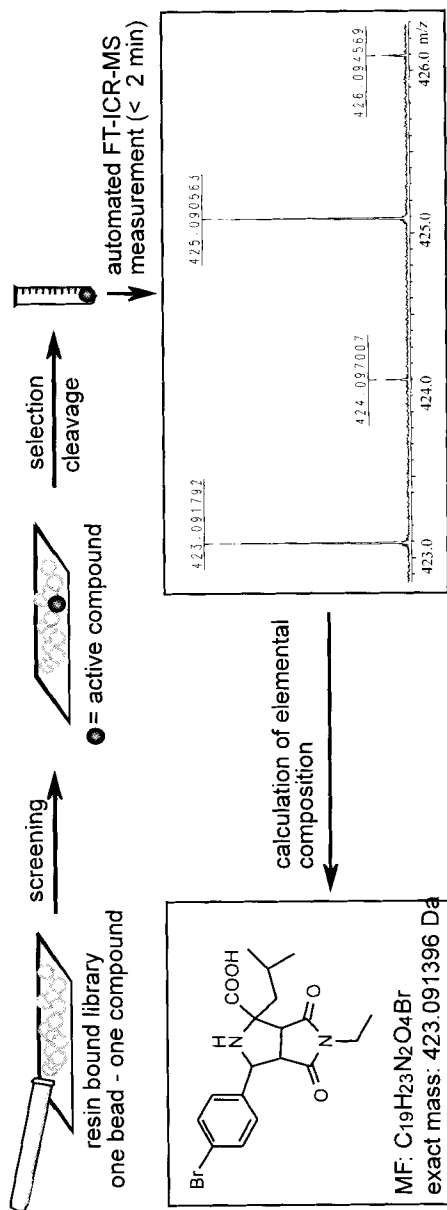
The system is also ideally suited for use in protein analysis, for the elucidation of natural products by  $\text{MS}^n$  experiments, or for metabolic studies.  $\text{MS}^n$  experiments in natural product elucidation [35], metabolic studies of drugs or the use in protein [36] and peptide [37] analysis clearly demonstrate the power of ICR-MS. The high flexibility with respect to the choice of ion sources is of importance. Thus, the advantages of both MALDI and ES techniques can be taken or the disadvantages avoided: MALDI yields mainly singly charged ions by spatially resolved ionisation, but matrix background can be a problem; ES yields mainly multiply charged ions in continuous ionisation and makes couplings with separation systems possible. Figure 20.14 illustrates two different ICR-MS measurements using ES ionization of the polypeptide melittin and MALDI ionization of the smaller polypeptide substance P.

### 20.3.2 Single Bead Analysis

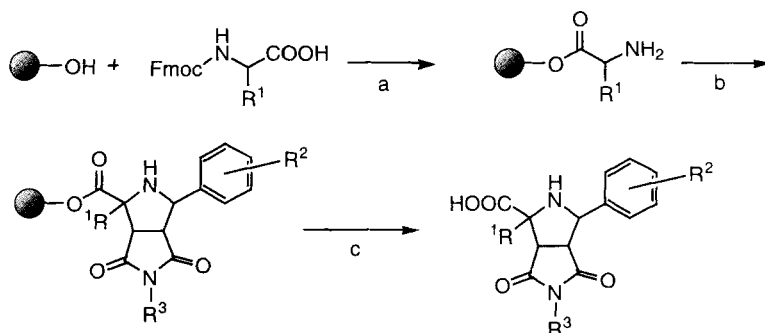
Biological activity of resin-bound compounds can be determined also directly on single beads for example by fluorescence labeling of the protein target. A more recent example uses synthetic resin-bound compound collections with incorporated fluorescence quencher systems in combination with proteases [38, 39]. The “active” single resin bead can be picked out with a micromanipulator and the product can be analyzed by FT-ICR-MS after cleavage from the resin. The high sensitivity and the accuracy of determination of the elemental composition allows an immediate structure determination without any time consuming deconvolution or tagging procedures. The procedure for the FT-ICR-mass spectrometric single bead analysis will be shown in section 20.3.2.2 using the pyrrolidine library explained in section 20.3.2.1.

#### 20.3.2.1 Synthesis of Pyrrolidine Libraries

The synthesis of highly substituted pyrrolidines was performed using commercially available monodisperse polystyrene Wang resin from Rapp Polymere (Tübingen, Germany). The diameter of the beads was 160  $\mu\text{m}$ .



**Fig. 20.15.** Strategy for a solid phase on-bead screening followed by high resolution single bead ESI-FT-ICR-mass-spectrometry.



**Fig. 20.16.** Synthesis of highly substituted pyrrolidines on monodisperse Rapp Wang-PS resin (diameter 160  $\mu\text{m}$ ). Coupling of an Fmoc-protected amino acid onto the Wang resin (TBTU/HOBT, DMF, rt). **a.** cleavage of the protecting group (DMF/piperidine); **b.** formation of imines, addition of substituted benzaldehydes (THF/TMOF, 1/1, rt), cycloaddition to bicyclic pyrrolidines (toluene, 110°C); **c.** cleavage from the solid support (DCM/TFA 1/1)

For the synthesis of single compounds the resin was coupled with Fmoc-protected amino acids using TBTU/HOBT as coupling reagents. After cleavage of the protecting group the free amino terminus was reacted with substituted benzaldehydes to the corresponding imines (16 h). The cycloaddition with substituted maleimides was performed using toluene at 110°C for 24 h. After the reaction sequence the resin was washed and dried in vacuum.

### 20.3.2.2 Result of Single Bead Analysis

The separation of the resin beads was performed using a micromanipulation apparatus from Zeiss Optics. The resin beads were spread on a surface and single beads were picked up with a glass capillary and transferred into HPLC vials. The pyrrolidines were cleaved from the single beads with 10  $\mu\text{l}$  of a mixture of DCM/TFA (1:1). After the evaporation of the cleavage solution, the products were diluted in 40  $\mu\text{l}$  ACN/water and placed in the autosampler for automated FT-ICR-MS analysis.

**Table 20.2.** Results from the measurement of 3 compounds:

Compd	Mass calcd $[M + H]^+$	Mass found $[M + H]^+$	abs. Intensity	$\Delta$ mass	$\delta$
LeuBrB	423,091396	423,091792	$0,2 \times 10^5$	0,000396	0,93
LeuBrC	549,001909	549,002323	$4,0 \times 10^6$	0,000414	0,75
LeuMeC	499,122696	499,122437	$1,0 \times 10^6$	0,000259	0,52
			mean value:	0,000356	0,74
			standard deviation:	0,000085	0,21

LeuMeB ( $R^1 = \text{CH}_2\text{CH}(\text{CH}_3)_2$ ,  $R^2 = 2,4\text{-dimethylphenyl}$ ,  $R^3 = \text{CH}_2\text{CH}_3$ ), LeuBrB ( $R^1 = \text{CH}_2\text{CH}(\text{CH}_3)_2$ ,  $R^2 = 4\text{-bromophenyl}$ ,  $R^3 = \text{CH}_2\text{CH}_3$ ), Leu BrC ( $R^1 = \text{CH}_2\text{CH}(\text{CH}_3)_2$ ,  $R^2 = 4\text{-bromophenyl}$ ,  $R^3 = 4\text{-bromophenyl}$ )



The sample consumption for the mass spectrometric identification of the compound on the selected single bead is very low (about 15 pmol). This allows one to use the remaining part of the solution for the biological tests.

The main advantages of the described screening procedure are:

- a) The application of the split/mix synthesis of large compound collections is possible without deconvolution after hit finding. All tagging procedures such as oligonucleotides, peptides, IR markers [40], polyhalogenated hydrocarbons and GC-MS [41] or radiofrequency tags can be avoided [42]. Considering the permanently high personal costs for deconvolution or tagging procedure the investment of an ICR-MS is justified.
- b) False positive hit findings or negative results which are very common for on-bead screening with split/mix libraries are immediately recognized.

In summary the advantages of single bead analysis are combined with those of library synthesis without creating new disadvantages.

## 20.4 Materials and Methods

The ES/MALDI-FT-ICR mass spectrometer of the Institute of Organic Chemistry at the University of Tübingen is from Bruker Daltonik GmbH, Bremen. The system is evacuated by efficient turbo molecular pumps which allows HPLC and GC coupling over long time periods. ES, NanoSpray and MALDI ionization sources allow individual adaptation to particular problems. Samples from combinatorial chemistry can be routinely analyzed with the 60°-ESI from Analytica of Branford Inc. (Branford, MA) via LC/ES-FT-ICR-MS as well as small amounts of valuable biological samples with nanoESI.

Using the ESI-/NanoSpray sources the generated ions can be intermediately trapped in a Hexapole (IRIS™, Hexapole Ion Guide) and pulsed in the analyzer cell. This allows the concentration of highly diluted samples in the gas phase. Depending on the ionization method the sample consumption is between several picomoles (pmol = 10<sup>-12</sup> mol; ES) and few femtomoles (fmol = 10<sup>-15</sup> mol, nanoESI, MALDI). The possibility of measuring the generated and trapped ions in the analyzer cell several times (remeasurements) by quadrupolar excitation and axialization allows mass spectrometric measurements from a few hundred molecules. This method is applied for extremely low amounts of sample where enhancement of the signal intensity and improvement of the S/N ratio is obtained and for ion-molecule reactions. The analyzer cell reflects the concept of the patented "infinity cell" [43].

The automated registration of daughter ion spectra (MS<sup>2</sup>) and tandem mass spectra (up to MS<sup>4</sup>) can be carried out with very high mass accuracy (< 3 ppm). Monoisotopic selection of the desired ions is possible. The resolutions ( $R = m/\delta m$ ) are in the range of 100 000 using the broad band mode (scanning a mass range of 2000 amu) and 380 000 in the high resolution mode (mass range of about 50 amu).

The mass accuracy of single compound measurements is below 0.5 ppm, which means that the measured mass of a molecule of 1 000 000 Da would differ by ±0.5 amu or an ion mass of 500 Da would have an uncertainty of ±0.00025 amu compared to the calculated exact mass.

For the calculation of the elemental compositions (Software: Bruker XMASS 5.0) the possible numbers for the elements were not restricted. The tolerance of the mass accuracy was limited to 3 ppm for libraries and 1 ppm for single compounds. As possible elements C, H, O, N, S, F were given and also Br, Cl.

All measurements were carried out on the Bruker Daltronik Apex<sup>TM</sup>II 4.7 ESI/MALDI/Fourier-Transform-Ionocyclotron Mass Spectrometer. The number of data points per measurement was 256 K in the mass range 100–1500 Da. The range was scanned one time per second and 16 scans were accumulated. The time shifted start of the automated measurements was possible by transfer of a TTL pulse of the HPLC system to the acquisition computer of the mass spectrometer at the time point of injection of the autosampler.

## 20.5 References

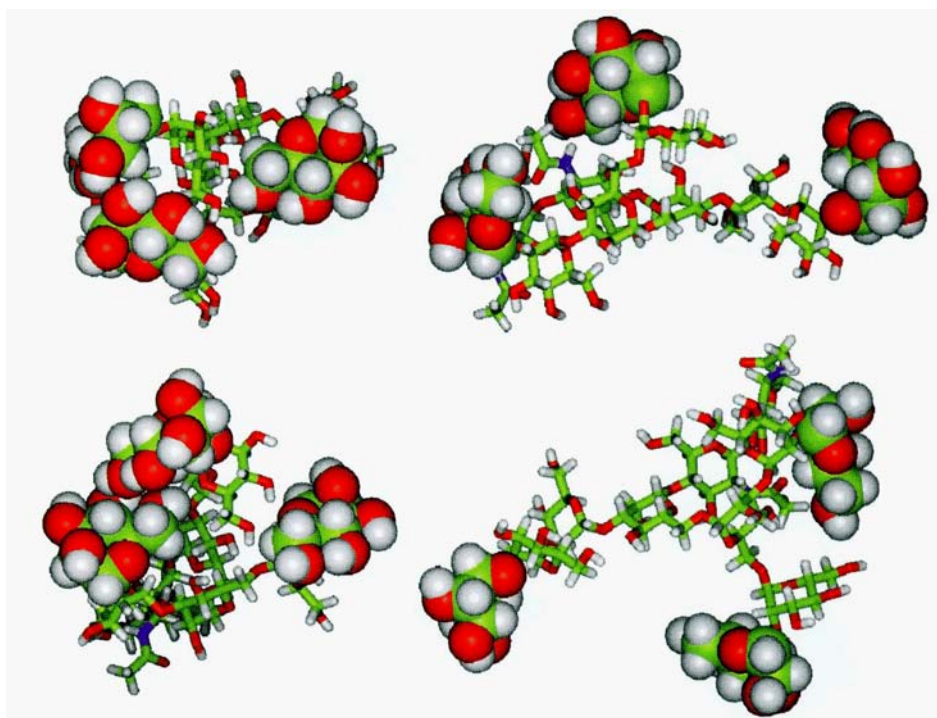
- [1] Balkenhohl, F., von dem Bussche Hünnefeld, C., Lansky, A. und Zechel, C., Combinatorial synthesis of small organic molecules. *Angew. Chem. Int. Ed.* **1996**, *35*, 2289–2337.
- [2] Früchtel, J. S. und Jung, G., Organic chemistry on solid supports. *Angew. Chem. Int. Ed.* **1996**, *35*, 17–42.
- [3] Gordon, E. M., Gallop, M. A. und Patel, D. V., Strategy and tactics in combinatorial organic synthesis. Applications to drug discovery. *Account. Chem. Res.* **1996**, *29*, 144–154.
- [4] Hermkens, P. H. H., Ottenheijm, H. C. J. und Rees, D. C., Solid-phase organic reactions II: A review of the literature Nov 95 Nov 96. *Tetrahedron* **1997**, *53*, 5643–5678.
- [5] Cody Reynolds, D. M., Hobbs DeWitt, S. H., Hodges, J. C., Kiely, J. S., Moos, W. M., Pavia, M. R., Roth, B. D., Schroeder, M. C. und Stankovic, C. J., *WO 94/08711* 1994.
- [6] Wigger, M., Nawrocki, J. P., Watson, C. H., Eyler, J. R. und Benner, S. A., Assessing enzyme substrate specificity using combinatorial libraries and electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1749–1752.
- [7] Smart, S. S., Mason, T. J., Bennell, P. S., Maeij, N. J. und Geysen, H. M., High-throughput purity estimation and characterisation of synthetic peptides by electrospray mass spectrometry. *Int. J. Pep. Protein Res.* **1996**, *47*, 47–55.
- [8] Haap, W. J., Metzger, J. W., Kempfer, C. und Jung, G., Composition and purity of combinatorial aryl ether collections analyzed by electrospray mass spectrometry. *Mol. Diversity* **1997**, *3*, 29–41.
- [9] Kienle, S., Wiesmüller, K. H., Brunjes, J., Metzger, J. W. und Jung, G., MS-Pep: a computer program for the interpretation of mass spectra of peptide libraries. *Fresenius J. Anal. Chem.* **1997**, *359*, 10–14.
- [10] Weller, H. N., Young, M. G., Michalczyk, S. J., Reitnauer, G. H., Cooley, R. S., Rahn, P. C., Loyd, D. J., Fiore, D. und Fischman, S. J., High throughput analysis and purification in support of automated parallel synthesis. *Mol. Diversity* **1997**, *3*, 61–70.
- [11] Shapiro, M. J. und Wareing, J. R., NMR methods in combinatorial chemistry. *Curr. Opin. Chem. Biol.* **1998**, *2*, 372–375.
- [12] Yan, B., Gremlich, H.-U., Moss, S., Coppola, G. M., Sun, Q. und Liu, L., A Comparison of Various FTIR and FT Raman Methods: Applications in the Reaction Optimization Stage of Combinatorial Chemistry. *J. Comb. Chem.* **1999**, *1*, 46–54.
- [13] Haap, W. J., Walk, T. B. und Jung, G., FT-IR Mapping - A New Tool for Spatially Resolved Characterization of Polymer-Bound Combinatorial Compound Libraries with IR Microscopy. *Angew. Chem. Int. Ed.* **1998**, *37*, 3311–3314.
- [14] Marshall, A. G., Hendrickson, C. L. und Jackson, G. S., Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.* **1998**, *17*, 1–35.
- [15] Walk, T. B., Trautwein, A., Richter, H. und Jung, G., ESI-Fourier-Transform-Ionenzyklotron-Massenspektrometrie (ESI-FT-ICR-MS) - schnelle höchstauflösende Analytik für kombinatorische Verbindungsbibliotheken. *Angew. Chem.* **1999**, *111*, 1877–1880; *Angew. Chem. Int. Ed.* **1999**, *38*, 1763–1765.
- [16] Sommer, H. und Thomas, H. A., Detection of Magnetic Resonance by Ion Resonance Absorption. *Phys. Rev.* **1950**, *78*, 806.

- [17] Sommer, H., Thomas, H. A. und Hipple, J. A., The Measurement of  $e/M$  by Cyclotron Resonance. *Phys. Rev.* **1951**, *82*, 697–702.
- [18] McIver, R. T., Jr., A Trapped Ion Analyzer Cell for Ion Cyclotron Resonance Spectroscopy. *Rev. Sci. Instrum.* **1970**, *41*, 555.
- [19] Comisarow, M. B. und Marshall, A. G., Frequency-Sweep Fourier Transform Ion Cyclotron Resonance Spectroscopy. *Chem. Phys. Lett.* **1974**, *26*, 489–490.
- [20] Comisarow, M. B. und Marshall, A. G., Fourier Transform Ion Cyclotron Resonance Spectroscopy. *Chem. Phys. Lett.* **1974**, *25*, 282–283.
- [21] Allemann, M., Kellerhals, H. P. und Wanczek, K. P., A New Fourier Transform Mass Spectrometer with a Superconducting Magnet. *Chem. Phys. Lett.* **1980**, *75*, 328–331.
- [22] Hillenkamp, F., Unsold, D., Kaufmann, R. und Nitsch, R., A High Sensitivity Laser Microprobe Mass Analyzer. *Appl. Phys.* **1975**, *8*, 341–348.
- [23] Severs, J. C., Hofstadler, S. A., Zhao, Z., Senh, R. T. und Smith, R. D., The interface of capillary electrophoresis with high performance Fourier transform ion cyclotron resonance mass spectrometry for biomolecule characterization. *Electrophoresis* **1996**, *17*, 1808–1817.
- [24] Hofstadler, S. A., Severs, J. C., Smith, R. D., Swanek, F. D. und Ewing, A. G., High performance Fourier transform ion cyclotron resonance mass spectrometric detection for capillary electrophoresis. *J. High. Res. Chromatogr.* **1996**, *19*, 617–621.
- [25] Carlin, T. J. und Freiser, B. S., Pulsed Valve Addition of Collision and Reagent Gases in Fourier Transform Mass Spectrometry. *Anal. Chem.* **1983**, *55*, 571.
- [26] Büchner, M., Nixdorf, A. und Grützmaker, H. F., Reactions of gaseous, halogenated propene radical cations with ammonia: A study of the mechanism by Fourier transform ion cyclotron resonance. *Chem. Eur. J.* **1998**, *4*, 1799–1809.
- [27] Anderson, J. S., Vartanian, H. und Laude, D. A., Evolution of trapped ion cells in Fourier transform ion cyclotron resonance mass spectrometry. *Trend. Anal. Chem.* **1994**, *13*, 234–239.
- [28] Amster, I. J., Fourier Transform Mass Spectrometry. *J. Mass Spectrom.* **1996**, *31*, 1325–1337.
- [29] Cody, R. B., Burnier, R. C. und Freiser, B. S., Collision Induced Dissociation with Fourier Transform Mass Spectrometry. *Anal. Chem.* **1982**, *54*, 96–101.
- [30] Bowers, W. D., Delbert, S.-S., Hunter, R. L. und McIver, R. T., Jr., Fragmentation of Oligopeptide Ions Using Ultraviolet Laser Radiation and Fourier Transform Mass Spectrometry. *J. Am. Chem. Soc.* **1984**, *106*, 7288–7289.
- [31] Little, D. P., Speir, J. P., Senko, M. W., O'Connor, P. B. und McLafferty, F. W., Infrared multiphoton dissociation of large multiply charged ions for biomolecule sequencing. *Anal. Chem.* **1994**, *66*, 2809–2815.
- [32] Wu, Q. Y., Multistage accurate mass spectrometry: A “basket in a basket” approach for structure elucidation and its application to a compound from combinatorial synthesis. *Anal. Chem.* **1998**, *70*, 865–872.
- [33] Trautwein, A. W., Süßmuth, R. D. und Jung, G., Hantzsch pyrrole synthesis on solid support. *Biorg. Med. Chem. Lett.* **1998**, *8*, 2381–2384.
- [34] Rodgers, R. P., White, F. M., Hendrickson, C. L., Marshall, A. G. und Andersen, K. V., Resolution, elemental composition, and simultaneous monitoring by Fourier transform ion cyclotron resonance mass spectrometry of organosulfur species before and after diesel fuel processing. *Anal. Chem.* **1998**, *70*, 4743–4750.
- [35] Heck, A. J. R. und Derrick, P. J., Selective fragmentation of single isotopic ions of proteins up to 17 kDa using 9.4 Tesla Fourier transform ion cyclotron resonance. *Eur. Mass Spectrom.* **1998**, *4*, 181–188.
- [36] Green, M. K., Vestling, M. M., Johnston, M. V. und Larsen, B. S., Distinguishing small molecular mass differences of proteins by mass spectrometry. *Anal. Biochem.* **1998**, *260*, 204–211.
- [37] Demirev, P. A. und Zubarev, R. A., Probing combinatorial library diversity by mass spectrometry. *Anal. Chem.* **1997**, *69*, 2893–2900.
- [38] Meldal, M., Svendsen, I., Breddam, K. und Auzanneau, F. I., Portion-mixing peptide libraries of quenched fluorogenic substrates for complete subsite mapping of endoprotease specificity. *Proc. Nat. Acad. Sci. USA* **1994**, *91*, 3314–3318.
- [39] Meldal, M. und Svendsen, I., Direct visualization of enzyme inhibitors using a portion mixing inhibitor library containing a quenched fluorogenic peptide substrate. I: Inhibitors for subtilisin Carlsberg. *J. Chem. Soc., Perkin Trans. 1*: **1995**, 1591–1596.
- [40] Haap, W. J., Walk, T. B. und Jung, G., FT-IR Mapping - A New Tool for Spatially Resolved Char-

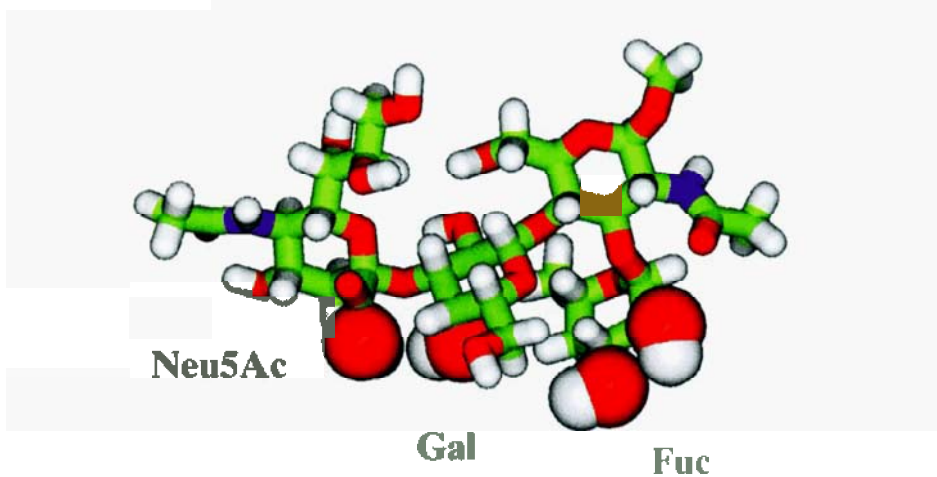
- acterization of Polymer-Bound Combinatorial Compound Libraries with IR Microscopy. *Angew. Chem. Int. Ed.* **1998**, *37*, 3311–3314.
- [41] Parandoosh, Z., Knowles, S. K., Xiao, X. Y., Zhao, C., David, G. S. und Nova, M. P., Synthesis coupled to scintillation proximity affinity screening. *J. Biomol. Screen.* **1998**, *3*, 293–298.
- [42] Xiao, X. Y., Parandoosh, Z. und Nova, M. P., Design and synthesis of a taxoid library using radiofrequency encoded combinatorial chemistry. *J. Org. Chem.* **1997**, *62*, 6029–6033.
- [43] Caravatti, P. und Allemann, M., The ‘Infinity Cell’: a New Trapped Ion Cell with Radiofrequency Covered Trapping Electrodes for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Org. Mass Spectrom.* **1991**, *26*, 514–518.

This Page Intentionally Left Blank

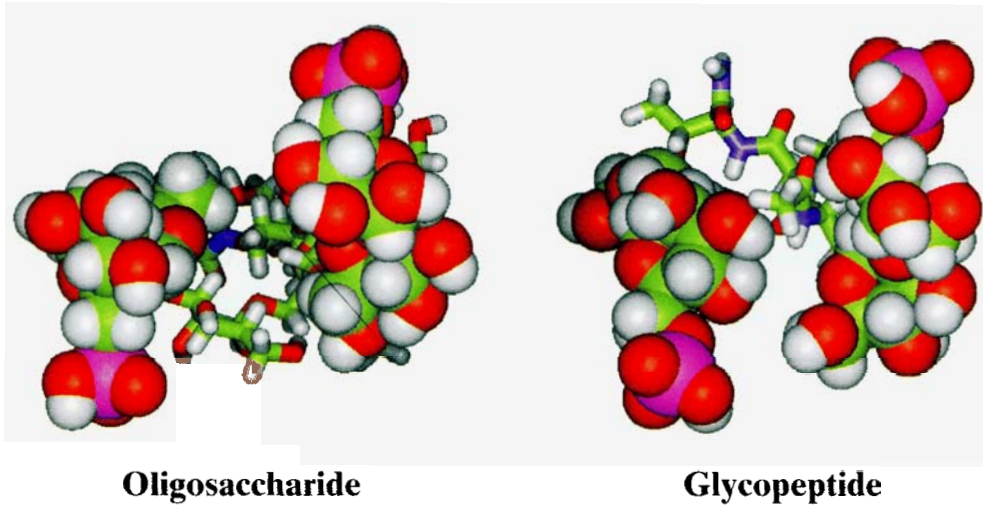
## Colour Plates



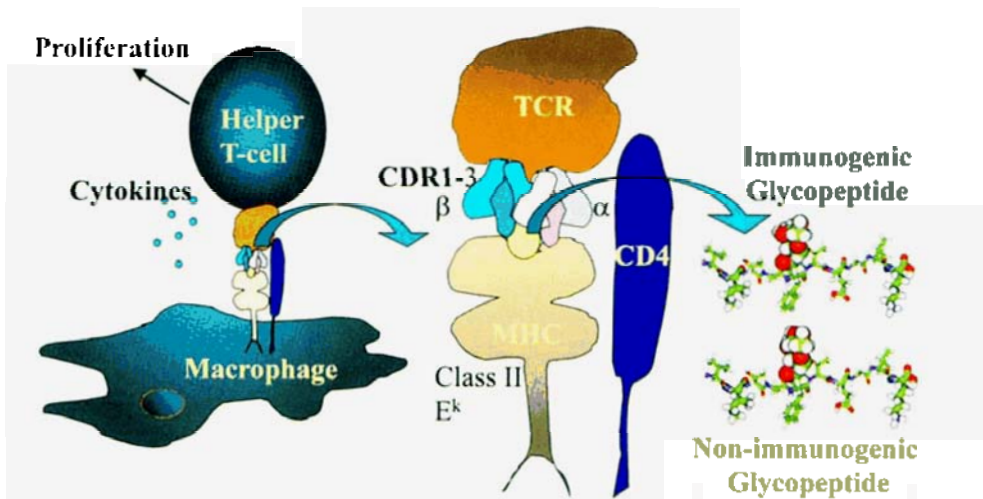
**Figure 8.3.** Structure models of four different conformer populations of a tri-antennary high-mannose undecasaccharide (M9) as inferred from NOE connectivities of fragments. Terminal mannoses are highlighted as CPK-models.



**Figure 8.4.** Bound conformation of the sialyl Lewis x tetrasaccharide with tightly bound hydroxyl groups shown as spheres.

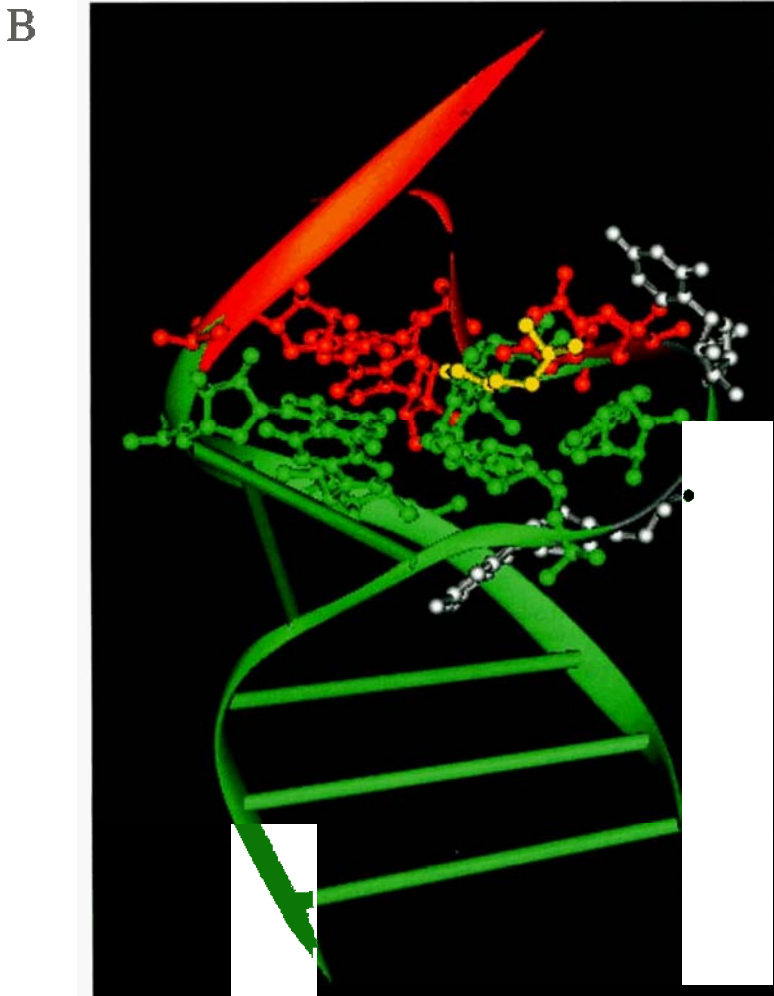
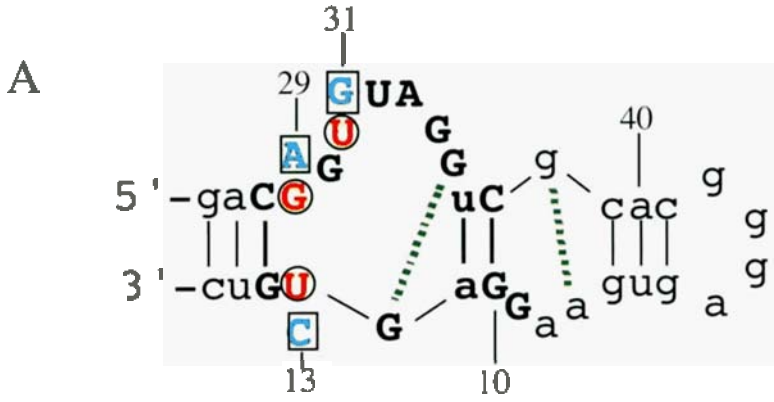


**Figure 8.8.** MD-models of phosphorylated high-mannose  $M_7$  and a glycopeptide mimic, as seen from the point of view of the mannose 6-phosphate receptor. Disaccharide phosphates are emphasized as spheres. The structure of the mimic was obtained after 40 ps at 400 K and 10 ps at 300 K using the Insight/Discover program package.

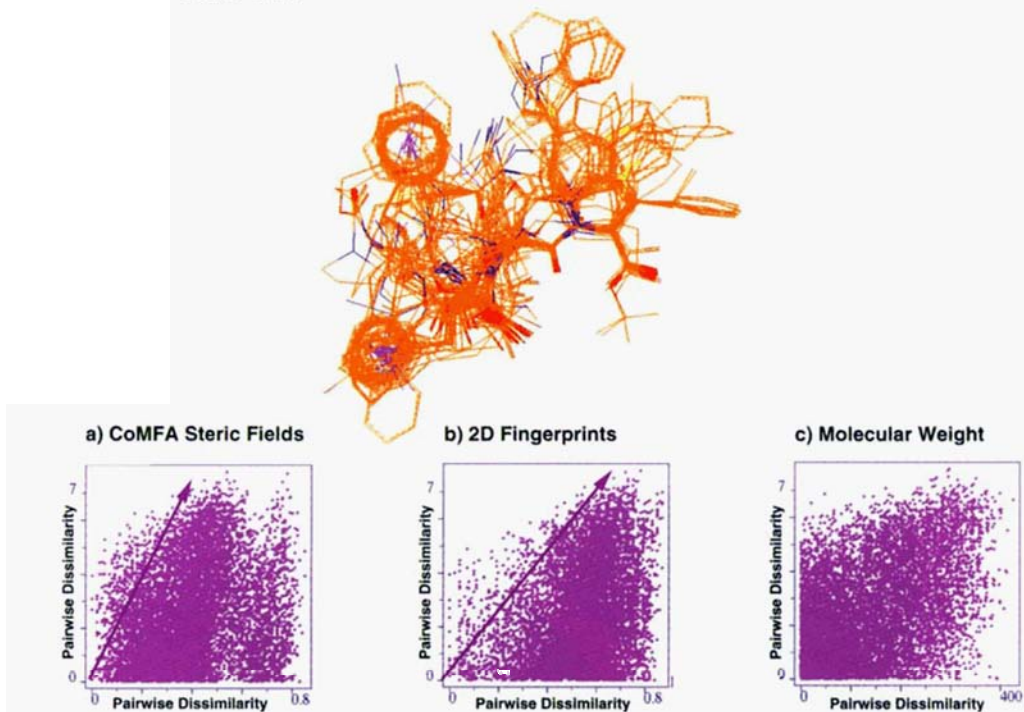


**Figure 8.9** (a) Mammalian histocompatibility complex (MHC) binding array of glycopeptides used in the study of glycan specificity of the T-cell receptor (TCR) recognition. Only the T<sub>n</sub> and T antigens were strong immunogens. (b) Glycopeptide/MHC/TCR interaction. The minor modification of substituting  $\alpha$ -GalNAc with  $\alpha$ -GlcNAc abolished immunogenicity.

**Figure 9.2.** (A) Secondary structure proposed previously for the citrulline- and arginine-specific aptamers, based on co-variations of selected sequences, on the chemical footprinting pattern obtained in the presence of the cognate amino acid, as well as in damage selection experiments. The bases which were conserved among different isolates are shown in upper case; variant bases are in lower case. The three nucleotides critical for arginine specificity (13, 29 and 31) are indicated by circles (for citrulline) and boxes (for arginine). (B) Tertiary structure of the L-arginine aptamer complex resolved by NMR spectroscopy. Yellow: L-arginine; red: the three mutations. (Illustration adapted from [9].)

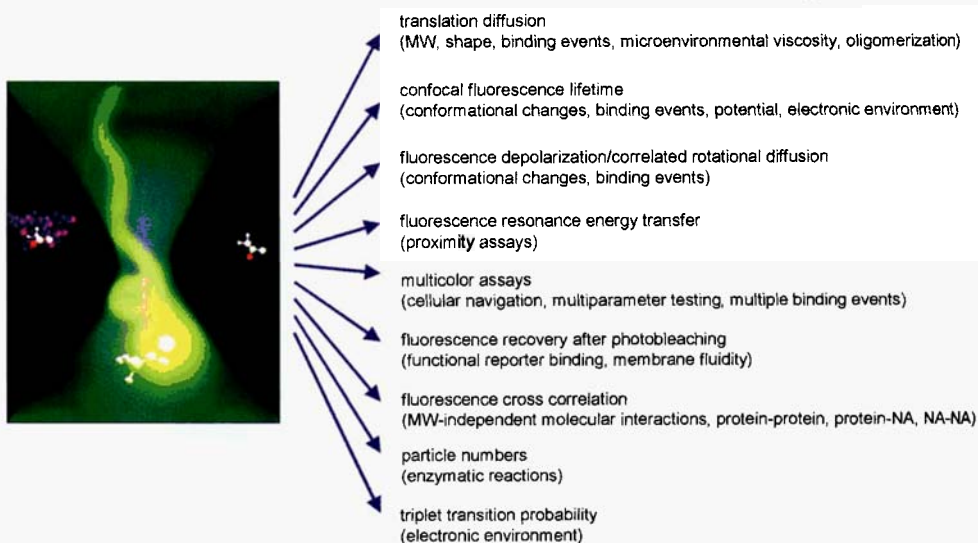






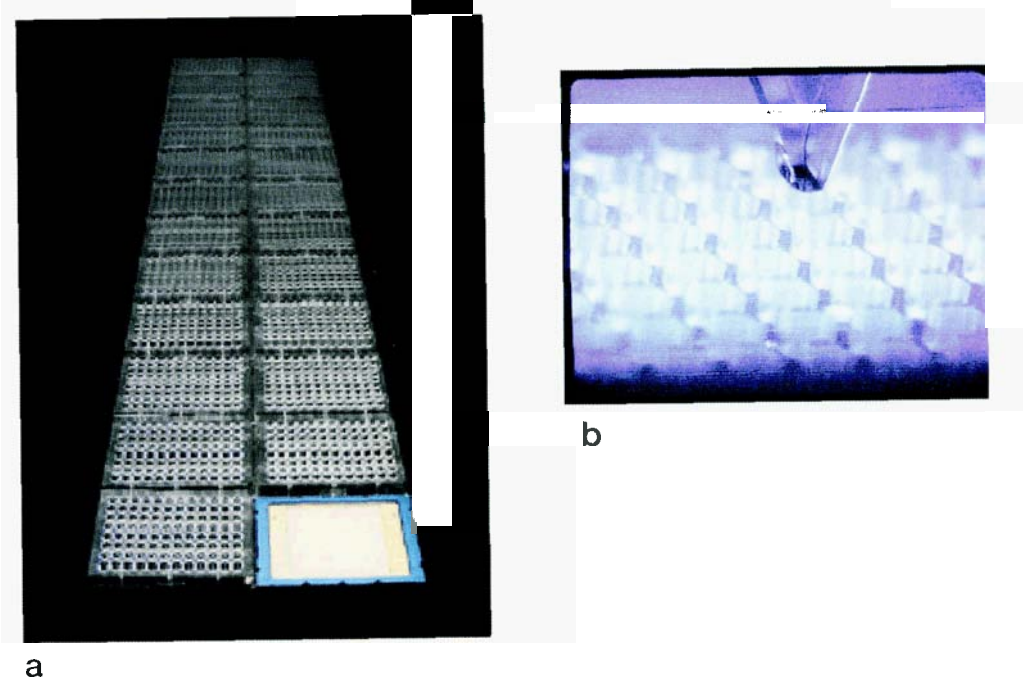
**Figure 13.8.** Comparison of pairwise biological distances versus descriptor differences for 138 angiotensin-converting enzyme (ACE) inhibitors (upper panel) as example for neighborhood plots: (a) CoMFA steric fields; (b) 2D fingerprints; (c) molecular weight.

### Confocal Fluorescence Detection: Multifunctional Analysis



**Figure 14.3.** A survey of read-out modes for confocal fluorescence detection technologies and their applications to biological systems. NA = nucleic acid.

## Nanodispenser and 2000-well Nanocarrier™



**Figure 14.4.** (a) A comparison of the reduction in space achieved by using miniaturized sample carriers. A total of 2000 samples may be carried by either the 23 microtiter plates seen in this picture, or by a single 2000-well nanocarrier (lower right). (b) The head of a nanodispenser situated over one of the wells of a 2000-well nanocarrier.

This Page Intentionally Left Blank

# Index

- acetoacetanilides 96
- acetonedicarboxylic acid 153
- acetophenones 91
- 4-acetyl-3,5-dioxo-1-methyl-cyclohexyl-carboxylic acid 188
- acetylene 96
- 4-acetyloxybenzyloxy resin 220
- acid-labile anchors 171
- acids for acylations 81
- acrylic acid linker 187
- actinorhodin 381, 391
- Actipat prediction programme 371, 372
- activation of acids 81
- activity pattern 366, 368
- acyl carrier protein 391
- [acyl-4-(4-oxymethyl)phenylacetoxymethyl]-3-nitrobenzoic acid 216
- acyl-2-[(oxymethyl)phenylacetoxyl]-propionyl-PS/DVB 216
- 2-acylamino-alkanoic acid amides 94
- acylation 81
  - of alcohols 81
  - of amines 81
- acyltransferases 391
- ADCC linker 187, 188
- addition reactions 87
- adenine 134
- ADPV linker 173
- affinity chromatography 321
- affinity fingerprints 417
- aldehydes by oxidation 63
- aldol reactions 67
- aliphatic polyamines 87
- 5-alkoxy-hydantoins 229
- N-alkoxy urea 39
- alkyl bromides 87
- alkylation 68, 87
  - of amides 69
  - of 2-arylsulfonylaminomethyl acrylic acids 243
  - of carbon atoms 67
- N-alkylation 69
- O-alkylation 68
- alkyl- $\beta$ -(N,N-dimethyl amino)- $\alpha$ -isocyanoacrylate 157
- 2-alkylidene arylidene- $\beta$ -ketoesters 250
- N-alkyl-oligocarbamates 263
- alloepitope mimic 362
- alloreactivity 361
- allyl-functionalized anchors 204
- allyloxy-pyrazoles 95
- amide 42
  - alkylations 69
  - bond formation 38
  - vibration 484
- amide-based linker 51
- amidoxime 52
- amine addition 59
- amino acid fluorenylmethyl ester 480
- amino acid pentafluorophenylester 481
- $\alpha$ -amino acids 128
- $\alpha$ -amino alcohols 481
- $\alpha$ -amino aldehydes 481
- 5-(2-aminoacetyl)pyrroles 246
- $\alpha$ -amino acid derivatives 147
- 1,3-amino alcohol 234, 235
- $\alpha$ -aminoalkylation 130
- 3-aminoaryl-2-methylidene sulfonylpropionic acids 233
- 2-amino-benzimidazoles 100
- 3-amino-benzophenones 83
- N,N-bis-(4-aminobutyl)-aminocarboxy-oxy-methyl-phenyloxyacetic acid 174
- amino-carbohydrates 147
- N-(2-aminoethyl)glycine monomers 280
- 5-aminoindoles 85
- $\alpha$ -aminoisobutyric acid peptides 280
- 7-amino-4-methylcoumarin-3-acetic acid 363
- 5-(4-aminomethyl-3,5-dimethoxy-phenoxy)-valeric acid 173
- 4-aminomethyl-3-nitrobenzoic acid 200
- 4-aminomethyl)phenoxymethyl-PS/DVB 173
- 3-amino-3-(2'-nitrophenyl)-2,3-dimethylpropionic acid 200
- 3-amino-3-(2-nitrophenyl)propionic acid 200
- 2-aminopropanol 173

- $\beta$ -aminopropionic acid oligomers 262
- 2-aminopyridines 247
- 2-aminopyrimidines 247
- amino-thiazoles 98
- aminothiazolopyrroles 246
- amplification cycle 321
- analysis of combinatorial libraries 499
- analyte 335
- analyzer cell, ICR-MS 563, 565
- anchor positions, MHC 358
- anchor, definition 168
- anchors cleaved by nucleophiles 187
- anchors cleaved by oxidation 219
- angucycline antibiotic 388
- aniline 65
- antennary glycans 293
- anthracyclines 390
- antibiotic production 381
- antigen recognition 372
- antigen, presentation of 355
- anti-Rev aptamers 327
- aptamer/ribozyme construct 323
- aptamers for small molecules 390
- Arbuzow rearrangement 242
- ARCoSyn work station 553
- arenesulphonyl linker 39
- L-arginine aptamer complex 323
- arginine-protecting groups 52
- aromatases 391, 395
- aromatic substitution, electrophilic 42
- aromatization 64
- aryl ethers 107, 483
- aryl zinc bromides 45
- N-arylation 45
- arylation of amines and alcohols 90
- arylhydrazines 97
- 3-arylsilyl-propionic acid 212
- arylsulfide 207
- arylsulfonates 217
- Asinger, 4CR 130
- association experiments 458
- association kinetics 457
- atmospheric pressure chemical ionization 504
- atom pair fingerprints 414
- attenuated total reflection spectroscopy 482
- autoantigen in multiple sclerosis 373
- autocorrelation vectors 415
- autoimmune antibodies 326
- autoimmune diseases 361
- automated purification 25
- automation 24, 543
  - concepts 545
  - of LC-MS 506
  - platform 553
- autosampler 506
- autosampler run, MS 513
- autosampler-assisted devices 549
- avermectin 381, 400
- Aza Wittig reaction 41, 60, 159
- $\alpha$ -azaamino acid monomers 277
- azapeptides 277
- azatides 277
- $\alpha$ -azido acids 280
- backbone amide linker (BAL) 178
- backbone modifications 267
- backbone-based libraries 4
- Baylis-Hillman reaction 60, 233
- bead measurements by NMR 535
- bead picker 453
- bench-top synthesizers 547
- benzamide-PS/DVB 207
- benzene triacetic acid 81
- benzhydramine-PS/DVB 173
- benzhydriyl-cation 171
- benzimidazole 57
- benzodiazepine-2,5-dione 57
- benzodiazepine-2-one 57
- benzodiazepines 42, 115
- benzodiazepines by U-4CR 158
- benzofuran 44, 58
- benzofurylpyrrolyl ketones 248
- benzoic acids 82
- benzopyran 58
- benzoxazinones 107
- benzoylation of amino methyl resin 487
- benzoylation of sulfonamides 107
- benzyl bromides 87
- O-benzyl hydroxamic acids 92
- benzyl-cation 171
- O-benzylhydroxylamine 282
- $\alpha$ -(benzyloxyamino)- acid esters 84
- p*-(benzyloxy)benzyl alcohol resin 488
- p*-benzyloxybenzylamine-PS/DVB 216
- benzyloxycarbonyl anchor group 480
- bergenin 86
- Bergmann-Bucherer, 4CR 130
- BHA anchor 173
- biased library 26
- bicyclomycin 142
- Biginelli reaction 116
- Biginelli, 4CR 130
- binding affinities 417
- binding experiments 458
- biosynthesis 381
  - of daunosamine 402
  - of microbial metabolites 381
  - of oligosaccharide antibiotics 403
- biosynthetic pathways 382
- biotechnology applications 553
- Bischler-Napieralski 56
- bis-thiazole 272
- Boc-iminodiacetic acid 83
- Boc-piperazines 90

- Boc-S-methylisothiourea 88  
borohydride resin 107  
boronic acid 43  
4-(2-(Br or Cl)phenoxy)acetic acid 189  
bromoacetic acid derivatives 87  
N-(N-bromoacetyl)-2-aminoethyl-polyacrylamide 188  
bromoacetyl-PS/DVB 188  
2-bromoacetyl-PS/3% DVB 189  
5-(2-bromoacetyl)-pyrroles 249  
5-(2-bromoacetyl)pyrroles as templates 244  
bromocarboxylic acid 492  
4-bromocrotonic acid 205  
4-(1-bromoethyl)-3-nitrophenyl-PS/DVB 202  
2-bromo-4-(hydroxymethyl(dimethylsilyl)benzotrile 212  
2-bromo-2-iodo-acetic acid 189  
 $\alpha$ -bromo ketones 87  
4-bromo-(or hydroxy)methyl-3-nitrobenzoic acid 201  
4-bromo-(or iodo)methylphenoxy-PS/DVB 174  
4-(2-bromopropionyl)-3-nitrophenylacetic acid 200  
4-(2-bromopropionyl)phenoxyacetic acid 201  
4-(2-bromopropionyl)phenyl-PS/DVB 189  
3-bromo-thiophene 95  
3-bromo-1-tosyl-pyrrol 95  
butane isomerization 468  
3-(*n*-butyl) 4-isopropoxy squarate 233  
N-tert-butyloxycarbonyl-N-methoxypropionic acid 217  
tert.-butylisocyanide 145
- calixarenes 336  
(2OS)-camptothecin 133  
capillary electrophoresis 505  
carbamate linker 49  
carbamate-based linkers 50  
carbamates 40  
carbohydrate ligands 293  
carbohydrate recognition domain 292  
carbohydrate-based therapeutic 305  
carbohydrate-binding proteins 291  
carbonates 40  
carbon-carbon bond formation 42  
carbonyl vibration 484  
4-carboxybenzaldehyde 174  
4-carboxybenzenesulfonamide 207  
3-carboxypropanesulfonamide 208  
carboxy-PS/DVB 190  
4-carboxytrityl alcohol 174  
catalysis 463  
catalyst development 466  
catalyst synthesis 470  
cathepsin D and E 361  
cation exchange resin 110  
cations 171  
cavity-forming structures 340  
cell lysis assays 360  
cellular assays 455  
chain length factor proteins 391  
characterization of combinatorial libraries 14  
characterization of resin-bound libraries 540  
chemical encoding strategies 21  
chemical ionization 503  
chemosensors 335  
chiral separation 337, 341  
 $\alpha$ -chloro ketones 87  
chloroacetaldehyde 156  
4-chloromethyl-3-nitrobenzyl-PS/DVB 202  
chloromethylloxymethyl-dimethylsilyl-bromobenzene 212  
chloromethyl-PS/DVB 175, 190, 218  
2-chloro-tritylchloride-PS/DVB 175  
Claisen reaction 53  
Claisen-type condensation 106  
cleavage  
– of alcohols 49  
– with hydroxylamine 51  
– from the polymeric carrier 169  
cluster analysis 416  
clustering 412  
cobalt(III)-complex 190  
collectins 293  
collision activated dissociation 567  
collision induced dissociation 567  
collision-induced decay 511  
column peptide synthesis 307  
combinatorial chemistry programmes 30  
combinatorial library 2, 4  
combinatorial mixtures 79  
components of a heterogeneous catalyst 465  
compound libraries 409  
compound selection techniques 415  
computational methods 412  
computer-assisted library design 410  
computer-controlled robot system 24  
condensation reactions 53, 91  
confocal fluorescence 448  
confocal microscopic optical systems 448  
contour plot HPLC/FT-ICR-MS 573  
control of synthesis 481  
convergent combinatorial synthesis 105  
cosmid libraries 388  
coupling reactions, Pd-cat. 45  
cross-metathesis 62  
crown ethers 338  
cryomagnet system 562  
cubane tetracarboxylic acid 81  
cyano absorptions 495  
cyanoborohydride 107  
3-cyanopyridine 251  
cyclases 391, 395  
cyclative cleavage 46

- cyclic anhydrides 83  
cyclic hexapeptide 153  
cyclization 46  
1,3-cycloaddition 237  
[2+2]cycloaddition 54  
[3+2]cycloaddition 54  
[4+2]cycloaddition 55  
cycloaddition  
– 1,3-dipolar 484  
– of the maleimides 491  
– with nitrile oxides 237  
cyclocondensations 154  
cyclodextrins 337  
cyclohexapeptide monolayers 346  
cyclohexapeptides 341, 342  
cyclohexenyl isocyanide 149  
cyclohexenyl isonitrile 94  
cyclooligocarbamates 264  
cyclopentapiperidinone 56  
cyclopeptide libraries 341  
cyclopeptides 342  
cyclophanes 337  
cyclopropanation 60  
cyclo-reversion 47  
cyclothioureas 270  
cyclotron equation 565  
cyclotron motion 564  
3-cysteiny-4,6-dinitroanilide of N-carboxyme-  
thyl-aminomethyl-PS/DVB 191  
cytolysis of target cells 360
- data processing 24  
database comparison histograms 433  
database for new prediction methods 371  
database interfaces 555  
daughter ion spectra 578  
daunomycin derivatives 390  
Dde-based primary amine linker 191  
debenzylations 50  
decarboxylation 47  
decentralized automation systems 549  
decoding method 21  
deconvolution 16  
degeneracy in antigen recognition 373  
dehydrogenation of cyclohexane 473  
dendrimer-supported combinatorial chemistry  
118  
deoxysugar biosynthesis 402  
descriptor validation studies 420  
descriptors 413  
design for global diversity 422  
design strategy 410  
desulphurization 47  
detection of side reactions 519  
detection principle, MS 564  
deuterated reactants 486  
 $\alpha,\alpha$ -dialkylated glycines 279
- L-2,3-diaminopropionic acid(phenyloxycar-  
bonyl), Dpr(Phoc)linker, 211  
2,6-diaminopurines 88  
3,5-diaryl-isoxazolines 97  
1,3-diaryl-1-propenones 96  
1,4-diazabicyclo[2.2.2]octane 234  
diazepine-2,5-dione 57  
diazo transfer 60  
dibenzoylmethanes 103  
1,4-dibromo-2,3-butanedione 244  
dicarboxylic acid esters 96  
2,6-dichloropurine 90  
2,4-dichloro-triazines 90  
2,6-dideoxy irisacharides 303  
Dieckmann condensation 106  
Diels-Alder-reactions, hetero 55  
2-diethoxy-phosphorylmethyl acrylic acids 242  
diffuse reflectance infrared Fourier transform  
spectroscopy, DRIFTS 496  
4,6-difluoro-1,3-dinitrobenzene 191  
dihydro thiazinones 115  
2,3-dihydro-4H-pyran 171  
1,6-dihydro-6-oxopyridine-2-carboxylic acid 159  
dihydropiperidine 56  
dihydroprotodaphniphyllite 133  
1,2-dihydropyran carboxylic acid 175  
dihydropyridine 56  
1,4-dihydropyridines 251  
dihydroquinazoline 57  
dihydroquinolinone 56  
1,4-dihydroxynaphthalines 233  
diimines, synthesis of 103  
diisooctylphthalate 522  
2,4-dimethoxy-4'-hydroxybenzophenone 176  
2,4-dimethoxy-5-((4-methoxyphenyl)-aminome-  
thyl)-phenylpropionic acid 176  
4-(2',4'-dimethoxyphenyl-Fmoc-aminomethyl)-  
phenoxyacetic acid 176  
4-(2,4-dimethoxyphenyl-Fmoc-aminomethyl)-  
phenoxymethyl-PS/DVB 177  
4-[2-(4,4'-dimethoxytriphenyl-methoxy)-  
ethylthiomethyl]-benzoic acid 208  
dimethyldioxirane 103  
dimethyldysidenin 142  
5-[dimethyl-(2-hydroxyethyl)-silyl]-valeric acid  
213  
4-(2,5-dimethyl-4-methylsulfinylphenyl)-4-hy-  
droxybutyric acid 208  
3-[dimethyl-(phenyl)-silyl]-propionic acid 213  
*o,o'*-dimethyl-xanthocillin 134  
diolefinic linkers 205  
3,3-diphenyl-2-indolones 103  
diphenylmethyl isocyanide 152  
diphenylphosphine 107  
2,6-dipiperazinediones 156  
5 – N,N-dipropargyle amide 485  
directed library 26

- dissimilarity 415  
 diversity  
   – analysis 409  
   – assessment 26  
   – descriptor 421  
 diversity-based design 432  
 domain replacement 398  
 doxorubicin 402  
 drug discovery 441  
 drug discovery process 2  
 drug screening 443  
 DSB resin 208
- Edman degradation 21  
 electrode array 472  
 electron impact 503  
 electrophilic aromatic substitution 61  
 electrospray ionization 503  
 electrospray ionization-mass spectrometry 508  
 electrospray mass spectrometry of peptide libraries 365  
 electrospray mass spectrometry 528  
 elementary analysis 570  
 $\beta$ -elimination 62  
 eloidisine 142  
 enamine 53  
 enamine-based linkers 51  
 encoding 21  
 endothelial growth factor 326  
 ene rearrangement 60  
 enones as templates 249  
 enzymatic acylation 85  
 enzymatic subunits 390  
 enzymatically cleavable anchors 219  
 enzyme acceptor 455  
 enzyme donor 455  
 epirubicin 402  
 epoxidation 64, 234  
 epoxidation of olefins 103  
 epoxides 88  
 erythromycin 381, 384  
 erythronolide macrolactone 384  
 Eschenmoser's salt 158  
 ES-MS-analysis 514  
 evaporators 549  
 EVO screen system 448, 451  
 exciting frequencies 566  
 expression of single enzymes 398  
 extraction by scavenger 113  
 extract-ion-function 525
- fast atomic bombardment 503  
 feature tree descriptor 418  
 $\alpha$ -ferrocenyl alkyl amines 146  
 $\alpha$ -ferrocenyl- $\beta$ -hydroxyethylamine 146  
 fingerprints, 2D 413, 424  
 flow injection 571
- fluorescence  
   – correlation spectroscopy 450  
   – cross correlation 449  
   – depolarization 449  
   – detection 449  
   – imaging 451  
   – polarization 451  
   – read-out technologies 448  
   – resonance energy transfer 449, 450  
 3-fluoro-4-(hydroxymethyl)phenoxyacetic acid 177  
 fluorous synthesis 13, 116  
 fluorous tin reagents 44  
 4-(N-Fmoc-1-aminoethyl)-2-methoxy-5-nitro-phenoxybutyric acid 202  
 N-Fmoc-2-aminomethylthiazole-4-carboxylic acid 272  
 9-Fmoc-amino-xanthen-3-ol 177  
 focused library 26, 411  
 5-(4-formyl-3,5-dimethoxy-phenoxy)-valeric acid 178  
 3-formyl-indol-1-yl-acetic acid 178  
 3-formyl-pyrrole 91  
 3-formyl-thiophene 91  
 Fourier transform ion cyclotron resonance 561  
 fragmentations in ES-MS 511  
 frenolicin 396  
 frequency shifts of quartz microbalance 344  
 fridamycin 389  
 fucosylated tripeptides 307  
 fuel cell catalysts 472  
 functional aptamers 327  
 function-oriented automation concept 548  
 function-oriented multicomponent systems 550
- $\beta$ -galactosidase 455  
 gas chromatography 521  
 gas-phase reaction 473  
 GC-EI-MS-analysis 521  
 GC-MS-analyses 529  
 gel-phase NMR spectroscopy 533  
 gene  
   – clusters 381, 388  
   – disruption 384  
   – probes 382  
   – replacement experiments 402  
 generic libraries 26  
 genetic algorithms 28  
 genetic engineering 382  
 glass transducers 346  
 global diversity 420  
 glutaric acid 196  
 glutaric acid-mono-4-alkoxyanilide 209  
 glutaric acid-mono-(3- or 4-)-[(4-hydroxymethyl)-phenoxy)-(*t*-butyl-phenyl)-silyl]-phenyl]-anilide 213  
 glutathione 142



- glycal procedures 302  
glycan specificity of T-cell receptor recognition 309  
glycohydrolases 302  
glycol linker 178  
glycopeptide 291  
– arrays 307  
– inhibitor 312  
– library 310  
– templates 305  
glycoproteins 291  
glycosamine 142  
glycosyl donors 301  
glycosylation 293  
– reaction 301  
gold electrodes 343  
gradient sputtering 469  
Grignard 103  
– reactions 55  
– reagents 51  
gripper changing system 552  
guanidation of amines 41  
guanidines 39, 41
- haloaromatic phenol ether diazoacetate 21  
halodediazotization 48  
halodesilylation 48  
HAL-resin 172, 180  
hammerhead ribozyme 323  
Hantzsch reaction 251  
Hantzsch, 4CR 130  
Heck reactions 44, 56, 58  
high-affinity reporter peptides 356  
high-performances size-exclusion chromatography 362  
high-throughput analysis 500, 506, 511, 571  
high-throughput experimentation 463  
high-throughput purification systems 507, 556  
high-throughput screening 15  
hit identification 15  
HMBA 194  
HMPB resin 180, 181  
Hoffmann elimination 50  
Horner-Emmons based cleavage 48  
Horner-Emmons reaction 62  
host-guest interactions 342  
HPLC purification 26  
HPLC-MS-analysis 522  
HPLC-MS-MS-analysis 516  
hybrid antibiotic 390  
Hycram resin 205  
Hycron linker 206  
hydantoin derivatives 57, 128  
hydantoin syntheses 229  
4-hydrazinobenzoic acid 209  
hydrazone-like linkers 51  
hydroboration 43  
hydrogen cyanide 134  
hydrogenations 92  
hydrogenolysis-labile anchors 218  
hydrolysis of phosphonium linker 47  
N-hydroxy glycine 283  
2-hydroxy-acetophenones 82  
N-(hydroxy-(alkyl or aryl)-dimethylformamidine 191  
N-hydroxyamide bond-containing peptides 282  
*p*-hydroxybenzhydramine 217  
5-(4-hydroxybenzyl)-oxazolidone 192  
4-hydroxybut-2-en-oxy-acetic acid 206  
4-hydroxybut-2-en-oxy-caproic acid 206  
4-(4-hydroxybutyl-dimethylsilyl)-2-trimethylstannyl-N-Bpoc-aniline 213  
9-hydroxy-9-(4-carboxyphenyl)fluorene 179  
9-hydroxy-9-(4-carboxyphenyl)xanthene 179  
hydroxyethylamine peptide bond isosteres 283  
hydroxyethylene-PS/DVB 192  
4-[(1-hydroxyethyl)-2-methoxy-5-nitro]-phenoxybutyric acid 202  
5-(2-hydroxyethyl)-mercaptoacetic acid 192  
4-hydroxyethyl-3-nitro-benzoic acid 193  
4-(hydroxyethylsulfonyl)-benzoic acid 192  
4-(9-hydroxyfluorenyl)benzyllic acid 179  
4-(6-hydroxyhexyl-simethylsilyl)-2-trimethylgermanyl-N-Bpoc-aniline 214  
N-hydroxylation 283  
8-hydroxy-1-mercaptopodibenzofuran 193  
4-hydroxy-2-methoxybenzyl alcohol 193, 179, 217  
4-hydroxymethylbenzoic acid 194  
hydroxymethylbenzoic acid, HMBA 180  
3-hydroxymethyl-2,3-dihydro-4H-pyran 180  
5-(4-hydroxymethyl-3,5-dimethoxy-phenoxy)-valeric acid 180  
4-(hydroxymethyl(dimethylsilyl)-1-bromobenzene 212  
3-(9-hydroxymethylfluorenyl)-acetic acid 195  
3-(9-hydroxymethylfluorenyl)-amido-butanedioic acid 195  
1-(9-hydroxymethylfluorenyl)-formic acid 194  
3-hydroxy-2-methylidene propionic acids 233  
4-hydroxymethyl-2-methoxy-5-nitro-phenoxybutyric / propionic acid 202  
4-hydroxymethyl-3-methoxyphenoxy-acetic acid 181  
4-(4-hydroxymethyl-3-methoxyphenoxy)-butyric acid 180  
4-hydroxymethylphenol 181, 195, 217  
4-hydroxymethylphenoxyacetic- or -propionic acid 182  
4-(4-(hydroxymethyl)phenoxy-methyl)-dimethylsilyl)-bromobenzene 214  
4-hydroxymethylphenylacetic acid 182, 195  
3-[(3-hydroxymethyl)-piperidylazo]-bromobenzene 196

- 3-(4-hydroxymethylphenyl)-3-(trimethylsilyl)-propionic acid 214  
 hydroxymethyl-PS/DVB 182, 194  
 hydroxymethylvinyl-PS/DVB 219  
 hydroxy-(2'-nitro)-benzhydryl-PS/DVB 203  
 3-(4-hydroxyphenyl)-propionic acid 196  
 4-hydroxyphenylsulfonylchloride 196  
 $\omega$ -(3-hydroxypropen-1-yl)-alkanoic acid 206  
 6-hydroxy-7-(3-propionic acid)-4,4,5,8-tetramethylhydrocoumarin 218  
 3-hydroxy-quinuclidine 234  
 4-hydroxytrityl alcohol 183
- IdaDC, iminodiacetic-acid double-cleavable linker 210  
 image analysis 497  
 imidazo[1,2-a]pyri(mi)dylypyrroles 247  
 imidazocarbonates 40  
 imidazole 57, 271  
 3-imidazol-4-yl-2-hydroxypropionic acid 209  
 imide 41  
 imine-formation 53  
 iminodiacetic acid linker 210  
 iminodiacetic diamides 83  
 $\alpha$ -iminodicarboxylic acid 156  
 imino-thiazoles 98  
 immobilization of a cyclohexapeptide 348  
 immune modulators 266  
 immune response 355  
 immunomodulating peptides 376  
 immunosuppressant FK 506 384  
 impregnation techniques 470  
 impurities from linker or resin components 509  
 incubation stations 549  
 indazole 57  
 indoles 44, 56  
 2-indolinones 91  
 indol-resin 178  
 indolyl-5-amines 81  
 influence of reagents on MS-analysis 509  
 influenza hemagglutinin 367  
 information logistics 550  
 infrared mapping of beads 490  
 infrared spectroscopy 479  
 infrared tagging 495  
 infrared-microscopy 479, 484  
 infrared-microscopy time resolved 486  
 infrared-Raman microscopy 497  
 infrared-Raman spectroscopy 497  
 inhibitors of galactosidase 307  
 insulin-dependent diabetes mellitus 361  
 intercellular adhesion molecules 329  
 intrabodies 330  
 intramers 330  
 intramolecular cyclization 42  
 intramolecular Wittig reactions 48  
 2-iodo glycoside 303  
 iodo-acetic acid 189  
 iodolactonization 48  
 ion cyclotron resonance 561  
 ion packages 566  
 ionization methods 562  
 ionization techniques 499  
 IR thermography 471  
 isatines 100  
 isatoic anhydride 112  
 isocyanates 39, 41  
 isocyanide chemistry 130  
 isocyanides 125  
   – synthesis of 125, 135  
 isonitrile chemistry 135  
 isoquinoline 44  
 isotope resolution 573  
 isoxazoles 54, 58  
 isoxazolidine library 491  
 isoxazolidines 484  
 isoxazoline library, MS 523  
 isoxazolines 237, 238  
 iterative deconvolution 16
- Kaiser oxime resin 197  
 $\beta$ -ketoesters 250  
 ketone 63  
 ketoreductases 391  
 ketoreduction 395  
 ketosynthase 391  
 Knoevenagel reaction 53, 250  
 Knorr resin 176  
 Kröhnke pyridine synthesis 251
- laboratory automation 545  
 laboratory technologies 553  
 $\beta$ -lactam derivatives 54, 151  
 $\gamma$ -lactam derivatives 151  
 ladder synthesis 300  
 landscape-plot, LC-MS 526  
 Langmuir adsorption 349  
 lanosterol 132  
 laser-induced resonant multiphon ionization 473  
 Lawesson's reagent 281  
 LC-MS couplings 529  
 lead discovery 409  
 lead finding 27  
 lead optimization 27  
 lectins 291  
 Leu-enkephalin azatide 277  
 libraries of compound mixtures 15  
 libraries, types of 26  
 library design 26, 410  
 library design procedure 428  
 library generator 311  
 ligand for  $\alpha_1$ -adrenergic receptor 259  
 ligand for MHC-I molecules 266  
 light-directed synthesis 10

- linker, definition 168  
lipase catalyst 85  
liquid chromatography 505  
liquid handling 452  
liquid-phase extraction 114  
liquid-phase synthesis 13  
lithiation 95
- macrocyclic peptides 339  
macrocyclization 153  
– by Heck reaction 44  
macrolactone 398  
macrolides 400  
magic angle spinning 533  
magneto-resistive materials 463  
major histocompatibility complex 264, 355  
malonic acid 196  
MAMP, alpha-methoxy phenyl resin 184  
Mannich reaction 53  
mannose ligands 293  
mannose 6-phosphate receptors 293  
MAS-NMR of saccharides 300  
mass accuracy 569, 571, 578  
mass spectrometric complications 510  
mass spectrometric systems 561  
mass spectrometric techniques 502  
mass spectrometry 499, 561  
– of oligosaccharides 298  
mass-sensitive sensors 342  
match search algorithm 419  
matrix-assisted laser desorption 503  
MBHA resin 185  
measles virus 372  
4-( $\alpha$ -mercaptobenzyl)-phenoxyacetic acid 183  
2-mercaptoethanol 197  
mercaptoethylamine resin 113  
4-mercaptophenol 197, 210  
3-mercapto-1,2-propanediol 183  
3-mercaptopropionic acid 197  
Merrifield-resin 172, 175, 190  
methodology for library construction 12  
4-methoxy-benzhydryl chloride-PS/DVB 184  
4-methoxy-benzhydrylamine-PS/DVB 183  
1-[methoxycarbonylmethyl]-pyridinium bromide 252  
4-methoxymethoxy- (or bromo-)phenyl-diisopropyl-hydroxysilane 214  
2-methoxy-5-(2-(2-nitrophenyl)dithio)-1-oxopropylphenylacetic acid 203  
4-[(4-methoxyphenyl)-aminomethyl]-phenoxyacetic acid 184  
4-methoxytritylchloride-PS/DVB 184  
4-methylbenzhydrylamine-PS/DVB 185  
N-methyl-benzylamine 81  
methylene ether isosteres 284  
2-methylidene- $\beta$ -alanines 238  
N-methyl-pyrrole-imidazole polyamides 272  
N-methylsuccinimide resin beads 493  
2-(4-methylsulfoxy)phenyl(aminomethyl)-5-methylsulfoxyphenoxyacetic acid 210  
O-methyltransferase 382  
4-methyltritylchloride-PS/DVB 185  
MHC class I ligands 267, 358  
Michael addition 58, 235  
– of aryl thiolates 250  
microcin B17 272  
microflow reactor 476  
microLC/ES-FT-ICR-MS 573  
micromanipulator 575, 577  
microreactor systems 446  
microwave-assisted reactions 102  
microwave irradiation 44, 92  
miniaturization 443  
mirror-image aptamers 325  
mirror-image in-vitro selection approach 324  
mithramycin 390  
Mitsunobu reactions 68, 69, 107, 236  
MOBHA resin 183  
Moffat oxidation of amino alcohols 103  
molecular descriptors 413  
molecular diversity 412  
molecular formula determination 570  
molecular receptors 335  
molecular similarity 416  
monitoring of reactions 480  
monolayer of cyclohexapeptides 347  
monomer approach to peptoid synthesis 259  
morpholinone derivative 284  
motif amino acids 359  
mucin oligosaccharides 294  
Mukaiyama aldol reaction 108  
multicolor assays 449  
multicomponent reactions 93, 125, 128  
multicomponent systems for automation 545  
multifunctional analysis 449  
multifunctional anchors 216  
multiphon infrared photodissociation 567  
multipin-technique 9  
multiple cleavable linkers 22  
multiple core structure libraries 229, 230  
multiple sclerosis 372  
Münchnone intermediate 158  
myasthenia gravis 326, 368  
myelin basic protein 373
- nanocarrier<sup>TM</sup> 452  
nanodispenser 452  
nanospray sources 578  
naphthalines 233  
1,4-naphthoquinone 233  
NBA resin 200  
Nbb resin 201  
NBHA resin 203  
neutral loss experiment 520

- nikkomycin 142
- (*o*-nitrobenzenesulfonyl chloride) 262
- 2-nitro-benzhydrylamine-PS/DVB 203
- p*-nitro-benzhydrylamine-PS/DVB 219
- p*-(4-nitrobenzophenoxime) resin 197
- nitrogen-containing heterocycles 56
- p*-nitrophenol chloroformate 49
- 1-nitrophenyl piperazines 81
- 2-nitrophenyl-1,3-propanediol 204
- 6-nitrovanillin 204
- NMR experiments 539
- NMR spectroscopy 533
- nonapeptide amide library 370
- nonapeptide libraries 359
- nosyl-like linkers 50
- NPE resin 193
- Npss resin 203
- nuclease-resistant aptamers 326
- nucleic acid aptamers 319
- nucleic acid libraries 319
- nucleophile, immobilized 59
- nucleophiles 81
- nucleophilic aromatic substitution 108
- nucleotide phosphoramidite 21
  
- off-bead analysis, MS 506
- olefin formation 61
- olefination reaction 64
- olefine metathesis 80
- olefinic linker 219
- oligocarbamates 263
- oligocycloureas 270
- oligosaccharides
  - arrays 301
  - libraries 291
  - mimetics 305
  - synthesis on solid phase 303
- oligotetrahydrofurans 273
- oligothiophenes 273
- oligoureas 268
- on-bead analysis 479
  - by MS 506
  - by NMR 533
- one-bead-one compound libraries 6
- one-pot U-4CR 142
- optical thickness 348
- optimization methods 28
- organo-lithium compounds 103
- orthogonal libraries 19
- 1,2,4-oxadiazoles 98
- oxalic acid 198
- oxazole 272
- oxazolones 84
- oxazolyl-thiazole 272
- oxidation 62
  - catalysts 464
  - of *p*-alkoxy-benzyl linker 49
  - of thioethers 103
  - of propene 464
  - reactions of hydrocarbons 472
- oxindole 44
- 2-oxindole 56
- oxygen-containing heterocycles 58
- oxytetracycline 381
- ozonolysis 62, 64
  
- PAA oligomers 267
- PAL resin 173
- palladium catalysts 56
- palladium-catalyzed
  - allylation 69
  - amination 45
  - C-bond formation 92
  - coupling reactions 45
- PAM resin, phenylacetamidomethyl resin 182, 195
- parallel reactor for catalysis 474
- parallel synthesis 7, 544
- parent ion scan 520
- parent library 410
- Passerini, P-3CR 131
- Pauson Khand reaction 56, 60
- PBTH 215
- Pd-catalysts 43, 56
- PEG-based resins 297
- penta-ethylene hexamine 114
- pentane-1,3,5-tricarboxylic acid 1,3-anhydride 198
- pentenylglycoside procedure 302
- peptide
  - alkaloids 142
  - antibiotic surfactin 401
  - biosynthesis 399
  - bond modifications 281
  - chemistry, by MCR 141
  - by gene replacement 401
  - lead structures 356
  - nucleic acids 143, 280
  - selection by MHC class I 358
  - synthetases 399
- peptide libraries
  - amide 362
  - $\beta$  279
  - MS 504
  - permethylated 261
  - libraries, synthesis of 362
- peptide ligands of MHC class I 282, 358
- peptide ligands of MHC class II 359
- peptide-binding groove of MHC class I 357
- peptide-MHC class II complexes 360
- peptidesulfonamides 340
- peptidolactone 401
- peptidomimetics 257
- peptidosteroidal receptors 340

- peptidylphosphonates 285  
peptoid synthesis 258  
 $\beta$ -peptoids 262  
peptomers 260  
peralkylated peptides 261  
pharmacophore definition triplets 414, 424  
phenol-PS/DVB/phenol-PEG 198  
phenyl boronic acid 92  
phenyl isocyanide 133  
phenylalanine ester 220  
phenylcarbamate 40  
2-phenyl-2', 2''-dithiane-3-hydroxyethyl-  
  phenoxymethyl-PS/DVB 210  
1,4-phenylene-diisocyanate 114  
9-phenylfluorene-9-yl-PS/DVB 185  
phenylglyoxylic acid 159  
3-N-phenyloxycarboxy-2,3-diaminopropionic  
  acid 211  
N-phenylpyrazole 254  
PhFL resin 185  
phosgene 49, 185  
phosphate ester 186  
phosphonate-containing peptide 285  
phosphonic peptides 285  
phosphoracid diester 220  
phosphoramides 42  
phosphorescent materials 463  
phosphorylated glycopeptides 305  
photoacoustic spectroscopy 497  
photocleavage 22  
photolithographic techniques 9  
photolysis-labile anchor 200  
phtalocyanines 339  
Pictet-Spengler reaction 57, 58, 480  
piperazinediones 156  
2,6-piperazinediones 150  
piperazines 87, 157  
pipetting robots 549  
pivaloylglycol 204  
plausibility control 555  
plumbemycin 142  
poly-N-acylated amines 266  
polyallylamines 341  
polyamine resin 112, 113  
polyaza macrocycles 90  
polyazacyclophanes 87  
polycyclizations 132  
polycycloadditions 132  
polyethylene pins 9  
polyisoxazolines 273  
polyketide synthase genes 393  
polyketide synthases 383, 391  
polyketides 278  
polymerase chain reaction 21  
polymer-bound aryl iodides 43  
polymer-supported aldehyde 112  
polymer-supported isocyanate 112  
polymer-supported phosphines 118  
polymer-supported synthesis 107  
poly-N-methyl pyrroles 271  
polymorphism of MHC class II 360  
polyoxins 142  
polypeptide synthetase 401  
 $\beta$ -polypeptides 278  
polystyrene 107  
  – NMR 537  
polystyrene hydroxymethyl resin 485  
Pon 211, 216  
pool sequencing 356, 365  
Pop 211, 216  
positional scanning 18  
precipitation reactions 471  
prediction algorithm 372  
prediction of epitopes 375  
predictions of HLA-ligands and T-cell epitopes  
  371  
pristinamycin 387  
privileged structures 26  
progesterone 132  
propane-1,3-dithiol 198  
propene oxidation 468  
protein antigens 361  
protodesilylation 47  
PS-benzyloxycarbonyl resin 480  
PS/DVB-di-*n*-butyltinhydride 215  
purification 25, 106  
  – using MS-detection 507  
pyrazole 53, 57  
pyrazole carboxylic acid 107  
pyrazole-3-carboxylic acid 83  
pyrazole-5-carboxylic acid 82  
pyrazolones 234, 240  
pyridines 98, 251  
pyridobenzodiazepinone 57  
pyridones 252  
pyrido[2,3]pyrimidines 251  
pyrimidine-5-carboxylic acid 82  
pyrimidines 57, 253  
pyrimidones 253  
pyrrole 56  
pyrrole amides, MS 567, 568  
pyrrole-imidazole tetramer 271  
pyrrole-library, MS 514  
pyrrolidine libraries, MS 575  
pyrrolidines 54, 252  
pyrrolinone-containing oligomers 274  
  
quadrupole analyzer 569  
quality control of polymer-bound libraries 494  
quantification by NMR 537  
quartz microbalance measurements 342  
quinazoline-2,4-dione 57  
quinoline 56

- radical reaction 60
- radiofrequency encoding system 7
- radius of similarity 426
- random library 26
- reaction apparatus 7
- reaction sequences 103
- reaction tables 38
- reagent management 557
- reagent mixture synthesis 10
- real-time monitoring of solid-phase reactions 488
- receptor 335
- receptor library 340
- recognition of MHC II-bound peptides 361
- recognition pattern of MHC molecules 358, 360, 374
- recognition by the T-cell repertoire 360
- reduced-amide bond peptides 282
- reducing agents 66
- reduction 65, 67, 92
  - of azide 66
  - of carbamate-linked amines 50
  - of ester-like compounds 49
  - of esters 51
  - of nitro groups 115
  - of Weinreb amides 51
- reductive alkylation 66
- reductive amination 52, 66, 89, 104
- redundancy check algorithm 413
- reflectometric interference spectroscopy 346
- Reisert-based amide linkage 47
- REM resin (regenerated Michael addition) 188
- resin capture 114
- resin-bound scandium catalyst 67
- resolutions 578
- resonance-enhanced multiphoton ionization 472
- resorcinarenes 338
- retro-inverso pseudopeptides 276
- rhodium(II)-catalyzed 1,3-dipolar cycloaddition 61
- ring closure metathesis 46
- ring formation 96
- ring-closing metathesis 62
- ring-containing peptides 272
- ring-opening metathesis 60
- Rink-'acetic acid'-amide 176
- Rink-acid resin 176
- Rink-amide resin 177
- robot platform ARCoSyn 552
- robot system 24
  
- saccharide antibiotics 402
- safety-catch anchor 207
- sample consumption 578
  - MS 571
- sample injector, MS 571
- sample logistics 550
- sample-oriented automation concepts 548
  
- sample-oriented multicomponent systems 551
- Sandmeyer-type reaction 48
- saponification 49, 115
- SASRIN super acid sensitive resin 172, 179
- scaffold-based libraries 4
- SCAL linker 210
- scale of solid-phase synthesis 12
- scanner/picker 453
- scanning mass spectrometer 473
- scheduling 549, 550
- Schöllkopf's isocyanide 158
- screening 441, 443
  - iterations 410
  - of a glycopeptide library 311
  - of a solid-phase glycopeptide library 312
  - of catalytic activity 468
  - of peptide libraries 358
  - scheme 23
  - strategies 444
  - technology 448
- selection experiment 321
- selection of building blocks 27
- selection techniques 413
- selenazolylypyrroles 246
- selenium compounds 62
- selenium-PS 215
- selenocysteine 328
- selenourea 247
- self-similarity plot 432
- semiconductor 547
- sensitivity of quartz microbalances 345
- sensor coating 336
- sensor signals 349
- sensors in automated workstations 547
- separation techniques with MS 505
- serial processing system 24
- serine/threonine 186
- Sharpless dihydroxylation 118
- sialylation 294
- sialyl-Le<sup>x</sup> tetrasaccharide 306
- Si-based traceless linkers 48
- siderophores 282
- Sieber-amide resin 177
- silicium-anchor 212
- similarity considerations 411
- similarity measure 416
- similarity principle 412
- sinefugin 142
- single bead analysis 575
- single bead IR spectroscopy 497
- single bead reaction monitoring 484
- single-bead assay 22
- software interface to robotic 554
- software-aided MS data evaluation 512
- solid support-bound EDC 107
- solid support-bound N-hydroxy-succinimide 108

- solid-phase, anchors 167
  - extraction 26, 109, 549, 553
  - glycopeptide template libraries 310
  - glycosylation of peptides 312
  - oligosaccharide library 304
  - organic reactions 35
  - bound reagents 106, 109
- soluble polymer approach 117
- sonication 84, 108
- spacer, definition 168
- specificity of human TAP 356
- spectra evaluation for MS 508
- spiramycin 381, 386
- spiroindoline 57
- split-and-combine method (IR analysis) 492
- split-pool synthesis 5
- split-search algorithm 419
- sputtering techniques 469
- squalene epoxide 132
- squaric acid as template 231
- stabilization assay 358
- stannanes 44
- statistical distribution of resin beads 495
- stereoselectivity of U-4CR 144
- steric fields 415
- Stille reaction 43, 231
- stimulation indices 374
- Streptomyces genetic tools 381
- structure-activity relationships 410, 454
- submonomer approach 258
- substance database 555
- N-substituted  $\beta$ -alanines 263
- N-substituted glycine monomers 258
- substitution reaction 67
- subtilisin 85
- succinaldehyde 153
- succinic acid 196
- sulfonamide-PS/DVB 211
- sulfonamidopeptides 264
- sulfonation of amines 86, 108
- sulfone 63
- sulfonic acid resin 107
- sulfonopeptides 264
- $\beta$ -sulfonopeptides 265
- sulfonylphenylhydrazine resin 211
- sulfoxide 63
- sulfoxide/ $Ti_2O$  procedure 302
- sulphonamide 39
- sulphonyl chlorides of resin 39
- sulphonylation 39
- sulphur-containing heterocycles 58
- superagonists 375
- superconducting magnet 562
- superconductors 463
- superposition algorithm 420
- support-bound hydroxybenztriazolyl group 108
- supports for oligosaccharide synthesis 296
- supramolecular recognition sites 336, 342
- surface-bound cyclopeptides 347
- surfactin antibiotics 401
- Suzuki coupling 43, 118
- Swissprot database 376
- synthesis
  - management 554
  - of catalysts 469
  - of mixtures 79
  - on silica wafer 9
  - on soluble polymers 117
  - in solution 13
- synthetic repertoire 36
- tagging reagents 21
- tagging techniques in oligosaccharide libraries 298
- tandem experiments 520
- tandem mass spectra 578
- tandem-domino reactions 132
- Tanimoto coefficient 414, 432
- TAP transporter 356
- target identification 443
- targeted libraries 410
- T-cell
  - antagonist 283
  - epitope 264, 266
  - mimotopes 362
  - proliferation 373
  - receptor contact positions 375
  - response 360
- tea-bag method 6
- templates 229
- testing of catalyst 466, 468, 471
- tetrabenz[a,c,g,i]fluorene 119
- tetracenomycin 381, 386, 389, 391
- tetracycline derivatives 137
- tetrafluoro-phthalic anhydride 113
- tetrahydro-carboline 57
- tetrahydrocyclopentapyridine 56
- tetrahydrofuran polyamides 274
- tetrahydroisoquinoline 56
- tetrahydroquinoxalin-2-one 57
- 1,2,3,4-tetrahydro- $\beta$ -carboline 480
- tetramethyl guanidine 114
- tetrangomycin 389
- tetragulol 389
- tetrazoles 131, 157
- thermistor 547
- thermocouple 547
- thermographic imaging device 471
- thiazanone 58
- thiazole 58, 98, 115, 272
- thiazole synthesis 158
- thiazolidine 53, 58
- thiazolidinone 53
- 4-thiazolidinone 58

- thiazolinones 115
- thiazolyl-oxazole 272
- thiazolylpyrroles 246
- thioacylating reagent 281
- thioamide pseudopeptides 281
- thioamides 247
- thioamino acid benzotriazole active ester 281
- thioesterification 401
- thiohydantoins 230
- thioisocyanates 39
- thiol addition 59
- thiols 199
- thionation 281
- thiophene 58
- thioureas 39, 247, 270
- Thorpe reaction 251
- T-lymphocytes 355
- tolerance to amino acid variations 370
- topological indices 415
- total correlation spectroscopy 534
- total ion currents 524
- traceless methods 47
- traceless-anchor 212
- transamination of amidine linker 50
- transducers 335
- transglutaminase 370
- translation diffusion 449
- translational invariance 367
- transmetallation 45
- transporter associated with antigen processing 357
- trialkylsilyl chloride-PS/DVB 215
- 1,3,5-triaryl-4,5-dihydropyrazoles 97
- triazine 57, 90
- triazene-PS/DVB 215
- triazole 57
- tributyltin isopropyl squarate 231
- trichloroacetimidate procedure 302
- 2,3,6-trichlorotriazine 90
- 3-(triethoxysilyl)-propylamine 199
- O-triflation 60
- O-triisopropylsilyl-serine 215
- triisoxazolines 273
- trimethylsilylation 522
- trimethylsilylenolthioether-PS/DVB 218
- triphenylphosphine-PS/DVB 199
- triphosgene-induced cyclization 270
- trityl-cation 171
- tritylchloride-PS/DVB 186
- Tröger's base 341
- tropinone 153
- tumor vaccines 361
- tweezer-like receptors 265, 340
- tylosin biosynthesis 382, 400
  
- Ugi 5-component condensation 94
- Ugi MCR reactions 125
- ultrahigh-resolution-measurement 574
- undecapeptide amide sublibraries* 362
- urdamycin 386
- ureapeptoids 271
- ureas and thioureas, synthesis of 84
- urethanes 40
- UV/visible video image of resin beads 491
  
- vacuum system, MS 563
- validation strategies 412
- vancomycin 339, 387
- vasopressin 325
- vinyllogous polypeptides 275
- virtual library 411
  
- Wang resin 172, 181, 195
  - IR 489
- willardiin 142
- Wittig olefination 107
- Wittig reaction 61
- Wittig-based cleavage 48
- Wittig-Horner reactions 108
- workstation 554
  
- xanthene library, MS 505
- xanthene tetracarboxamides 80
- xanthene tetracarboxylic acid 81
- xanthine 57
- xanthocillin 134
- xylocain 135
  
- zeolite synthesis 470
- zinc reagents 45