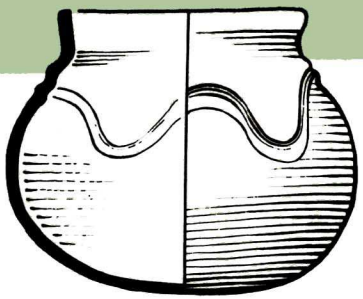
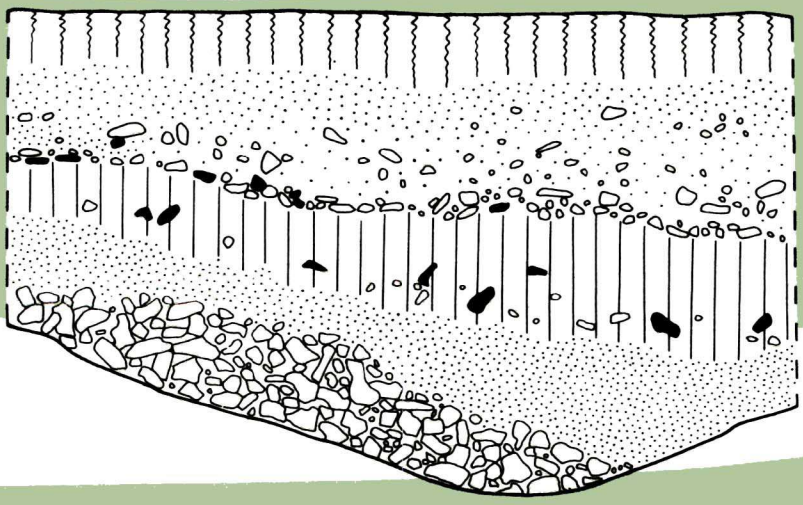
RSC PAPERBACKS

ARCHAEOLOGICAL CHEMISTRY

A. Mark Pollard
and Carl Heron





ARCHAEOLOGICAL CHEMISTRY

RSC Paperbacks

RSC Paperbacks are a series of inexpensive texts suitable for teachers and students and give a clear, readable introduction to selected topics in chemistry. They should also appeal to the general chemist. For further information on selected titles contact:

Sales and Promotion Department
The Royal Society of Chemistry
Thomas Graham House
The Science Park
Milton Road
Cambridge CB4 4WF, UK
Telephone : + 44 (0) 1223 420066
Fax: + 44 (0) 1223 423623

Titles Available

Water *by Felix Franks*
Analysis – What Analytical Chemists Do *by Julian Tyson*
Basic Principles of Colloid Science *by D. H. Everett*
Food – The Chemistry of Its Components (Third Edition)
by T. P. Coultate
The Chemistry of Polymers *by J. W. Nicholson*
Vitamin C – Its Chemistry and Biochemistry
by M. B. Davies, J. Austin, and D. A. Partridge
The Chemistry and Physics of Coatings
edited by A. R. Marrion
Ion Exchange: Theory and Practice, Second Edition
by C. E. Harland
Trace Element Medicine and Chelation Therapy
by D. M. Taylor and D. R. Williams
Archaeological Chemistry *by A. M. Pollard and C. Heron*

How to Obtain RSC Paperbacks

Existing titles may be obtained from the address below. Future titles may be obtained immediately on publication by placing a standing order for RSC Paperbacks. All orders should be addressed to:

The Royal Society of Chemistry
Turpin Distribution Services Limited
Blackhorse Road, Letchworth
Herts SG6 1HN, UK
Telephone: + 44 (0) 1462 672555
Fax: + 44 (0) 1462 480947

RSC Paperbacks

ARCHAEOLOGICAL CHEMISTRY

A. MARK POLLARD and CARL HERON
Department of Archaeological Sciences
University of Bradford
Bradford BD7 1DP, UK

THE ROYAL
SOCIETY OF
CHEMISTRY
Information
Services

ISBN 0-85404-523-6

A catalogue record for this book is available from the British Library

Published by The Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge CB4 4WF, UK

Typeset by Computape (Pickering) Ltd, Pickering, North Yorkshire, UK
Printed by Hartnolls Ltd, Bodmin, Cornwall, UK

Foreword

Archaeological science is a discipline which is growing rapidly in its scope and maturity. The application of techniques deriving from the natural sciences to archaeological research is, of course, not new and the present volume rightly emphasizes some of the earlier initiatives in the field of archaeological chemistry. But among the indicators of this increasing maturity is the quality of the science on the one hand, and a growing awareness of the problems of interpretation on the other.

This work by Pollard and Heron performs a valuable service in defining more carefully than hitherto and then richly exemplifying the field of archaeological chemistry, and it is fitting that it should be published by The Royal Society of Chemistry. Hitherto it is perhaps the discipline of physics which has come in for a good deal of the limelight since it holds what is almost a monopoly in two important fields of archaeological science: dating and prospecting. Some of the techniques of dating may indeed be included within the province of chemistry: amino acid racemization is dealt with very fully here, and obsidian hydration is mentioned. But many of the most useful dating methods are dependent upon radioactivity, a subject area often relinquished to physics. And again many of the techniques of field prospection – resistivity surveys, use of the proton magnetometer, *etc.* – fall within the same province.

As defined here it is, very reasonably, chemical *analysis* which lies at the heart of archaeological chemistry. Therefore, following the historical introduction, a wide-ranging survey of analytical techniques is undertaken which certainly offers the best overview currently available to the archaeologist. This is followed by what at first sight seems a comprehensive survey of examples and applications. On closer examination, however, it emerges that nearly every chapter is based in good measure upon the original research of one of the authors, supplemented by a

further overview. Thus Heron's work on organic residues, notably resins, is put to very good use. Pollard's researches in the characterization of ceramics, in glass studies, and in the use of brass and other metals likewise form the basis for much wider reviews of important fields.

One of the great strengths of the book in my view is its clear perception that archaeology and archaeological science, although interrelated and interdependent, are not at all the same thing. For the interpretation of the data of archaeological science is a complicated and difficult matter, and it involves two separate operations. The first, the elucidation of rather concrete, factual questions – 'how old is this sample?', 'where did it come from?', 'how has it been modified through burial in the ground?' – fall within archaeological science proper. But this does not make them easy questions: the difficulties in the interpretation of the data from lead isotope analysis obtained in order to determine the places of origin of metals used in the Aegean is a case in point. Another such quesiton would be the age and authenticity of the materials from the French site of Glozel – perhaps wisely not dealt with here, but still an unresolved problem in the field of archaeological science. The second operation is to achieve a clearer realization of the implications of such concrete conclusions for the understanding of human behaviour in the past. There the archaeologists and the archaeological scientists have yet to find more effective ways of working together: the thoughtful caution in the book makes an excellent start.

Archaeological chemistry is clearly not the same field as materials science, although the two overlap, nor as metallurgy: Chapter 6 draws a distinction between 'the chemical study of metals' and (I presume) the application of metallurgical techniques to elucidate the history of the working of individual artefacts. And this suggests to me that there lies here another sub-field which has not yet emerged, fully fledged, in its own right: the archaeology of technology, or archaeo-technology. For the time will come when we shall wish to focus more clearly upon processes and procedures of manufacture. The concept of the '*chaine operatoire*' now commonly applied to the study of lithic artefacts and to the sequence of operations undertaken in their manufacture has been applied already to ceramics and can certainly be applied also to other products of pyrotechnology. Texts are available from Mesopotamia which give recipes for various preparation procedures, for instance those of metals, and very odd reading they make for modern eyes. This suggests to me that the study of ancient technology, giving more attention to what the ancient smiths, potters, glassblowers, and dyers *thought* they were doing – this is to say, laying greater emphasis upon cognitive aspects – will soon be a possibility. But in order to do this a

prerequisite is to get the modern science right, and that is what Pollard and Heron are so systematically doing.

It is a great pleasure to introduce a work which, more clearly than ever before, delineates an important field within archaeological science and thus makes an important contribution to the discipline as a whole.

Colin Renfrew

Cambridge
January 1996

Preface

Archaeological chemistry is the application of chemical knowledge in order to help solve problems in archaeology. It is, however, much more than a straightforward application of existing techniques to new and interesting questions. Many of the chemical questions posed are unique to archaeology, although, somewhat surprisingly, many other disciplines share, from time to time, similar concerns. Our premise is that archaeological chemistry requires a thorough understanding of the background of both halves of the story, and often mastery of information from related disciplines such as biochemistry and geochemistry. This book is therefore aimed at two groups: chemists who are interested in new applications, and also archaeologists, particularly those on undergraduate and postgraduate courses encompassing aspects of scientific archaeology. It will also be of interest to geochemists, materials scientists, and forensic scientists. Perhaps the most important message it contains is the need to tackle the fundamental issues of chemical change in archaeological materials if scientific analysis is to make major contributions to the study of the past.

The continued expansion of scientific applications in the study of the past is one reason for writing this book. If the title of the book is not entirely new, then we believe our approach is – the adoption of a thematic structure. There are several reasons for this. Firstly, the majority of previous texts in archaeological science have tended to emphasize techniques at the expense of applications. Nowadays, there are so many techniques that such an approach would be unduly laborious, although some of the more important techniques are summarized in Chapter 2. Secondly, through an examination of particular themes, it is possible to document the successes and failures of past applications and assess the impact of scientific analysis on specific archaeological problems. It is also possible to see contemporary debates in terms of competing scientific

views and to suggest how these might be resolved from a knowledge of the underlying principles. In Chapter 1, we provide a short historical context, and return to general issues and future challenges in the final chapter.

We are aware of a Eurocentric bias in most of the chapters, although the issues raised will be applicable to other contexts. The majority of chapters feature the results of primary research carried out separately by the authors in collaboration with other colleagues. It is not possible to do justice to every application of chemistry to archaeology in a single volume; some of the criteria for inclusion are discussed above and in Chapter 1. As such, this book does not pretend to be a truly balanced review of archaeological chemistry and some readers may feel that some heinous crime of omission has been committed. We hope, however, that there is enough of interest here to justify the attempt to adopt a unified approach to archaeological chemistry as a worthwhile area of endeavour in its own right.

Mark Pollard and Carl Heron

Bradford
December 1995

Contents

Acknowledgements

The authors are grateful to many people over the past twenty years, all of whom have contributed in various ways to the production of this book. These include Dr G.A. Cox, Professor O.S. Heavens and Professor J.A.D. Mathew (Department of Physics, University of York); Professor Teddy Hall, Dr Robert Hedges, Dr Richard Gillespie, Dr John Gowlett, Helen Hatcher (Research Laboratory for Archaeology and the History of Art, Oxford University); Dr Robin Symonds (Museum of London); Dr Cath Mortimer (Ancient Monuments Laboratory, English Heritage); Professor Bob Gillard, Dr Pete Edwards and the troops (School of Chemistry, University of Wales, Cardiff); Professor Pete Williams, Dr Richard Thomas (School of Chemistry, University of Western Sydney); Dr Richard Evershed (Department of Chemistry, University of Bristol); Dr John Goad (Department of Biochemistry, University of Liverpool); Dr Gerry McDonnell, Dr Paul Budd, Dr Arnold Aspinall, and Geoff Gaunt (Department of Archaeological Sciences, University of Bradford).

We are also indebted to the following individuals: Miranda Schofield for drawing the figures; Rebecca Stacey for compiling the chemical structures; Professor Curt Beck (Vassar College, New York), Dr Vasily Kilikoglou (Demokritos, Athens), Professor Dr Klaus Ruthenberg (Coburg, Germany), and Dr Robert Tykot (Harvard University), for providing papers in advance of publication; Paula Mills, Dr Cathy Batt, Dr Randolph Haggerty, Dr Vikki Carolan, Dr Suzanne Young (Harvard University), and Dr Olwen Williams-Thorpe (Open University) for reading Chapters or parts thereof; and the many authors and publishers who allowed reproduction of figures used in this book. Finally, our thanks go to Rebecca, Yannick and our respective families for considerable forbearance during the production of this book.

Chapter 1

The Development of Archaeological Chemistry

INTRODUCTION

In its endeavour to understand human behaviour primarily through the material remains of past societies, archaeology has interacted more and more with the sciences of physics, chemistry, biology, and of the Earth. In truth, it is a test to conjure the name of any scientific discipline which has not at one time or another provided information of direct use for the archaeologist (Pollard, 1995). Indeed, many would consider archaeology itself, a discipline which involves the systematic collection, evaluation, and analysis of data and which aims to model, test, and theorize the nature of past human activity, to be a science. Furthermore, they might argue that it is possible to arrive at an objective understanding of past human behaviour, and in that sense archaeology is no different from other scientific disciplines, given the obvious differences in methodology. As Trigger (1988; 1) has reminded us, from a different perspective, archaeologists have a unique challenge:

> '*Because archaeologists study the past, they are unable to observe human behaviour directly. Unlike historians, they also lack access to verbally encoded records of the past. Instead they must attempt to infer human behaviour and beliefs from the surviving remains of what people made and used before they can begin, like other social scientists, to explain phenomena.*'

The claim that archaeology is a science is clearly not universally held. Many archaeologists suggest that the study of human behaviour in the past is restricted by science with its apparent rigidity of scientific method and dubious claims of certainty and must continue to reside with the humanities. Undoubtedly, archaeology is one of the few disciplines which bridges the gulf between the humanities and the sciences.

In our view, one of the fundamental enquiries in archaeology is the

relationship between residues, artefacts, buildings and monuments, and human behaviour. From the period of production, use or modification of materials (whether natural or synthetic) to the time when traces are recovered by archaeologists, the material output of humans is altered by a plethora of physical, chemical, and biological processes, including those operating after deposition into the archaeological record. A significant part of the evidence is lost, displaced, or altered significantly. Inferring the activities, motivations, ideas, and beliefs of our ancestors from such a fragmentary record is no small task. In fact, it is a considerable challenge. Although there are notable exceptions, archaeology in the last 150 years has been transformed from a pastime pre-occupied with the embellishment of the contemporary world (or at least a minuscule portion of it) with treasure recovered from 'lost civilizations' (still a view which predominates in some media, such as the cinema), to a discipline which relies on painstaking and systematic recovery of data followed by synthesis and interpretation. However, the development of archaeology has not been one uniform trajectory. There have been, and still are, numerous agendas which encompass the broad range of archaeological thought, and many uncertainties and disagreements concerning the direction of the discipline remain. Collectively, the sciences provide archaeology with numerous techniques and approaches to facilitate data analysis and interpretation, enhancing the opportunity to extract more information from the material record of past human activity. Specifically, chemistry has as much to offer as any other scientific discipline, if not more.

The sheer diversity of scientific analysis in archaeology renders a coherent and comprehensive summary intractable. In a recent review, Tite (1991) has packaged archaeological science rather neatly into the following areas:

- Physical and chemical dating methods which provide archaeology with absolute and relative chronologies.
- Artefact studies incorporating (i) provenance, (ii) technology, and (iii) use.
- Environmental approaches which provide information on past landscapes, climates, flora, and fauna as well as diet, nutrition, health, and pathology of people.
- Mathematical methods as tools for data treatment also encompassing the role of computers in handling, analysing, and modelling the vast sources of data.
- Remote sensing applications comprising a battery of non-destructive techniques for the location and characterization of buried features at the regional, microregional, and intra-site levels.

- Conservation science, involving the study of decay processes and the development of new methods of conservation.

Although in this volume we focus on the interaction between chemistry and archaeology or *archaeological chemistry*, it is relevant, in part, to most if not all of the areas proposed by Tite. For example, although many subsurface prospecting techniques rely on (geo)physical principles of measurement (such as localized variations in electrical resistance and small variations in Earth magnetism), geochemical prospection methods involving the determination of inorganic and biological markers of anthropogenic origin (*i.e.*, chemical species arising as a direct consequence of human action) also have a role to play. Throughout this book, archaeological chemistry is viewed not as a straightforward application of routine methods but as a challenging field of enquiry, which requires a deep knowledge of the underlying principles in order to make a significant contribution.

EARLY INVESTIGATIONS

It would not be possible to write a history of chemistry without acknowledging the contribution of individuals such as Martin Heinrich Klaproth (1743–1817), Humphry Davy (1778–1829), Jöns Jakob Berzelius (1779–1848), Michael Faraday (1791–1867), Marcelin Berthelot (1827–1907), and Friedrich August von Kekulé (1829–1896). Yet these eminent scientists also figure in the early history of the scientific analysis of antiquities. Perhaps the primary motivation for their work was curiosity, which resulted from their dedication to the study and identification of matter and the way in which it is altered by chemical reaction. In addition to his significant contributions to analytical and mineralogical chemistry, Martin Heinrich Klaproth determined the approximate composition of some Greek and Roman coins, a number of other metal objects, and a few pieces of Roman glass. Klaproth was a pioneer in *gravimetry* – the determination of an element through the measurement of the weight of an insoluble product of a definite chemical reaction involving that element. His first paper entitled '*Mémoire de numismatique docimastique*' was presented at the Royal Academy of Sciences and Belles-Lettres of Berlin on July 9th, 1795. The coins were either copper or copper alloy. In producing compositional data on ancient materials, Klaproth had first to devise workable quantitative schemes for the analysis of copper alloys and glass. His scheme for coins has been studied by Caley (1949; 242–43) and is summarized briefly below:

'*After the corrosion products had been removed from the surface of the metal to be*

analysed, a weighed sample was treated with "moderately concentrated" nitric acid and the reaction mixture was allowed to stand overnight ... the supernatant liquid was poured off and saved, and any undissolved metal or insoluble residue again treated with nitric acid ... If tin was present as shown by the continued presence of a residue insoluble in nitric acid, this was collected on filter paper ... (this) was simply dried in an oven and weighed ... a parallel control experiment was made with a known weight of pure tin. It was found from this that 100 parts of dried residue contained 71 parts metallic tin, in other words the gravimetric factor was 0.71.

The filtrate from the separation of the tin was tested for silver by the addition of a saturated solution of sodium chloride to one portion and the introduction of a weighed copper plate into another.

Lead was separated from the solutions ... by evaporation to a small volume. The separated lead sulfate was collected and either weighed as such or reduced to metallic lead in a crucible for direct weighing as metal.

(Copper) was determined as metal from the filtrate from the lead separation by placing in it a clean iron plate. The precipitated copper was then collected, dried, and weighed.'

In addition to Klaproth's pioneering work in quantitative analysis, he made a major contribution to mineralogical chemistry and discovered many elements in the process. His efforts did not go unrewarded, since he became Berlin's first Professor of Chemistry.

In 1815, Humphry Davy published a paper on the examination of ancient pigments collected at Rome and Pompeii. In addition to reviewing evidence for natural pigments, he was also able to identify a synthetic pigment later to be called *Egyptian Blue*, formed by fusing copper, silica, and naturally occurring natron (sodium carbonate). A report by H. Diamond, published in the journal *Archaeologia* in 1867 includes a section on a Roman pottery glaze studied by Michael Faraday in which the presence of lead in the sample provided the first indications on chemical grounds of the use of lead glaze in antiquity. In addition to his significant contributions to modern chemistry during the first half of the 19th Century, Berzelius became interested in the composition of ancient bronzes. Similarly, Kekulé carried out analysis of an ancient sample of wood tar that may have comprised, in part, compounds with aromatic or benzene rings, the structure of which he subsequently proposed in 1865.

In addition to the diverse activities of these well known scientists, efforts made by a number of other investigators during the 19th Century are worthy of note. Frequently they sought to examine ancient metal objects (Caley, 1949, 1951, 1967) with a view, initially, to understand their composition and the technology needed to produce the artefacts,

although other questions began to emerge. As these investigations continued, mostly in isolation from one another, prehistoric archaeology was making its first steps towards a systematic enquiry into the study and chronology of early materials. In 1819, Christian Thomsen assigned the artefacts in the Danish national collection into successive ages of stone, stone and copper, bronze, early iron, and later iron. This relative chronology was based on comparisons of material-type, decoration, and the context of recovery, and it marked a major development in the study of ancient materials which prevails in archaeology today (see Trigger, 1989; 73–79 for a more detailed consideration).

As early as the mid-19th Century, the Austrian scholar J.E. Wocel suggested that correlations in chemical composition could be used to provenance or identify the source of archaeological materials and even to provide relative dates of manufacture and use. During the 1840s, C.C.T.C. Göbel, a chemist at the University of Dorpat in Estonia, began a study of large numbers of copper alloy artefacts from the Baltic region, comparing those recovered from excavations with known artefacts of prehistoric, Greek, and Roman date. He concluded that the artefacts were probably Roman in origin. With the work of Göbel, scientific analysis progressed beyond the generation of analytical data on single specimens to, as Harbottle (1982; 14) has emphasized, '*establishing a group chemical property.*' The French mineralogist Damour proposed that the geographical source of stone axes could be located by considering the density and chemical composition of a number of rock types, including jade and obsidian found '*dans les monuments celtiques et chez les tribus sauvages*', as his papers of 1864 and 1866 were entitled (Caley, 1951; 66). Damour also exhorted archaeologists to work with specialists from other disciplines such as geology, zoology, and palaeontology (Harbottle, 1982; 14). Perhaps he was aware of the interdisciplinary research programmes comprising zoologists, geologists, and archaeologists then being carried out in Scandinavia on ancient shell mounds along the coast of Denmark (Klindt-Jensen, 1975; 71–73). Damour's primary interest was jade. Some 13 decades later, questions as to the possible source(s) of jade axes in prehistoric Europe remain.

The appearance of the first appendices of chemical analysis and references to them in the text of a major excavation report represents the earliest significant collaboration between archaeologists and chemists. Examples include the analysis of four Assyrian bronzes and a sample of glass in Austen Henry Layard's *'Discoveries in the Ruins of Nineveh and Babylon'* published in 1853 and Heinrich Schliemann's *'Mycenae'* first published in 1878 (so distinguished was the publication of the 1880 edition that William Gladstone, then British Prime Minister, wrote the

preface). The reports in the appendices of both these works were overseen by the metallurgist, John Percy, at the Royal School of Mines in London. Between 1861 and 1875, Percy wrote four major works on metallurgy which included significant sections on the early production and use of metals (Percy, 1861, 1864, 1870, 1875). These books remain important sources even today. Analysis of metal objects from Mycenae showed the extensive use of native gold and both copper and bronze, the latter used predominantly for weapons. Percy wrote in a letter to Schliemann dated August 10th, 1877 that '*Some of the results are, I think, both novel and important, in a metallurgical as well as archaeological point of view.*'

The effort made by Otto Helm, an apothecary from Gdansk, Poland, to source amber towards the end of the 19th Century constitutes one of the earliest systematic applications of the natural sciences to archaeology. It can be said that this enquiry was advanced with a specific archaeological problem in mind: determining the geographical source of over 2000 amber beads excavated by Schliemann at Mycenae. In the excavation monograph, Schliemann noted that '*It will, of course, for ever remain a secret to us whether this amber is derived from the coast of the Baltic or from Italy, where it is found in several places, but particularly on the east coast of Sicily.*' Helm based his approach on the succinic (butanedioic) acid content of Baltic amber (known since the mid-16th Century from the studies by Georg Bauer (1494–1555), who is better known to metallurgists as *Agricola*), but did not undertake a systematic study of fossil resins from other sources in Europe. His motivation lay, at least partly, in disproving the hypothesis of an Italian mineralogist, Capellini, who suggested that some of the earliest finds of amber in the south could have been fashioned from local fossil resins. A full account of the investigations made and the success claimed by Helm along with the eventual shortcomings has been compiled by Curt Beck (Beck, 1986) who in the 1960s published, with his co-workers, the results of some 500 analyses using infrared (IR) spectroscopy which demonstrated for the first time successful discrimination between Baltic and non-Baltic European fossil resins (Beck *et al.*, 1964, 1965). Unless severely weathered, it is usually possible to demonstrate that the vast majority of amber from prehistoric Europe derives from material originating in the Baltic coastal region.

The French chemist Marcelin Berthelot was active in chemical analysis in the late 19th Century, investigating some 150 artefacts from Egypt and the Near East. According to Caley (1967; 122), Berthelot may have been '*less interested in the exact composition of ancient materials than in obtaining results of immediate practical value to archaeologists.*' This was coupled with an interest in the corrosion of metals and the degradation of organic materials, which prompted a series of experimental studies based

on prolonged contact of metal objects with air and water. Although Berthelot published some 42 papers in this field, many of them remained unaltered, in title or content, from journal to journal. For this at least he perhaps deserves credit from contemporary academics for enterprise!

Towards the end of the 19th Century, as archaeological excavation became a more systematic undertaking, the results of chemical analysis became more common in reports and new suggestions began to appear. As early as 1892, A. Carnot suggested that fluorine uptake in long-buried bone might be used to provide an indication of the age of the bone (Caley, 1967; 122), although the feasibility of the method was not tested until the 1940s. The increasing numbers of antiquities brought about more emphasis on their restoration and conservation. The pioneer in this field was Friedrich Rathgen, who established a laboratory at the State Museum in Berlin and later published the first book ('*Die Konservierung von Alterthumsfunden*') dealing with practical procedures for the conservation of antiquities, including electrolytic removal of corrosion from ancient artefacts and the use of natural consolidants (such as pine resin and gelatin) in the conservation process. Developments in the examination of archaeological materials in Europe began to be applied to New World artefacts. In Sweden, Gustav Nordenskiöld submitted pottery sherds collected at Mesa Verde, Colorado for petrological examination (thin section analysis). The results appeared in his volume '*Cliff Dwellers of the Mesa Verde*' published in 1893. One of the first wet chemical investigations of ancient ceramics (Athenian pottery from the Boston Museum of Fine Arts) was carried out at Harvard and published in the *Journal of the American Chemical Society* in 1895 by T.W. Richards (Harbottle, 1982; 17).

THE GROWTH OF SCIENTIFIC ARCHAEOLOGY IN THE 20TH CENTURY

The 1920s and 1930s saw the addition of instrumental measurement techniques, such as optical emission spectroscopy (OES; see Chapter 2), to the repertoire of the analyst. The principal interest at the time was understanding the level of technology represented by finds of ancient metalwork, especially in terms of alloying, and systematic programmes of analysis were initiated in Britain and Germany leading to substantial analytical reports (*e.g.*, Otto and Witter, 1952). As a result of the rapid scientific and technological advances precipitated by the Second World War, the post-war years witnessed a wider range of scientific techniques being deployed in the study of the past. Eventual reconstruction as a result of war damage was preceded by a major expansion of archae-

ological excavation which produced very large quantities of artefacts. The development of radiocarbon dating by Willard Libby in 1949 paved the way for establishing absolute chronologies throughout the world. Although the impact was not immediate, radiocarbon dating eventually allowed sites to be dated in relation to one another and enabled cultural sequences to be established independent of cross-cultural comparisons (based on artefact typologies) with areas dated by historical methods (Renfrew, 1973).

Other materials such as faience beads and ceramics were incorporated into analytical programmes. Faience comprises a core of finely powdered quartz grains cemented by fusion with a small amount of alkali and lime. The core is coated with a glaze of soda-lime and coloured in the range blue to green with copper compounds. Faience was first produced in the Near East although Egyptian faience became very important between the 4th and 2nd Millennia BC. During the 2nd Millennium BC, faience was distributed across prehistoric Europe and occurs in England and Scotland. In 1956 Stone and Thomas reported on the use of OES to '*find some trace element, existent only in minute quantities, which might serve to distinguish between the quartz or sand and the alkalis used in the manufacture of faience and glassy faience in Egypt and in specimens found elsewhere in Europe.*' (1956; 68). This study represented a clear example of the use of chemical criteria to determine whether faience beads recovered from sites in Britain were imported from Egypt or the Eastern Mediterranean. For many years, it had generally been assumed that faience manufacture and other technological innovations originated in the east and diffused westwards. Although the initial results suggested that OES could not be used unequivocally, the data were subsequently re-evaluated statistically by Newton and Renfrew (1970) who suggested a local origin on the basis of the quantities of tin, aluminium, and magnesium present in the beads. This was augmented by re-analysis of most of the beads, using neutron activation analysis (NAA), by Aspinall and co-workers (1972). They confirmed that the tin content of British beads is significantly higher than that found in groups of beads from elsewhere and that a number of other trace elements showed promise. However, the belief in a local origin for the British beads is by no means universally held and only investigations of larger sample groups can hope to resolve the issue. This use of the highly sensitive technique of NAA was by no means the first. Sayre and Dodson (1957) applied NAA in their study of ancient Mediterranean ceramics and, in conjunction with gamma ray spectroscopy, NAA was used by Emeleus and Simpson (1960) to test the applicability of trace element analysis to locate the region of origin of Samian sherds which could not be placed stylistically (see Chapter 4).

In Britain, the term '*archaeometry*' was coined in the 1950s by Christopher Hawkes in Oxford to describe the increased emphasis on dating, quantification, and physico-chemical analysis of archaeological materials. A journal with the same name was launched in 1958 and textbooks by M.J. Aitken (1961) and M.S. Tite (1972) illustrated the full potential of emerging applications. In 1974, the first volume of another periodical dedicated to scientific work in archaeology (*Journal of Archaeological Science*) was published.

During the late 1950s and early 1960s, a number of individuals advocated strongly a new and refreshing approach to archaeology, although the roots of this impetus are evident in earlier writings. Progressive thinking in anthropology and the social sciences as well as the explicit use of models by geographers had largely left archaeology lagging behind. This transformation, which became known as the '*New Archaeology*', represented an explicit effort on the part of a number of archaeologists who emphasized optimistically the potential for explaining past human action rather than simply describing it. Such was the optimism of the early 1960s that it was felt that all human behaviour could be embodied within laws of cultural phenomena. Patterning in '*material culture*' and the '*archaeological record*' could be used to explore behavioural correlates regardless of time or place. Not surprisingly, the philosophy of science played a significant role in providing the terminology for statistical and quantitative approaches in archaeology (Trigger, 1989). The New Archaeology rejuvenated research into prehistoric trade and exchange. Invasion or diffusion of peoples was no longer viewed as the principal instigator of cultural change. Alternatively, internal processes within society were emphasized, although evidence for 'contact' arising from exchange of artefacts and natural materials (as well as the transmission of ideas) were seen as important factors from which scientific analysis might hope to evaluate change in economic and social systems. This increased interest in the distribution of materials initiated a golden era in archaeometry as a wide range of scientific techniques were deployed in the hope of chemically characterizing certain rock types, such as obsidian (see Chapter 3) and marble (Rybach and Nissen, 1964) as well as ceramics (*e.g.*, Catling *et al.*, 1963), metals (Junghans *et al.*, 1960), glass (Sayre and Smith, 1961) and natural organic materials, such as amber (Beck *et al.*, 1964). These characterization studies were aimed at '*the documentation of culture contact on the basis of hard evidence, rather than on supposed similarities of form*' (Renfrew, 1979; 17). The substantial data sets generated by these techniques could now also be subjected to statistical treatment using computers. Characterization studies remain an important research area in archaeological science,

uitilizing a range of chemical properties incorporating trace element composition, biomarker composition, mineralogy, and scientific dating, including isotopic measurements. In a review of chemical characterization, Harbottle (1982; 15) reminded practitioners that:

> *'... with a very few exceptions, you cannot unequivocally source anything. What you can do is characterize the object, or better, groups of similar objects found in a site or archaeological zone by mineralogical, thermoluminescent, density, hardness, chemical, and other tests, and also characterize the equivalent source materials, if they are available, and look for similarities to generate attributions. A careful job of chemical characterization, plus a little numerical taxonomy and some auxiliary archaeological and/or stylistic information, will often do something almost as useful: it will produce groupings of artefacts that make archaeological sense. This, rather than absolute proof of origin, will often necessarily be the goal.'*

The geographical source of the materials under investigation included quarries, mines or clay deposits, and sites of production where materials are modified or fabricated. If the material remains unaltered during preparation or modification, for example, when flakes of obsidian are removed from a large core of the rock, then the bulk composition of the artefact is unaltered from the source material, although subtle changes may occur (such as in the case of a *hydration layer* – see Chapter 3). However, in the case of synthetic materials such as ceramics, metals, and glass, production may bring about significant changes in the composition of the finished artefact with respect to the composition of the raw materials. The whole question of provenance becomes a complex issue (Tite, 1991; 143–144 for important comments, and Cherry and Knapp, 1991).

Until recently, archaeology has generally paid more attention to the analysis of inorganic artefacts – natural stone, metal, glass, ceramic material, and so on – reflecting an interest in the most obviously durable artefacts in the archaeological record. In recent years, increasing attention has been directed at biological materials; natural products such as waxes and resins, accidental survivals, such as food residues, and, above all, human remains, including bone, protein, lipids, and, most recently of all, DNA. Some of the methodology for this work has been imported not only from chemistry, biochemistry, and molecular biology, but also from organic geochemistry, which has grown from a discipline with a principal interest in elucidating the chemical origins of oil and coal into one which studies the short-term alteration and long-term survival of a very wide range of biomolecules (*e.g.*, Engel and Macko, 1993). Another related discipline in this quest for ancient biomolecular information is molecular palaeontology. These disciplines are widely recognized as having much to offer each other, particularly in the recovery of genetic information from animals and plants. In terms of specific archaeological interest, the

ability to extract DNA has considerable significance. Hitherto, extraction of nucleic acids from bone some 25 000 years old has been claimed. Preserved soft tissue and seed remains also yield extractable DNA. Specific DNA sequences can be targeted, amplified using the *polymerase chain reaction* (PCR), and compared with sequences in other individuals and modern specimens. However, ancient DNA is severely damaged and fragmented. Contamination of aged samples and extracts with modern DNA is a serious problem and whilst the study of DNA in archaeological samples will constitute a major area of future activity in the discipline (Thomas, 1993), current research will continue to focus on the authentication of samples of ancient DNA (Richards *et al.*, 1995).

Preservation of a wider range of biomolecules has been demonstrated in a number of archaeological contexts. In particular, proteins preserved in human bone have been subject to immunological investigation (*e.g.*, Smith and Wilson, 1990; Cattaneo *et al.*, 1992). The survival of protein residues on stone tool surfaces (*e.g.*, Loy, 1983) hints at the possibility of characterizing artefact use and identifying utilization of specific resources and dietary items, although the specificity of the approaches used remains contentious (*e.g.*, Eisele *et al.*, 1995; Tuross and Dillehay, 1995). Certain biomolecules, such as lipids, are more durable than nucleic acids and, although some chemical alteration is to be expected, specific identifications can be made on aged samples (*e.g.*, Heron *et al.*, 1994; Evershed, 1993).

The voluminous literature on bone chemical investigations generated during the last two decades represents one of the significant growth areas of archaeological chemistry (*e.g.*, Price, 1989; Grupe and Lambert, 1993; Sandford, 1993). Quantitative analysis of certain trace elements (such as strontium, barium, zinc, and lead) incorporated into bone mineral and, less frequently, teeth and hair has been used to assess diet, nutrition, health status, and pathology. Similarly, dietary inferences have been made through measurement of light stable isotope ratios of carbon and nitrogen in bone collagen and the carbon isotope composition of bone apatite. The realization that maize and other grasses have a high carbon-13 content has been used to characterize long term consumption of maize in the Americas, specifically its introduction and expansion in eastern North America (van der Merwe and Vogel, 1978). However, the recognition of significant compositional and mineralogical alteration during long-term burial (often labelled *diagenesis*) has brought about a re-evaluation of bone chemical investigations; the onus of proof is now on the analyst to demonstrate that the data are not geochemical artefacts reflecting more on the complex interaction between bone and the burial environment than any dietary or other signal which may have accumu-

lated during life. Undoubtedly, trace element compositions are highly susceptible to a wide range of post-depositional alterations, including exchange between ions in the soil solution and bone mineral (Sandford, 1993).

The limitations of the types of sample analysed in archaeological chemistry can be considerable. Typically samples are usually far from ideal from the analytical point of view – small, fragmentary, and, particularly in the case of biological samples, often considerably degraded and potentially contaminated not only during deposition but also, significantly, once the sample is recovered (post-excavation) due to storage media, handling, and airborne particles. The ubiquitous problem of degradation and contamination makes archaeological chemistry a challenging field, and not one which can be regarded as just another routine application of analytical chemistry. Recently, parallels with forensic science have been drawn (Hunter *et al.*, 1996).

CURRENT STATUS AND SCOPE OF ARCHAEOLOGICAL CHEMISTRY

Although archaeological science is recognized as a fundamental component of the inquiry into past human behaviour and development, the demand for relevant data has been echoed on many occasions. Scientific analysis should be much more than a descriptive exercise which simply documents the date, morphology, or composition of ancient materials. As DeAtley and Bishop (1991; 371) have pointed out (see also Trigger, 1988) no analytical technique has '*built-in interpretative value for archaeological investigations; the links between physical properties of objects and human behaviour producing the variations in physical states of artefacts must always be evaluated.*' The demand for meaningful scientific data also needs to viewed against the changing approaches to the study of the past as the discipline of archaeology has evolved. The historical relationship between scientific approaches and techniques and prevailing theoretical views regarding past human behaviour has been reviewed by Trigger (1988; 1) who states that '*archaeologists have asked different questions at different periods. Some of these questions have encouraged close relations with the biological and physical sciences, while other equally important ones have discouraged them.*' During the 1960s, archaeology embraced the sciences, not only the techniques but the terminology of scientific methods for explaining human behaviour. During the last decade and a half, general attitudes in society towards science have been shifting towards a more critical stance. Against this backdrop, the contribution of scientific analysis to the study of the past has come under increasing scrutiny. Certain approaches, commonly

labelled *post-processual*, to archaeological thinking from the mid-1980s onwards have stressed relativism and subjectivism in interpretation as well as the ideological and symbolic roles of material culture to a greater extent than before, and, evidently, the limited contribution of scientific data to this inquiry has been emphasized (see, for example, Hodder, 1984; Thomas, 1991).

Although the majority of archaeologists acknowledge the contribution of scientific dating and analytical techniques to increasing the information potential of the past, the central concern prevails that scientific studies often proceed in a context devoid of a specific archaeological problem (Yoffee and Sherratt, 1993; 4–5). However, it would be misleading to suggest that chronological, compositional, or locational data generated by scientific techniques have no role to play in providing foundations for interpretations of past human behaviour. As Cherry and Knapp (1991; 92) have remarked, scientific analysis could '*help arbitrate amongst competing cultural hypotheses*' although they could find little evidence for such an approach. Perhaps success in archaeological science is difficult to measure. For some, it may be the implementation of a robust or elegant scientific methodology; for others it may be the degree of integration within archaeological problems. Ideally, it should display characteristics of both. The promotion of genuine interdisciplinary studies rather than multidisciplinary investigation lies at the heart of archaeological endeavour (DeAtley and Bishop, 1991).

THE STRUCTURE OF THIS VOLUME

It is hoped that the foregoing provides a short historical context to this volume. Indeed we return to this discussion briefly in the final chapter. In the intervening chapters, we have selected a number of themes which exemplify past and current research in archaeological chemistry. The themes presented in this volume (representing only a small component of what is called archaeological science) span many diverse areas of chemistry. There are many reasons for adopting a thematic approach. Firstly, the majority of previous texts in archaeological science have tended to emphasize techniques at the expense of applications which have produced relevant archaeological information. Nowadays, there are so many techniques that such an approach would be unduly laborious. Techniques of analysis are summarized as briefly as possible in Chapter 2. The remaining chapters range in scope, but each includes a discussion of some of the underlying science. Chapter 3 reviews and updates the classic characterization studies undertaken with the aim of locating the source of the volcanic glass obsidian. Chapter 4 reviews the

structural chemistry of clays, and illustrates the power of chemical studies of ceramics with an example from Roman Britain and Gaul. Chapter 5 discusses the structure and chemistry of archaeological glass, together with a review of some work on the atmospheric corrosion of Medieval window glass. Two chapters focus on the chemical study of metals. The first is Chapter 6, which considers the 'traditional' chemical analysis of brass objects, including a discussion of the use of this knowledge for the authentication of brass scientific instruments. Chapter 7 focuses on the chemistry of resins and aims to consider the future role of analytical organic chemistry applied to seemingly unavailing amorphous deposits surviving on artefact surfaces. Chapter 8 continues the theme of organic chemistry in archaeology, with a consideration of the racemization of amino acids in bones and teeth, with an example drawn from the controversial question of the dating of the arrival of the earliest humans in the New World. Finally, Chapter 9 returns to metals, and tackles the currently topical field of lead isotope geochemistry, in particular with a critical review of its role in locating the source of metals in the Mediterranean Bronze Age.

Perhaps more importantly we should enumerate some of the areas not covered. The chemistry of ancient bone is neglected, simply because of the plethora of good quality reviews which have been published in the last few years. Those who wish to seek out recent work should consult the following useful edited volumes and conference proceedings by Price (1989), Grupe and Lambert (1993), and Sandford (1993). Without doubt, the field of DNA studies in archaeology is one of the fastest growing areas in archaeological science – so fast, in fact, that in our view any review attempted here would be out of date before it appeared in print. A collection of papers on the extraction and study of ancient DNA can be found in the recent book by Hummel and Herrmann (1993) although the advances have been so rapid that perusal of the appropriate scientific journals is essential. Every reader will probably feel that some heinous crime of omission has been committed. This book does not pretend to be a truly balanced review of archaeological chemistry – the size of a single volume precludes any attempt to do that. What we have attempted to do is present a range of studies which have been important archaeologically, are interesting from a chemical standpoint, and have interested the authors at one time or another.

FURTHER READING

A comprehensive history of scientific analysis applied to the study of past people and materials is lacking. The papers by Caley (1949; 1951; 1967)

remain useful for summaries of the early applications of chemistry to archaeology and the paper by Trigger (1988) is essential reading. The contributions of Berzelius, Davy, Faraday, and others to the development of chemistry are summarized by Hudson (1992). The recent text by Renfrew and Bahn (1991) serves as a very useful introduction, covering the scope and aims of modern archaeology, including many scientific applications. For a more detailed consideration of the development of archaeology, see Trigger (1989). Current debates on the theory of archaeology can be found in a series of essays in Yoffee and Sherratt (1993) and in Hodder *et al.* (1995).

A collection of recent scientific studies, largely relating to museum objects, including dating, authenticity, metalwork, ceramics, and glass, can be found in the edited volume by Bowman (1991). Many conference proceedings (especially those entitled *Archaeological Chemistry*, produced by the American Chemical Society – Beck, 1974; Carter, 1978; Lambert, 1984; Allen, 1989; Orna, in press) contain a very wide range of chemical studies in archaeology. Of the several books covering the chemical aspects of archaeological science, most are too superficial to be useful. The book by Goffer (1980) is probably the best, giving a very broad introduction to archaeological chemistry, covering basic analytical chemistry, the materials used in antiquity, and the decay and restoration of archaeological materials. Scientific dating methods, such as radiocarbon dating, electron spin resonance, thermoluminescence dating, and so on are not covered in this volume, even though chemistry is intimately involved in many of the methods. There is a long history of relevant texts: from Zeuner's *Dating the Past* (four editions between 1945 and 1957) to Aitken's excellent *Science-based Dating in Archaeology* (1990).

REFERENCES

Aitken, M.J. (1961). *Physics and Archaeology*. Interscience Publishers, New York. (2nd Edition, 1974, published by Oxford University Press, Oxford).

Aitken, M.J. (1990). *Science-based Dating in Archaeology*. Longman, London.

Allen, R.O. (ed.) (1989). *Archaeological Chemistry IV*. Advances in Chemistry Series 220, American Chemical Society, Washington, D.C.

Aspinall, A., Warren, S.E., Crummett, J.G. and Newton, R.G. (1972). Neutron activation analysis of faience beads. *Archaeometry* **14** 41–53.

Beck, C.W. (ed.) (1974). *Archaeological Chemistry*. Advances in Chemistry Series 138, American Chemical Society, Washington, D.C.

Beck, C.W. (1986). Spectroscopic studies of amber. *Applied Spectroscopy Reviews* **22** 57–110.

Beck, C.W., Wilbur, E. and Meret, S. (1964). Infrared spectra and the origins of amber. *Nature* **201** 256–257.

Beck, C.W., Wilbur, E., Meret, S., Kossove, D. and Kermani, K. (1965). The infrared spectra of amber and the identification of Baltic amber. *Archaeometry* **8** 96–109.

Bowman, S. (ed.) (1991). *Science and the Past.* British Museum Press, London.

Caley, E.R. (1949). Klaproth as a pioneer in the chemical investigation of antiquities. *Journal of Chemical Education* **26** 242–247; 268.

Caley, E.R. (1951). Early history and literature of archaeological chemistry. *Journal of Chemical Education* **28** 64–66.

Caley, E.R. (1967). The early history of chemistry in the service of archaeology. *Journal of Chemical Education* **44** 120–123.

Carter, G.F. (ed.) (1978). *Archaeological Chemistry II.* Advances in Chemistry Series 171, American Chemical Society, Washington, D.C.

Catling, H.W., Blin-Stoyle, A.E. and Richards, E.E. (1963). Correlations between composition and provenance of Mycenaean and Minoan pottery. *Annual of the British School at Athens* **58** 94–115.

Cattaneo, C., Gelsthorpe, K., Phillips, P. and Sokol, R.J. (1992). Reliable identification of human albumin in ancient bone using ELISA and monoclonal antibodies. *American Journal of Physical Anthropology* **87** 365–72.

Cherry, J.F. and Knapp, A.B. (1991). Quantitative provenance studies and Bronze Age trade in the Mediterranean: Some preliminary reflections. In *Bronze Age Trade in the Mediterranean*, ed. Gale, N.H., Studies in Mediterranean Archaeology XC, Paul Åströms Förlag, Jonsered, pp. 92–119.

DeAtley, S.P. and Bishop, R.L. (1991). Toward an integrated interface for archaeology and archeometry. In *The Ceramic Legacy of Anna O. Shepard*, eds. Bishop, R.L. and Lange, F.W., University Press of Colorado, Boulder, Colorado, pp. 358–380.

Eisele, J.A., Fowler, D.D., Haynes, G. and Lewis, R.A. (1995). Survival and detection of blood residues on stone tools. *Antiquity* **69** 36–46.

Emeleus, V.M. and Simpson, G. (1960). Neutron activation analysis of ancient Roman potsherds. *Nature* **185** 196.

Engel, M.H. and Macko, S.A. (eds.) (1993). *Organic Geochemistry: Principles and applications.* Plenum Press, New York.

Evershed, R.P. (1993). Biomolecular archaeology and lipids. *World Archaeology* **25** 74–93.

Goffer, Z. (1980). *Archaeological Chemistry.* Wiley-Interscience, New York.

Grupe, G. and Lambert, J.B. (eds.) (1993). *Prehistoric Human Bone: Archaeology at the Molecular Level.* Springer-Verlag, Berlin.

Harbottle, G. (1982). Chemical characterization in archaeology. In *Contexts for Prehistoric Exchange,* eds. Ericson, J.E. and Earle, T.K., Academic Press, New York, pp. 13–51.

Heron, C., Nemcek, N., Bonfield, K.M., Dixon, J. and Ottaway, B.S. (1994). The chemistry of Neolithic beeswax. *Naturwissenschaften* **81** 266–269.

Hodder, I. (1984). Archaeology in 1984. *Antiquity* **58** 25–32.

Hodder, I., Shanks, M., Alexandri, A., Buchli, V., Carman, J., Last, J. and Lucas, G. (eds.) (1995). *Interpreting Archaeology: Finding Meaning in the Past.* Routledge, London.

Hudson, J. (1992). *The History of Chemistry.* Macmillan, London.

Hummel, S. and Herrmann, B. (eds.) (1993). *Ancient DNA: Recovery and Analysis of Genetic Material from Paleontological, Archaeological, Museum, Medical and Forensic Specimens.* Springer-Verlag, New York.

Hunter, J.R., Roberts, C.A. and Martin, A. (eds.) (1996). *Studies in Crime: An introduction to Forensic Archaeology.* Seaby/Batsford, London.

Junghans, S., Sangmeister, E. and Schroder, M. (1960). *Kupfer und Bronze in der fruhen Metallzeit Europas.* Mann Verlag, Berlin.

Klindt-Jensen, O. (1975). *A History of Scandinavian Archaeology.* Thames and Hudson, London.

Lambert, J.B. (ed.) (1984). *Archaeological Chemistry III.* Advances in Chemistry Series 205, American Chemical Society, Washington, D.C.

Loy, T.H. (1983). Prehistoric blood residues: Detection on stone tool surfaces and identification of species of interest. *Science* **220** 1269–71.

Newton, R.G. and Renfrew, C. (1970). British faience beads reconsidered. *Antiquity* **44** 199–206.

Orna, M.V. (ed.) (in press). *Archaeological Chemistry V.* Advances in Chemistry Series, American Chemical Society, Washington, D.C.

Otto, H. and Witter, W. (1952). *Handbuch der Altesten Vorgeschichtlichen Metallurgie in Mitteleuropa.* Leipzig.

Percy, J. (1861). *Metallurgy. Volume I: Fuel; Fire-clays; Copper; Zinc; Brass.* Murray, London.

Percy, J. (1864). *Metallurgy. Volume II: Iron; Steel.* Murray, London.

Percy, J. (1870) *Metallurgy. Volume III: Lead.* Murray, London.

Percy, J. (1875). *Metallurgy. Volume IV: Silver; Gold.* Murray, London.

Pollard, A.M. (1995). Why teach Heisenberg to archaeologists? *Antiquity* **69** 242–247.

Price, T.D. (ed.) (1989). *The Chemistry of Prehistoric Bone.* Cambridge University Press, Cambridge.

Renfrew, C. (1973). *Before Civilization: The Radiocarbon Revolution and Prehistoric Europe.* Jonathan Cape, London.

Renfrew, C. (1979). *Problems in European Prehistory.* Edinburgh University Press, Edinburgh.

Renfrew, C. and Bahn, P. (1991). *Archaeology: Theories, Methods, and Practice.* Thames and Hudson, London.

Richards, M.B., Sykes, B.C. and Hedges, R.E.M. (1995). Authenticating DNA extracted from ancient skeletal remains. *Journal of Archaeological Science* **22** 291–299.

Rybach, L. and Nissen, H.U. (1964). Neutron activation of Mn and Na traces in marbles worked by the Ancient Greeks. In *Proceedings of Radiochemical Methods of Analysis.* International Atomic Energy Agency, Vienna, pp. 105–117.

Sandford, M.K. (ed.) (1993). *Investigations of Ancient Human Tissue: Chemical analyses in anthropology.* Food and Nutrition in History and Anthropology, Volume 10, Gordon and Breach, Langhorne, Pennsylvania.

Sayre, E.V. and Dodson, R.W. (1957). Neutron activation study of Mediterranean potsherds. *American Journal of Archaeology* **61** 35–41.

Sayre, E.V. and Smith, R.V. (1961). Compositional categories of ancient glass. *Science* **133** 1824–1826.

Smith, P.R. and Wilson, M.T. (1990). Detection of haemoglobin in human skeletal remains by ELISA. *Journal of Archaeological Science* **17** 255–268.

Stone, J.F.S. and Thomas, L.C. (1956). The use and distribution of faience in the Ancient East and Prehistoric Europe. *Proceedings of the Prehistoric Society* **22** 37–84.

Thomas, J. (1991). Science or anti-science? *Archaeological Review from Cambridge* **10** 27–36.

Thomas, K.D. (1993). Molecular biology and archaeology: A prospectus for inter-disciplinary research. *World Archaeology* **25** 1–17.

Tite, M.S. (1972). *Methods of Physical Examination in Archaeology.* Seminar Press, London.

Tite, M.S. (1991). Archaeological science – past achievements and future prospects. *Archaeometry* **31** 139–151.

Trigger, B.G. (1988). Archaeology's relations with the physical and biological sciences: a historical review. In *Proceedings of the 26th International Archaeometry Symposium,* eds. Farquhar, R.M., Hancock, R.G.V. and Pavlish, L.A., University of Toronto, Toronto, pp. 1–9.

Trigger, B.G. (1989). *A History of Archaeological Thought.* Cambridge University Press, Cambridge.

Tuross, N. and Dillehay, T.D. (1995). The mechanism of organic

preservation at Monte Verde, Chile, and one use of biomolecules in archaeological interpretation. *Journal of Field Archaeology* **22** 97–110.

van der Merwe, N.J. and Vogel, J.C. (1978). ^{13}C content of human collagen as a measure of prehistoric diet in Woodland North America. *Nature* **276** 815–816.

Yoffee, N. and Sherratt, A. (eds.) (1993). *Archaeological Theory: Who Sets the Agenda?* Cambridge University Press, Cambridge.

Zeuner, F.E. (1957). *Dating the Past.* Methuen, London (4th edn.).

Chapter 2

Analytical Techniques Applied to Archaeology

INTRODUCTION

The purpose of this chapter is to give a brief but largely non-mathematical introduction to some of the many analytical techniques used in modern archaeological chemistry. The vast majority of applications use standard instrumentation, which is well described elsewhere in the analytical chemistry literature (*e.g.*, Ewing, 1985, Sibilia, 1988, Skoog and Leary, 1992, Christian, 1994), but often some accommodation has to be made for the unique nature of many archaeological samples, which has been discussed in the previous chapter. The analytical techniques are described in groups, which are based on factors such as the region of the electromagnetic spectrum they employ, or the type of information obtained. Inevitably, most attention has been given to describing those techniques which have appeared most often in the archaeological literature. This might give a slightly 'dated' feel to some sections, although, where possible, recent developments are discussed and compared with established techniques.

In order to give some background to several of the techniques, the next section gives a description of the theory of the relationship between the electronic structure of the atom, electronic transitions, and the electromagnetic spectrum, since this is the basis of analytical spectroscopy. For those unfamiliar with the simple theory of the structure of the atom, a summary is given in Appendix 1 (p. 347). The third section of this chapter describes the techniques which use the visible (or near visible) region of the spectrum – optical emission spectroscopy (OES), atomic absorption spectroscopy (AAS) and inductively coupled plasma emission spectroscopy (ICP). The next section focuses on the techniques which use *X*-rays [*X*-ray fluorescence (XRF), scanning electron micro-

scopy (SEM), and proton-induced *X*-ray emission (PIXE)]. We then change emphasis, and look at those techniques which use mass spectrometry (including a discussion of the newer 'hyphenated techniques', which couple a mass spectrometer to another device to give more sensitive detection, such as ICP-MS). A section follows considering the range of chromatographic techniques which have become increasingly important with the growing interest in organic and biological remains in archaeology. Finally, there is a very brief review of some other analytical techniques which, although mainstream and important elsewhere, have received less attention in archaeological chemistry. This includes infrared spectroscopy (IR, which is rapidly growing in importance), electron spin resonance (ESR), and nuclear magnetic resonance (NMR), and some of the thermal methods of analysis, including differential thermal analysis (DTA) and differential scanning calorimetry (DSC).

ELECTRONIC TRANSITIONS, THE ELECTROMAGNETIC SPECTRUM, AND ANALYTICAL SPECTROSCOPY

Appendix 1 reviews the traditional Bohr model of the atom, and the associated electronic structure. It shows briefly how the Periodic Table of the elements can be built up from the known rules for filling up the various electron energy levels, without going into detail about the shapes of orbitals, or the subtleties of the electronic structure of the transition metals. Some of this is discussed in more detail in Chapter 5, in connection with the colouring effect of transition metal ions in glasses.

The Bohr model of the atom shows that electrons can only occupy orbitals whose energy are 'fixed' (quantized), and that each atom is characterized by a particular set of energy levels. These will be different in detail between atoms of different elements, because each element, by definition, has a different nuclear charge and orbital electron configuration. Interaction between the nuclear charge and the orbital electron energies ensures that each element has a unique pattern of orbital energy levels, despite the fact that the same notation is used for different elements. Electron transitions between allowed energy levels, as described in Appendix 1 (pp. 347–354), either absorb or emit a fixed amount of energy, corresponding to the energy difference between the energy levels of the atom. This fixed amount of energy, symbolized as ΔE, where ΔE is $E_2 - E_1$, is the energy difference between the two levels having energies E_1 and E_2. This quantum of energy manifests itself as electromagnetic radiation, whose energy and wavelength are related by the equation:

$$E = h\upsilon = hc/\lambda$$

where υ is the *frequency*, λ is the *wavelength*, c is the speed of light, and h is Planck's constant (see Appendix 1).

A key prediction from the Bohr model is therefore that atoms can only emit or absorb electromagnetic radiation in fixed units or quanta, corresponding to the energy differences between electron orbitals. If the exact energy and the number of electrons occupying each orbital is known, then it should be possible to predict exactly the allowed transition energies between orbitals, and therefore, via particle–wave duality, the wavelengths corresponding to these electronic transitions for any particular atom. Furthermore, calculations suggest that some at least of these allowed transitions should give rise to radiation which falls into the visible region of the spectrum. Early spectroscopic experiments on the simplest of atoms (hydrogen) showed that the emission spectrum obtained by passing an electrical discharge through hydrogen gas and dispersing the light emitted with a prism did indeed yield a relatively small number of discrete lines in the visible region, whose wavelengths were calculable using a formula (the *Rydberg equation*) derived from the models outlined above. Figure 2.1 shows a schematic diagram of the electronic energy levels in the hydrogen atom, the allowed transitions, and the resulting spectral lines. More complicated (multi-electron) atoms give more emission lines, and corrections need to be applied to the Rydberg formula, but the theory is generally found to hold good. Thus it was realized that each element in the Periodic Table has a unique line emission or absorption spectrum in the visible region of the spectrum, directly reflecting the unique orbital electronic structure of the atom. Spectral analysis of elements became an established 'fingerprinting' technique in a range of applications, although the heavier elements can display a bewildering total number of emission lines – up to 4 500 for iron, but not all in the visible region.

An important observation was that the emission lines are not confined to the narrow visible region of the electromagnetic spectrum. Instrumental detection showed that discrete lines are also present in the infrared and ultraviolet wavelengths, and eventually also in the *X*-ray region. It became clear that the wavelength of the line simply corresponded to the energy difference between the particular electron energy levels involved, and therefore that different regions of the electromagnetic spectrum could give information from different energy 'depths' within the orbital structure of the atom. The correspondence between these regions and the wavelength is shown in the lower half of Figure A1.2 in Appendix 1 (p. 353), ranging from microwave frequencies which relate to the relatively low energy rotations of molecules around their bonds up to *X*-rays, which arise from high energy transitions between the deepest

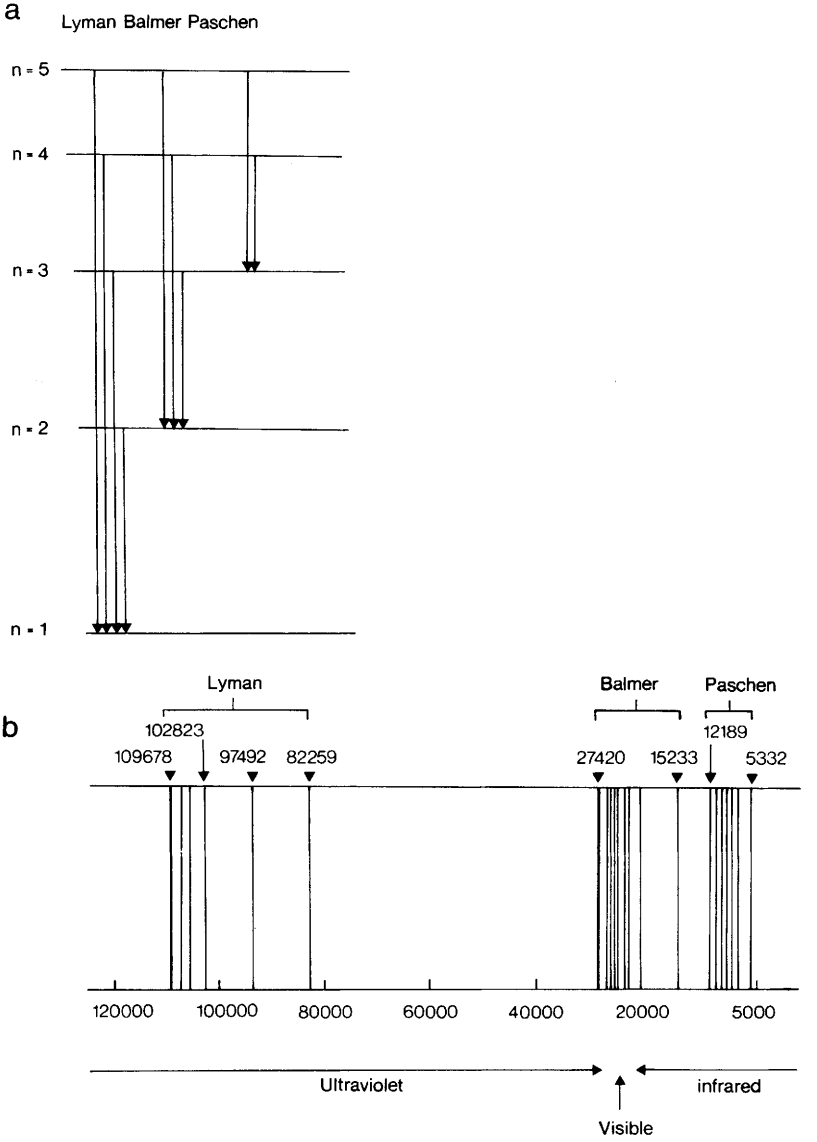


Figure 2.1 *Electronic orbitals and the resulting emission spectrum in the hydrogen atom. (a) Bohr orbitals of the hydrogen atom and the resulting spectral series, (b) emission spectrum of atomic hydrogen. The spectrum in* (b) *is calibrated in terms of wavenumber* ($\bar{\nu}$)*, which is reciprocal wavelength. The Balmer series, which consists of those transitions terminating on the second orbital, give rise to emission lines in the visible region of the spectrum*

electronic orbitals in large atoms. Spectroscopy has therefore provided a wide range of tools capable of giving a great deal of information about the electronic structure of the atom.

Most pertinent to the current discussion is the fact that each atom has a unique pattern of emission lines around the visible region of the spectrum, thereby allowing chemical identity to be established by simply comparing patterns. Moreover, via a relationship known as *Beer's Law* (or the *Beer–Lambert Law*), a quantitative link is established between the number of atoms involved and the intensity of the emission lines. Thus, in the hydrogen experiment described above, if the number of hydrogen atoms in the electrical discharge chamber was doubled, the intensity (strength) of the observed emission lines would also theoretically be doubled. Under normal circumstances, there is therefore a simple linear relationship between the intensity of an emission line and the number of atoms which give rise to that emission. This is the basis of quantitative spectroscopy.

A number of generalizations are necessary before further discussion. Atoms which lose or gain one or more electrons in the course of chemical bonding become electrically charged and are called *ions*: an ion will not have exactly the same emission spectrum as its parent atom, since the electronic energy levels will adjust to allow for the change in the number of electrons orbiting the nucleus. It is therefore important in spectroscopy to distinguish between atomic and ionic spectra. By extension, when atoms or ions combine to form *molecules*, the molecular energy levels are different from both the ionic and atomic levels, and molecular emission spectra are consequently different again. In general, molecular spectra do not show the sharp lines characteristic of atomic or ionic spectra, tending more towards broad bands. There is also a very simple relationship between the wavelengths which atoms, ions, or molecules will emit, when excited by some external stimulus such as an electrical discharge or an increase in temperature, and the wavelengths which will be absorbed if the atom, ion, or molecule is exposed to electromagnetic radiation. They all absorb at exactly the same wavelengths as they emit. In the hydrogen example, therefore, if instead of passing an electric discharge to cause emission, the gas was illuminated by light of all wavelengths, the light which is transmitted through the gas would be the same as that illuminating the gas, but without the wavelengths seen in the emission spectrum – thus the normal 'rainbow' of dispersed white light would appear to have a number of dark lines in it, at wavelengths identical to those in the emission spectrum. Beer's law still applies to the absorption case – *i.e.*, the strength of the absorption is directly proportional to the number of atoms, ions, or molecules involved in the absorption process. These relatively simple concepts underpin much of quantitative analy-

tical chemistry, and often allow the identity and relative amounts of a wide range of elements present in a sample to be determined from a single sample of a few milligrams.

TECHNIQUES BASED ON OPTICAL WAVELENGTHS

Optical Emission Spectroscopy

Although now largely outmoded as an analytical technique in its original form, it is worthwhile to start our consideration of the three intimately related techniques of optical emission, atomic absorption, and inductively-coupled plasma emission with a brief description of *optical emission spectroscopy* (OES; Britton and Richards, 1969), since the principles involved in OES transplant simply to the more recent techniques. An outline diagram of the components of a large quartz spectrograph is shown in Figure 2.2. The sample in powder form is placed in a hollow graphite cup (in the position marked 'sample' in the figure) and a graphite electrode is brought close to the cup, allowing an electric spark to be

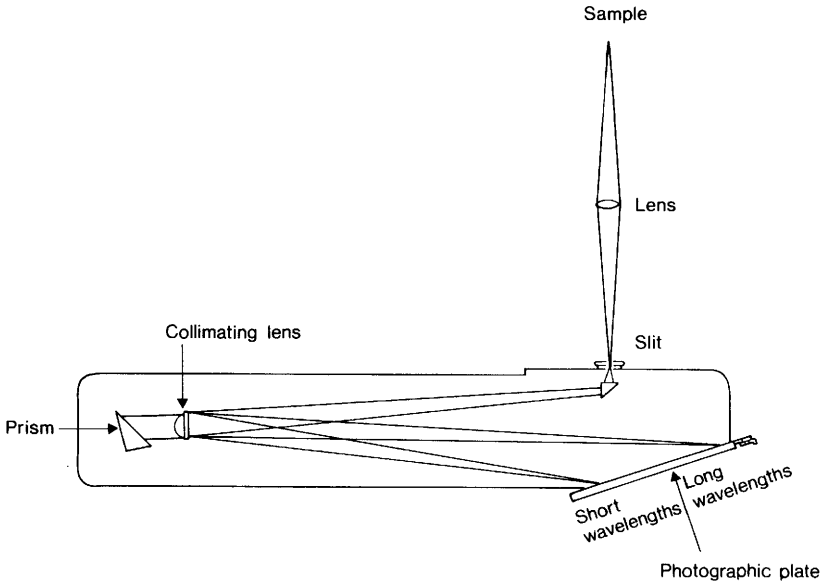


Figure 2.2 *Schematic drawing of an optical emission spectrograph. Light from the sample is focused onto the input slit of the spectrograph and is then dispersed via a prism (or diffraction grating) and recorded on a photographic plate*
(Adapted from Britton and Richards, 1969; figure 108, by permission of Thames and Hudson Ltd.)

struck between the two, thus volatilizing the sample and causing it to emit light. This light is focused through a series of lenses onto a large quartz prism (or in later models a diffraction grating) which disperses the light into its constituent wavelengths. This light is then focused onto the recording medium, which, in the original spectrometers, was simply a photographic plate, which is subsequently developed in the conventional manner. Careful measurement from the plate enables the wavelength of each line recorded to be determined, allowing the elements present to be identified, and optical densitometry techniques allow the intensity of particular emission lines to be measured, thus allowing quantification of the amount of each element present. In practice, a single line from each element known to be of interest is selected for quantification, taking into account the relative theoretical intensities of all the lines from the element, and the number of spectral overlaps known to occur in the relevant sample. By calibrating the spectrometer against known concentrations of each element of interest, it is therefore possible to produce quantitative data. One particular advantage of this technique is that elements which may be present but might not have been expected can be detected by careful observation of unusual lines on the photographic plate. In the later terminology of instrumental analytical techniques, OES is described as having a *simultaneous* detection system, since all elements are recorded at the same time. The plates can also be stored, thus enabling reference at a later date should it be necessary. The disadvantages are many (which is why it is no longer used!). The exact positioning of the graphite electrodes affects the reproducibility of the measurements, and the photographic procedure is notoriously difficult to standardize. Both of these problems can be allowed for to some extent by 'spiking' the sample with a known concentration of an element not present in the sample (often lithium), and referencing all intensity measurements to that element. Using this approach, it is possible to measure virtually any element present in a sample of 10 mg, in concentrations between 0.001% and 10%. In practice, because of the complexity of the emission spectra of the elements, the maximum number of elements measurable from a single exposure is about 20, with a coefficient of variation (defined as the standard deviation divided by the mean, multiplied by 100) of between 5 and 25% for major and minor elements (Jones, 1986; 24 and 883–888).

Atomic Absorption Spectrometry

In archaeological chemistry, OES was the standard method of analysis for pottery, obsidian, faience, and metals from the 1950s through to

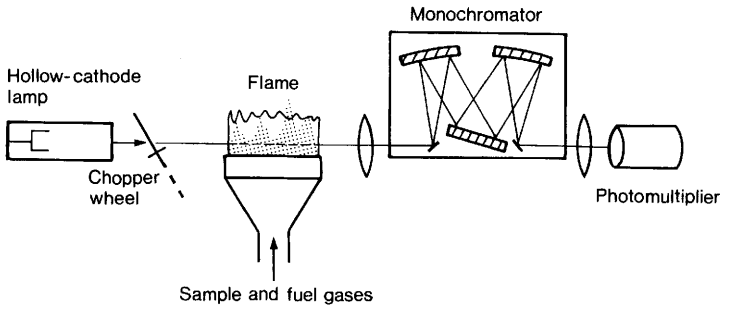


Figure 2.3 *Schematic diagram of an atomic absorption spectrometer. Light of wavelengths characteristic of the sample of interest is generated in the lamp and passes through the flame containing the atomized sample. The light is quantitatively absorbed in the flame, and the wavelengths are separated in the monochromator. The intensity of the transmitted light is measured in a photomultiplier* (Adapted from Ewing, 1985; Figure 5–9, with permission of The McGraw-Hill Companies)

about 1980, and there is a great deal of literature referring to its use. It was gradually replaced as an analytical technique by atomic absorption spectrometry (AAS). This differs principally in that it requires the sample to be in a liquid (normally aqueous, *i.e.* dissolved in water) form, thus making the sample preparation stage somewhat more complicated, and that it can only measure one element at a time (a so-called *sequential* operation). A schematic diagram of a simple atomic absorption spectrometer is shown in Figure 2.3. Since in its basic mode of operation AAS is an absorption (as opposed to emission) technique, light which is characteristic of the element to be determined has to be generated. This is produced in a special lamp called a hollow-cathode lamp, which consists of a glass (or quartz) envelope filled with a noble gas at low pressure. Within the envelope are two electrodes, one a wire and the other a cup made from (or lined with) the element of interest. On application of a few hundred volts between the electrodes, the cathodic cup begins to glow with the characteristic radiation of the element, since the voltage is enough to cause electrons to travel from the wire to the cup; on impact, the atoms of the cup are excited and emit their characteristic radiation. In the spectrometer, the light from this lamp is then guided towards the photomultiplier detector through the long axis of a long thin gas flame produced at a specially-shaped burner. The liquid sample is aspirated into the flame along with the combustible gases. The fuel and oxidant gases are pre-mixed in a chamber designed to ensure complete homogenization of the gases, and the flow of these gases causes the liquid

sample to be sucked into the chamber via a capillary tube – no additional pumping is necessary. The design of the mixing chamber also ensures that the sample liquid is effectively atomized into a fine aerosol before it enters the flame (see Ewing, 1985; Chapter 5).

On entering the flame, the sample is almost immediately converted into an atomic vapour. This is then in the ideal form to absorb characteristic radiation from the light source shining through the flame, and the amount of radiation absorbed at a particular wavelength is directly proportional to the concentration of that particular element in the flame. It is sometimes necessary for certain elements to increase the temperature of the flame in order to ensure complete atomization. This is done by changing the mixture of gases used, from the normal compressed air/acetylene mixture, which burns at about 2 200 °C, to nitrous oxide/acetylene, giving a temperature of up to 3 000 °C. The latter gases require a slightly different design of burner, so the change-over cannot be made at will.

In the simplest version of AAS, the light from the hollow-cathode lamp shines directly through the flame onto a slit and monochromator device, which disperses the light into its constituent wavelengths and selects a particular wavelength for transmission onto the detector, which is a photomultiplier tube capable of quantitatively converting the light intensity it receives into an electric current which is easily measured. In its most basic form of operation, the light intensity passing through the flame is first of all measured without any sample being aspirated, and then with the sample introduced. The difference between the two is the absorption due to the atoms in the sample, and this absorption is calibrated by measuring the absorption due to the aspiration of a solution containing a known concentration of that element. The concentration of the element in the sample solution can then be calculated, which in turn allows the calculation of the concentration in the original solid sample, since the weight of sample dissolved to make the solution is carefully recorded. The instrumentation for AAS has now advanced considerably from this rather laborious procedure. *Double-beam instruments*, where the light from the hollow-cathode lamp is split by a rotating mirror into a sample and a reference beam, allow the simultaneous comparison of absorbed and unabsorbed intensities. Signal noise reduction techniques ('signal chopping', using the 'chopper' wheel which also serves as the mirror) are now also used, so that 'flicker' in the light produced by the flame (a source of noise in the signal measured in the above techniques) can be eliminated. Further background noise reduction techniques are also routinely employed, such as the use of polarized light. Details of these can be found in Ewing (1985; 109–123). It has

even proved possible to remove altogether the biggest source of signal noise and irreproducibility – the flame itself – by using what is known as an *electrothermal* or *graphite furnace* technique. In this the sample solution is injected directly into a small electrically-heated chamber which replaces the burner, where the temperature is rapidly increased in a programmed manner so as to produce fast and reproducible atomization. Light from the hollow-cathode lamp passes directly through the furnace chamber.

Once a flame atomic absorption spectrometer has been set up for a particular element and all of the variables (*e.g.*, gas flow, burner position relative to the light beam, *etc.*) have been optimized to produce the best conditions, it provides a rapid and effective means of analysis, capable of analysing many tens of samples per hour (especially if the instrument is equipped with an autosampler). Analytical precision is generally good, with a coefficient of variation of between 1 and 5% for a wide range of elements. Detection limits in solution are typically between 1 and 100 parts per million (*ppm*, roughly equivalent to mg of element per litre of solution), depending on the element, analytical conditions, and the particular absorption line selected (Hughes *et al.*, 1976). One particular advantage of solution techniques of analysis in general is that the standards used for calibration can be accurately made up from commercially-available standard solutions, allowing accurate matching between the composition of the standards and the unknowns. This allows the analyst to undertake good *quality assurance procedures* by having separate calibration standards and quality control samples (usually samples of known composition, to validate the analysis).

Disadvantages include problems of reproducibility between 'runs' on the same element, due to inconsistencies in the setting of instrumental parameters such as gas flow, although this can usually be minimized (or at least quantified) by extensive use of quality control standards. Problems of calibration drift can also be encountered during a 'run' as a result of small changes in operating conditions. These are monitored by regular referral to one or more of the standards being used. In computer-controlled instruments this drift can be automatically compensated for by adjusting the calibration curve being used, but this is an unsatisfactory procedure if the drift is significant. Other limitations of AAS derive from the sequential nature of the operation. Although multi-element lamps exist (*e.g.* a combined lamp for calcium and magnesium), they are not always as spectrally clean as single element lamps, and the basic procedure remains the sequential use of a single lamp for each element to be determined in the sample. This results in a lengthy analytical procedure if several elements are required from the same sample, requiring a separate aspiration for each sample for each

element. One further result of this is that, unlike OES, it is very unlikely that any unexpected elements would be noticed in a sample, since the procedure is designed only to give information about the elements sought. As with most analytical techniques, a great deal of skill and experience is needed to produce reliable data, taking into account problems such as chemical and spectral interferences in the flame, *etc.* One further major consideration is that the sample must be in liquid form, and for certain types of material (such as glass, pottery, and certain metals) this can sometimes pose considerable problems, since some materials (*e.g.*, pottery) require aggressive dissolution conditions involving hydrofluoric acid. In the analysis of metals, some elements such as tin are difficult to keep in solution once the sample has been dissolved. Standard texts are available, covering a wide range of dissolution techniques for various samples, such as van Loon (1980, 1985).

Despite the obvious inference to be made from the name, atomic absorption spectrometers can in some circumstances also be used as emission spectrometers. Instrumentally this is simply achieved by switching off the lamp, and letting the light emitted by the excited atoms in the flame pass through the monochromator onto the detector. Using Beer's Law, the intensity of the light emitted is proportional to the concentration of that element in the flame. It might be wondered why the machine is not permanently used in this so-called *emission mode*, since it is simpler than using it in the absorption mode, but the answer lies in the behaviour of different elements in flames of different temperatures. For atomic absorption, the optimum state of the elements in the sample is completely atomized, but with all the electrons in their lowest possible energy levels – the *ground state.* In this condition, the maximum possible number of atoms are available to absorb the correct frequency from the hollow-cathode lamp, thus giving maximum analytical sensitivity. If conditions in the flame are slightly more energetic, the flame itself will excite a proportion of the atoms into a higher energy state, preventing them from absorbing the incoming light, and actually cause them to emit light of the correct atomic frequency as they return to their ground state. If there are a significant number of excited atoms in the flame, then it is better to switch off the lamp and do the analysis in the emission mode. If conditions in the flame are even more energetic, then some of the atoms will lose (or gain) electrons and become ions, thus changing the emission spectrum completely, and making them 'invisible' at the atomic frequencies used for the analysis. In this case, the sensitivity of the method deteriorates badly, although there are techniques available for suppressing this ionization. Some elements, such as the alkali metals sodium and potassium, are relatively easily excited and ionized, and these elements

are often best measured using the emission mode. Other elements, such as aluminium and titanium, require a great deal of energy to get them even into the atomic state, and can only be measured using the absorption mode with the hottest flame available. There is comprehensive literature available on dissolution techniques for AAS, optimal measuring conditions for each element, and for each type of sample.

Inductively Coupled Plasma Emission Spectrometry

As noted above, there may sometimes be a need to raise the temperature of the flame well above that easily obtained in a conventional gas burner, in order to ensure that some of the more refractory compounds are fully dissociated. In atomic absorption, the upper limit is about 3–4 000 °C, but temperatures in excess of 8–10 000 °C are achievable using a device called an *inductively coupled plasma atomic emission spectrometer* (ICP-AES). This is essentially the same as an atomic absorption spectrometer operated in the emission mode, with the exception that the gas burner is replaced by a *plasma torch*, capable of supporting the combustion of argon at these very high temperatures. Clearly, at such high temperatures, any normal material would rapidly melt and fail, and the ingenuity of the plasma torch is the fact that it is made out of relatively ordinary materials (silica tubing, with a melting point of around 1 700 °C), but designed in such a way as to support the plasma. A typical plasma torch consists of three concentric silica tubes, with copper coils wound around the outside at the top (Figure 2.4). The argon gas which is to form the plasma is injected vertically through the central tube, but a larger volume enters between the two outer envelopes at a tangential angle, and spirals up between the outer casing, acting as a coolant. When ignited, the plasma passing through the centre of the torch glows white-hot, but is lifted away from the silica tubing by the toroidal flow of cooling gas. The heating is maintained by a high-power radio frequency (RF) alternating current which is passed through the copper coils surrounding the torch, which causes the charged particles in the plasma to flow through the gas in a circular path by induction. The friction caused by this rapid motion through the gas holds the temperature at several thousand degrees, and ensures that the plasma is sufficiently ionized to respond to the RF heating. One complication is that the ignition of argon is not easy – it cannot be done using a match! Instead, an external spark is passed through the argon which causes some of the gas to ionize, and enables it to respond to the RF heating. Frictional heating soon raises the temperature sufficiently for the ionization to become self-sustaining.

As with atomic absorption, a liquid sample is injected into the flame as

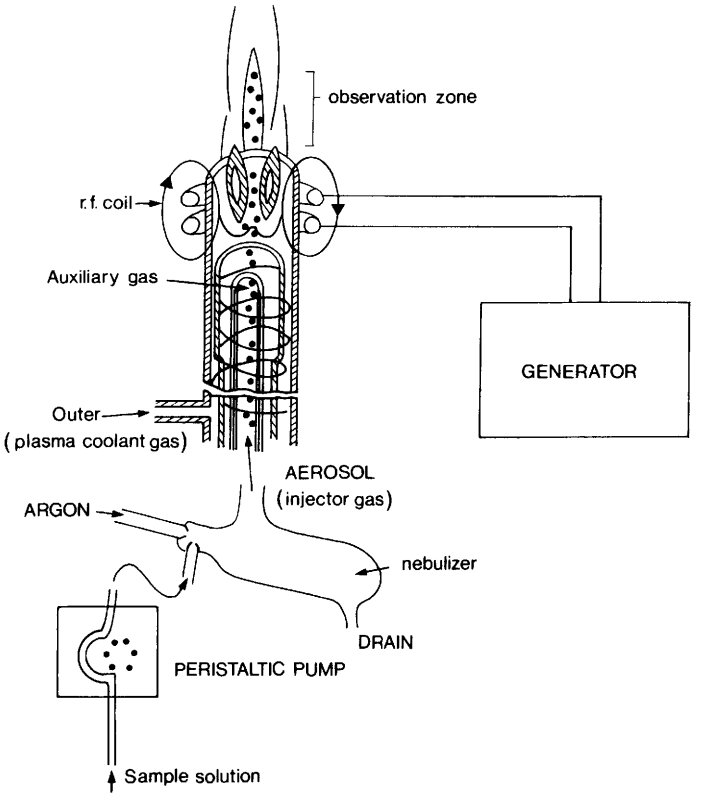


Figure 2.4 *A plasma torch for an inductively coupled plasma emission spectrometer.* (Reprinted from *Applied Geochemistry*, vol. 1, J.N. Walsh and R.A. Howie, 'Recent developments in analytical methods: uses of inductively coupled plasma source spectrometry in applied geology and geochemistry', pp. 161–171, Figure 1, copyright 1986, by kind permission from Elsevier Science Ltd., The Boulevard, Langford Lane, Kidlington, OX5 1GB, UK)

a solution carried by the fuel argon, although the design of the ICP usually involves a pump to suck up the sample and inject it into the argon stream. At the high temperatures of the torch, all compounds are completely dissociated and normally in an excited state, so that they strongly emit characteristic lines, which are subsequently dispersed using a diffraction grating and slit system similar to that used in atomic absorption. Detection is by photomultiplier tube, as with AAS, but there are some further innovations incorporated into modern instrumentation. Some machines which are used for quality control in industrial production, or for the routine analysis of a particular material, are set up

with a bank of photomultipliers, each one positioned (at a fixed angle relative to the diffraction grating) for a particular wavelength corresponding to a particular element. In this manner simultaneous measurements of up to twenty elements in a single sample can be achieved, making the equipment in many ways analogous to the old optical emission spectrometers with their simultaneous detection capability. More usually for research purposes the ICP is equipped with a single computer-controlled detector, which can automatically perform sequential analysis of several tens of elements whilst the sample is being aspirated into the machine. The detector is moved to the correct position for a particular element, measures the emission intensity for a pre-set time, and then moves on to the next element according to a pre-set programme in the computer. Although strictly sequential in operation, the combination of software-controlled analysis programmes and automated sampling makes the machine quasi-simultaneous. It would be fair to say that, over the past five years or so, ICP has largely replaced AAS as the industrial standard for the multi-element analysis of solution samples (Thompson and Walsh, 1989). Obviously, for archaeology, there is still the problem of putting the sample into solution, and keeping it there. Further developments have enhanced its performance even more, such as the use of a mass spectrometer as the detection system, and the recent development of a laser ablation system, which allows a solid sample to be precisely vaporized by a high power laser, which clearly has a great deal to offer archaeological chemists.

A major benefit of the ICP torch which has been exploited over the past few years is that it can be connected up to a mass spectrometer (see pp. 61 and 66) to give the very powerful technique of *inductively coupled plasma mass spectrometry* (ICP-MS). The temperature of the torch (up to 10 000 °C) is sufficient to ionize approximately 50% of the atoms in the sample, making the plasma an ideal source of ions which, given the appropriate interface to deal with the high temperatures involved, can be extracted from the plasma and injected directly into a mass spectrometer. When used in this mode, individual charged ions are separated according to their mass and charge, and effectively can be counted individually. Not only is this more sensitive than the measurement of emission intensities (thus lowering even more the detection limits for certain elements), but it also allows the analyst to monitor the concentration of individual isotopes of a particular element, which is useful for monitoring artificial isotopic tracers in biological systems. It also enables the direct measurement of the ratios of specific isotopes for a particular element, such as lead. Some of the advantages and uses of this approach are discussed in Chapter 9.

There have been several comparative reviews of the performances of AAS and ICP, a recent one of which is by Slavin (1992). Figure 2.5 shows a comparison of the analytical sensitivity measured as the lowest amount detectable in parts per billion in the sample solution (1 in 10^{-9}, or $\mu g\ l^{-1}$) for four techniques – ICP-MS, furnace AAS (AAS using electrothermal atomization), flame AAS, and ICP-AES. Although in general ICP-AES usually has a limit of detection comparable with flame AAS, it can be seen that an element by element comparison often reveals significant differences. In general, the refractory elements (Al, Ti, *etc.*) are better detected by ICP-AES, whereas the heavy (non-refractory) elements are better by flame AAS. For example, Ba by ICP-AES has a detection limit of around 0.5 ppb (parts per billion), compared with about 8 ppb by flame AAS, whereas Pb is better detected by AAS (10 ppb compared with about 50 ppb). The use of an electrothermal furnace with AAS generally gives an order of magnitude improvement in limits of detection, and the use of a mass spectrometer on an ICP gives a further order of magnitude improvement. Equally, if not more, important are the analytical precisions of the various techniques – a measure of the repeatability of the experimental method. In routine applications, the precisions are usually quoted as around 0.5% for flame AAS, 1.5% for ICP-AES, 3% for furnace AAS and 2–3% for ICP-MS, although individual laboratories may be able to improve substantially on these figures for particular applications (Slavin, 1992). Recently, a review has been published concerning the intercomparability of silicate analyses carried out by ICP-AES and AAS for the purposes of comparing data on archaeological ceramics (Hatcher *et al.*, 1995). This concluded that the results from both methods were sufficiently close that common data banks could be established, providing adequate care had been taken to include certificated standards within each run to monitor performance on the more difficult elements.

The various forms of instrumentation for ICP are substantially more expensive to buy than AAS instruments, and are generally regarded as requiring more skilled operation. Nevertheless, over the past few years it has gradually become the technique of choice for large-scale analytical programmes where the samples can be brought into solution without too much difficulty. Recently, attention has also been paid to two methods of analysing samples in the solid state, namely *slurry nebulization* and *laser ablation* (Jarvis *et al.*, 1992; Chapter 10). A slurry is a uniform suspension of small particles, which can be introduced into the ICP in just the same way as a solution, providing it is not too viscous, and the suspended particle load is not too great (less than 2 mg ml^{-1} for ICP-MS). Problems occur in ensuring uniform grinding and particle size distribu-

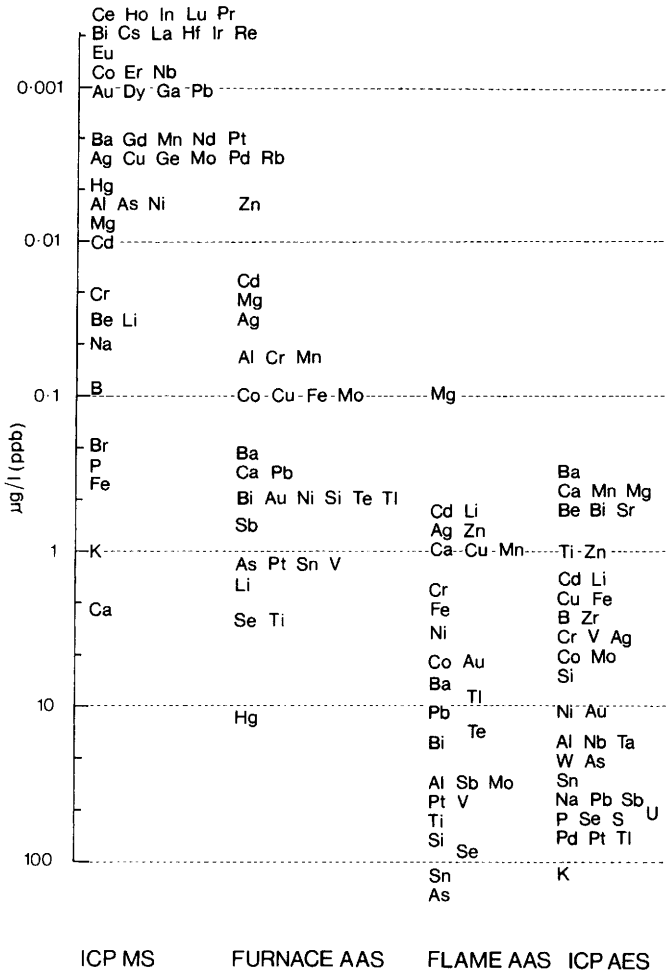


Figure 2.5 *Comparison of detection limits for four related analytical techniques ICP-MS, furnace AAS, flame AAS, and ICP-AES*
(Slavin, 1992; Figure 1, by permission of *Spectroscopy Europe* and the author)

tion, but pilot studies with geological material using slurry nebulization appear to suggest that there is no compromise of analytical performance (Jarvis *et al.*, 1992; 287). Laser ablation is potentially even more attractive, since it offers the possibility of spatially-resolved microanalysis of solid samples, in a manner similar to electron microscopy (described below), but with greater sensitivity and the potential for isotopic analysis.

In this technique, a high energy pulsed laser beam is directed onto a solid sample, with a beam diameter of less than 25 μm. The pulse vaporizes about 1 μg of material, to leave a crater 50 μm deep. The vaporized sample is swept into the ICP torch via a carrier gas. Current equipment allows solid samples, approximately flat, of 35 mm diameter to be accommodated in the laser ablation unit. The analytical sensitivity in terms of minimum detection limits is poorer for laser ablation than for solution analysis, and the detection limit for a particular element (or isotope) depends on the solid matrix which is being analysed. Nevertheless, the potential advantages of such a system far outweigh these problems, and laser ablation ICP-MS (LA-ICP-MS) is likely to become a powerful technique in archaeology, as in all other materials sciences.

TECHNIQUES USING *X*-RAYS

The *X*-ray region of the electromagnetic spectrum consists of wavelengths between 10^{-9} and 10^{-15} m. *X*-ray spectroscopists still use the non-SI standard unit of the Ångstrom (Å) which is defined as 10^{-10} m (so that 10 Å = 1 nm). In these units, *X*-rays used in analytical work range between 1 and 10 Å in wavelength. Electronic energy levels deep within the orbital electron structure of the heavier elements have such high energy differences that transitions occurring between these levels give rise to quanta whose energy (or wavelength) cause them to lie within the *X*-ray region of the spectrum. The situation is exactly analogous to that described above for optical spectra, except for the higher energy of the transitions involved. It is, however, complicated by the fact that *X*-ray spectroscopists, for historical reasons, insist on labelling the energy levels differently to those used by chemists! They designate the innermost orbitals K, L, M, N, *etc.*, corresponding to the principal energy levels n = 1, 2, 3, 4 described in Appendix 1, and use a different notation for defining the sub-shells of each energy level. Thus, 2*s* is designated L_I, but the 2*p*-orbital is split into two levels, labelled L_{II} and L_{III}. The correspondence between these different systems is set out in most books on *X*-ray spectroscopy, such as Jenkins (1988; Chapter 2), and is incidentally illustrated non-mathematically in Figures 2.7 and 2.8.

If an electron is removed (by a means discussed below) from one of the inner energy levels of one of the heavier elements – in practice, generally those above sodium in the Periodic Table – a *vacancy* or a *hole* is created in the electronic structure. Two competing processes can occur to rectify this unstable arrangement, one resulting in the emission of an electron (the *Auger process*), and one in the emission of an *X*-ray. In the *X*-ray process, internal rearrangement of the outer electrons results in one of

the electrons from a higher energy level dropping down to fill the vacancy. The energy difference between the two levels is carried away as an *X*-ray of energy E, as defined by the usual equation:

$$E = E_K - E_L = hc/\lambda$$

In the Auger process, an outer electron drops down to fill the vacancy as before, but instead of emitting a photon, a third electron is ejected, whose kinetic energy is approximately given by the difference between the energy levels involved:

$$E \approx E_K - E_L - E_M$$

(assuming that the vacancy is created in the innermost K shell, with an L electron dropping down and an M electron being emitted as an Auger electron). The Auger process is termed a *radiationless transition.* Figure 2.6 illustrates these two processes – in both cases the hole is in the K shell, and an L electron drops down to fill the vacancy. The probability that an inner shell vacancy will de-excite by either of these processes depends on the energy level of the initial vacancy, and on the atomic weight of the atom. The fluorescent yield, ω, is defined as the number of *X*-ray photons emitted per unit vacancy, and is a measure of the probability

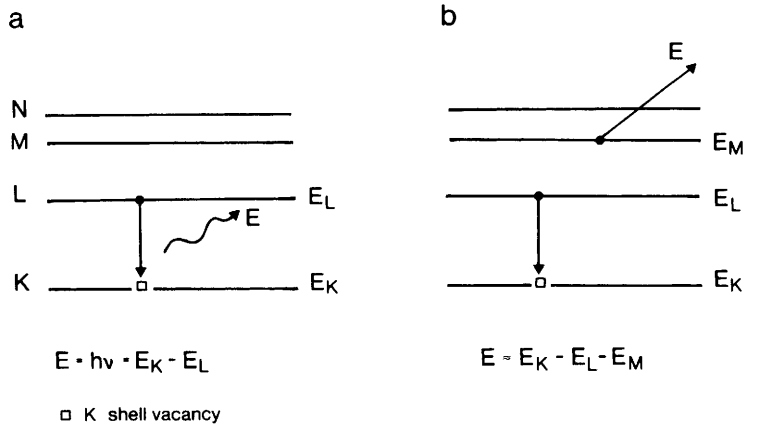


Figure 2.6 *The* X*-ray emission and Auger processes. An inner vacancy in the K shell de-excites via one of two competing processes* – (a) X*-ray emission, in which an L electron drops down and the excess energy is carried away by an* X*-ray photon, or* (b) *the Auger process, in which an L electron drops down, but the excess energy is carried away by a third electron – in this case from the M shell*

(value between 0 and 1) of a particular vacancy resulting in an emitted *X*-ray. Fluorescent yields are defined for each energy level (ω_K, ω_L, *etc.*), but in practice the K shell Auger processes are only really significant for the lighter elements. L and higher level vacancies, however, are more likely to result in Auger electrons than in *X*-ray photons.

As with optical emission lines, selection rules apply, defining those transitions which are allowed. The details are available in standard texts (*e.g.*, Jenkins, 1988), but the net results for all elements is that vacancies created in the K shell give rise to two separable emission lines – a stronger one, labelled the K_α line, which results from the $2p$–$1s$ transitions, and a weaker one labelled the K_β, resulting from $3p$–$1s$, $4p$–$1s$, $3d$–$1s$ and $4d$–$1s$ transitions. Although this shows that several transitions may contribute to these α and β lines, the resolution of most detection systems is insufficient to separate the fine detail within the α and β lines, and most tabulations of emission line energies list only the average value. The separation in energy between the K_α and K_β lines varies from element to element, increasing with atomic weight, and the intensity of the K_β is typically only 10% of the K_α. Figure 2.7 shows the detailed attributions of the K transitions in tin and the resulting *X*-ray emission lines at normal analytical resolution. The L spectra are considerably more complicated, but are usually only resolved into three lines, labelled L_α, L_β, and L_γ. These are emission lines which have originated with a vacancy created in the second (n = 2, or L) shell. The L_α line is the strongest, resulting from some $3d$–$2p$ transitions (specifically M_{IV}–L_{III} and M_V–L_{III}). The L_β line, principally due to another $3d$–$2p$ transition (M_{IV}–L_{II}), but including many others (up to seventeen separate transitions may contribute to the L_β peak), is normally only slightly weaker in intensity (perhaps 70% or more of the L_α). The L_γ is considerably weaker (typically 10% of L_α), and due largely to a $4d$–$2p$ transition (N_{IV}–L_{II}). Figure 2.8 shows the L_α transitions in gold and the resulting spectrum. The details of the relative intensities of each transition in *X*-ray emission depend on the quantum mechanical transition probabilities. Some, however (just as in optical transitions), are theoretically 'forbidden' by transition rules, such as $3d$–$1s$, but they can occur and appear as very weak lines. Other lines can appear in the spectrum, such as *satellite lines*, which result from transitions in doubly-ionized atoms – the Auger process, for example, leaves the atom in such a state, but these are usually very weak, and are not normally used for bulk chemical analysis.

The precise energy of an Auger electron (as given approximately by the equation above) is particularly sensitive to the chemical state of the atom from which it is ejected, since the outer orbitals from whence the

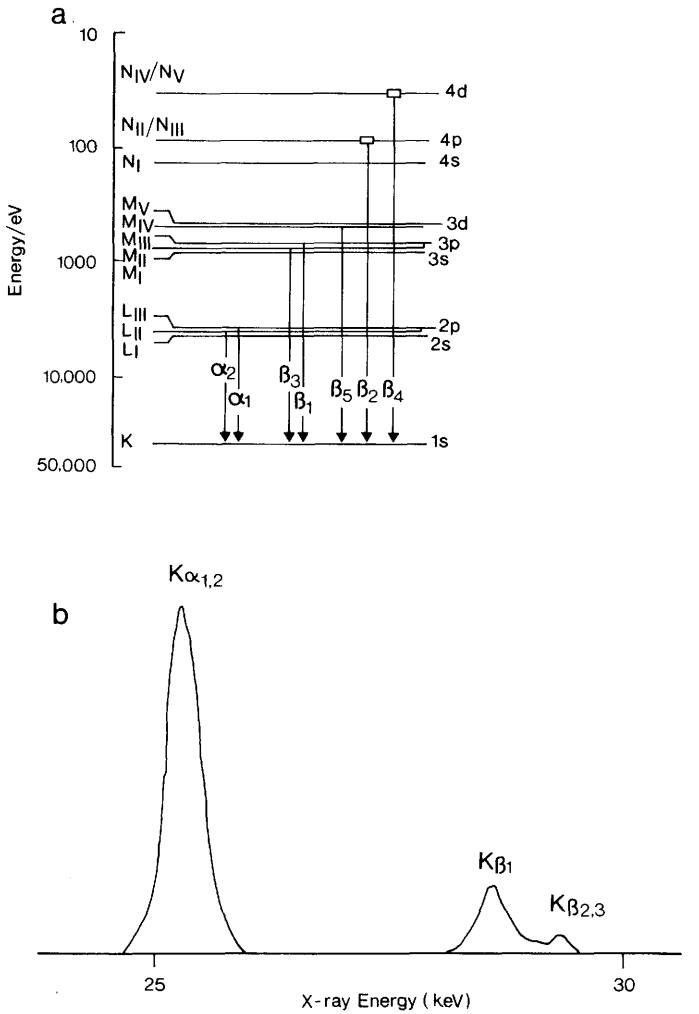
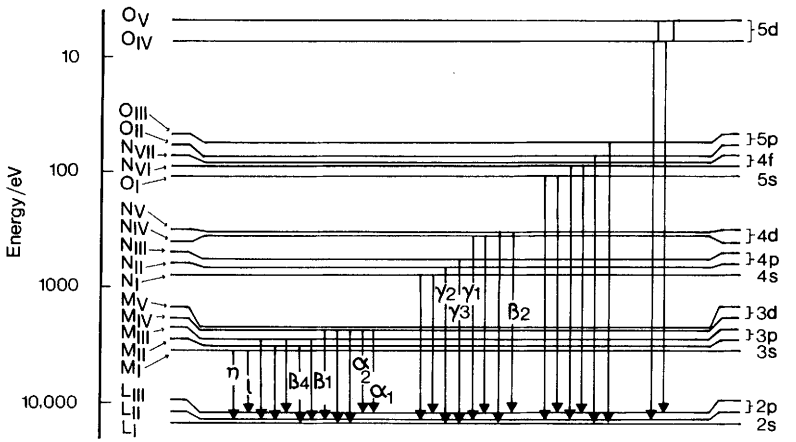
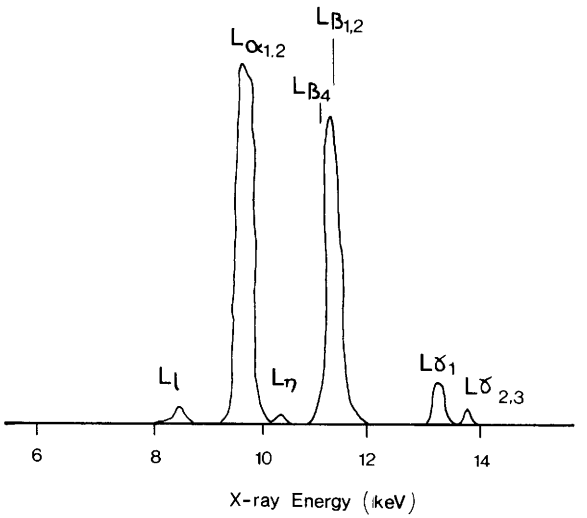


Figure 2.7 *Electronic transitions giving rise to the K spectrum of tin. The K spectrum is normally resolved into two lines, K_α and K_β (as shown), which are composed of unresolved* X*-rays from two and five transitions respectively*
(After Jenkins, 1974: Figure 2–4. © John Wiley & Sons, Ltd. Reprinted by permission of the publisher)

Auger electron originated are often involved in chemical bonding. Hence *Auger electron spectroscopy* (AES – the study of such electrons) is extremely valuable for looking at the chemical state of the surfaces of solids. It is possible, for example, to differentiate between clean, oxidized, and carbon-covered surfaces of metals (Briggs and Seah, 1990).



a



b

Figure 2.8 *Electronic transitions giving rise to the L spectrum of gold. The L spectrum is considerably more complicated with three main lines normally resolved as shown in the accompanying spectrum – L_α (arising from two transitions), L_β (with up to 17 contributing transitions) and L_γ (up to 8 transitions), plus a number of 'forbidden' transitions*
(After Jenkins, 1974; Figure 2–11. © John Wiley & Sons, Ltd. Reprinted by permission of the publisher)

The extreme surface sensitivity of Auger spectroscopy arises from the fact that these Auger electrons have a very low kinetic energy (usually less than 1500 eV), and so only emerge with interpretable information if they originate from the top 25 Å of the solid (*i.e.*, from the top two or three atomic layers). This is a particularly valuable attribute when studying the extreme surfaces of materials, but has restricted its applications in archaeology. One of the few examples is a detailed study of the corroded surfaces of Medieval glass (Dawson *et al.*, 1978). In marked contrast to this, because they arise from inner shell transitions, the chemical environment has virtually no influence on the energy of the emitted *X*-rays resulting from the competing de-excitation process (apart from in specialized studies of the lightest elements in the solid state). Consequently, the emission spectra of *X*-rays are uniquely and quantitatively characteristic of the parent atom, making techniques which involve the study of such *X*-rays very powerful for chemical analysis.

Three analytical techniques which differ in how the primary vacancies are created share the use of such *X*-rays to identify the elements present. In one (X-*ray fluorescence*, or XRF), the solid sample is irradiated by an *X*-ray beam (called the *primary beam*), which interacts with the atoms in the solid to create inner shell vacancies, which then de-excite via the emission of *secondary* or *fluorescent* *X*-rays – hence the name of the technique. The second uses a beam of electrons to create the initial vacancies, giving rise to the family of techniques known collectively as *electron microscopy*. The third and most recently developed instrumentation uses (usually) a proton beam to cause the initial vacancies, and is known as *particle-* (or *proton-*) *induced* X-*ray emission* (PIXE).

X-Ray Fluorescence Spectrometry

In *X*-ray fluorescence, the primary *X*-rays (those used to irradiate the sample) are usually produced by an *X*-ray tube, although in certain circumstances powerful radioactive sources can be used to do the same job – these give rise to portable machines, but the irradiation intensity is much lower. An *X*-ray tube has an anode made from a metal (often tungsten or molybdenum) which emits *X*-rays efficiently when bombarded with electrons. Thus a tungsten anode will emit the characteristic *X*-ray lines of tungsten, but in addition to line spectra, a solid is capable of emitting a continuous *X*-ray spectrum when bombarded with electrons. This is because the high energy electrons impacting the target material can give up their energy in a stepwise series of processes, each resulting in an *X*-ray, and thus the output contains a range of *X*-ray energies (termed *bremsstrahlung*), up to a maximum limit set by the

accelerating voltage applied to the electrons in the *X*-ray tube. The output of such a tube therefore consists of a continuum up to this maximum, superimposed upon which is the line spectrum of the target material. It is important to know what the target material is when using an *X*-ray tube, because its characteristic lines will almost certainly be detected in the secondary *X*-ray spectrum of the sample, and must be discounted.

When the primary *X*-rays from the tube strike the sample, two processes take place – *scattering* and *absorption*. Scattering may be elastic (*coherent*, or *Rayleigh scattering*), in which case the scattered ray has the same wavelength as the primary beam, or inelastic (*incoherent*, or *Compton scattering*), which results in longer wavelength *X*-rays. Coherent scattering results in the primary spectrum from the *X*-ray tube being 'reflected' into the detector (hence the appearance of the characteristic tube lines in the resulting spectrum), and can sometimes also result in diffraction phenomena. Incoherent scattering results in scattered *X*-rays which have longer wavelengths (lower energies) than the primary beam, and sometimes gives rise to a broadened inelastic peak at the lower energy side of the coherently scattered characteristic tube lines, as well as contributing to the general background. Vacancies occur in the atoms of the sample as a result of energy absorption from the primary beam, when part of the primary energy is transferred to the atom, resulting in the ejection of an orbital electron. When an electron is ejected from an atom as the result of the impact of an *X*-ray photon, *photoelectric absorption* is said to have occurred, and the ejected electron is termed a *photoelectron*. Study of these photoelectrons is the basis of another surface sensitive chemical analytical technique called *X*-ray photoelectron spectroscopy (XPS, also referred to as *electron spectroscopy for chemical analysis*, or ESCA), which has been used sparingly in archaeology (*e.g.*, to study the black coatings applied to ceramic vessels; Gillies and Urch, 1983). Thus, on passage through matter, the incident *X*-ray beam is attenuated as a result of these processes. According to Beer's Law, the intensity of the beam [$I(\lambda)$] after travelling a distance x through the solid is given by:

$$I(\lambda) = I_o \exp(-\mu\rho x)$$

where μ is the *mass absorption coefficient* of the material, of density ρ, and I_o the intensity of the primary beam. The mass absorption coefficient (absorption per unit mass) is a function of the atomic number of the material, and varies with the wavelength (*i.e.*, it is a function of λ). It can be calculated at any wavelength for a complex material by simply summing the mass absorption coefficients of all the elements present, weighted by their fractional abundance. Several tabulations of mass

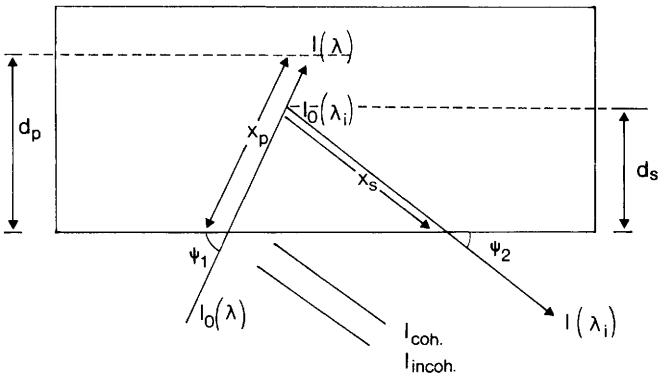


Figure 2.9 *Interaction of the primary X-ray beam with a solid sample. See text for discussion*
(After Jenkins, 1974; Figure 3–3. © John Wiley & Sons, Ltd. Reprinted by permission of the publisher)

absorption coefficients are available over the normal range of *X*-ray wavelengths (*e.g.*, Jenkins, 1988). These need to be calculated for the sample concerned in order to produce fully quantitative analytical data by XRF, but this is now normally done by the associated software controlling the machine.

Figure 2.9 shows a schematic diagram of the interaction of the primary *X*-ray beam of intensity $I_o(\lambda)$ incident at an angle of ψ_1 with a flat sample. At some depth d_s the primary beam, whose intensity has been attenuated according to Beer's Law (the *primary absorption*), creates a vacancy which gives rise to a secondary (fluorescent) *X*-ray with the characteristic wavelength λ_i, which travels towards the detector in the direction characterized by angle ψ_2. On its way out of the solid, it has to travel a distance x_s through the absorbing medium, and its intensity is also reduced, again according to Beer's Law (the *secondary absorption*). The amount of attenuation experienced by the secondary *X*-ray beam depends on the absorbance of the matrix, and the distance travelled, which depends on the angle ψ_2. Eventually attenuation is so severe that *X*-rays generated at or greater than the depth d_p cannot escape from the solid – this is termed the *escape depth* and is an important factor in XRF – it limits the depth from which analytical information can be obtained. It clearly depends on the nature of the solid matrix – absorption will be much greater in metallic lead, for example, which has a very high average atomic number (and therefore high absorption) compared with something like glass, which has a low average atomic number. Calculations show that for the lighter elements such as sodium in glass, for

example, the intensity of the secondary radiation is reduced by 90% after passing through only 14 μm of glass. A similar reduction is observed with the characteristic radiation of calcium after passage through 122 μm of glass (Cox and Pollard, 1977). For radiation from the heavier elements in glass, such as that emitted from lead, the escape depth is in excess of half a millimetre. Although not as surface sensitive as AES and XPS, XRF is still essentially a surface analytical technique, but the exact depth from which information can be obtained depends on the elements of interest and the nature of the solid matrix.

The secondary *X*-radiation coming from the surface of a solid sample which is being irradiated by *X*-rays contains a number of components. Of greatest interest to the analyst is the line spectra of the elements contained in the sample – this is the basis of the identification and quantification of the sample chemistry. These line spectra are superimposed on an elastic and inelastic scattered version of the primary irradiation from the *X*-ray tube, including the characteristic lines of the tube target material, plus a continuous background arising from unspecific processes within the sample. A typical XRF spectrum from a ceramic sample is shown in Figure 4.11. The prime requirement of an *X*-ray fluorescence spectrometer is therefore that it can resolve the separate peaks, identify them (either by measuring their wavelength or energy) and measure their area in order to quantify the data. Two approaches are possible, which mirror the particle-wave duality of electromagnetic radiation – *energy dispersive* X-*ray fluorescence* (EDXRF, or EDAX) and *wavelength dispersive XRF* (WDXRF). Figure 2.10 shows a schematic comparison of the two techniques. In the former, the secondary *X*-ray emitted by the excited atom within the sample is considered to be a

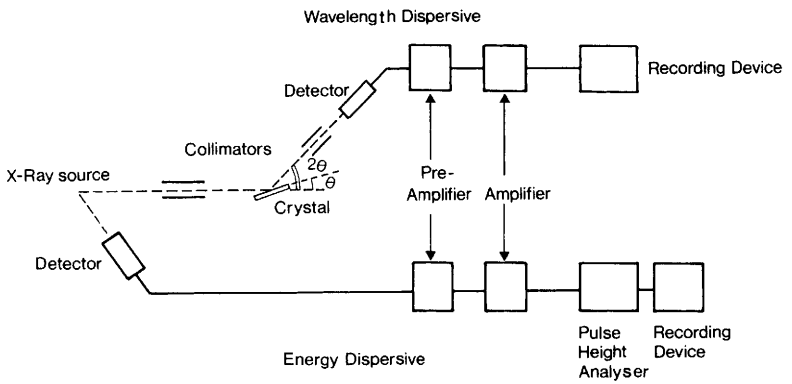


Figure 2.10 *Comparison of EDXRF and WDXRF*

particle (an *X*-ray photon), whose energy is characteristic of the atom from whence it came. An EDXRF system consists of a solid state device which measures the energy of the photon, and counts the number of photons with known energies. The heart of the system is (usually) a single crystal of silicon, doped (or 'drifted') with lithium to reduce the electronic impurities in the crystal to an absolute minimum (although modern technology allows single crystals to be produced of sufficient purity not to require doping). The crystal is kept at liquid nitrogen temperatures to prevent the lithium diffusing out, and to reduce the electronic noise in the device. When an *X*-ray photon strikes the crystal, the whole energy of the photon is dissipated into the crystal by the creation of a large number of *electron-hole* pairs in the semiconductor – for a Si(Li) detector, each electron-hole pair requires 3.8 eV to form, and therefore the number of pairs created is given by the energy of the incident photon divided by 3.8. A voltage is applied across the crystal via (usually) gold contacts, and the electrons created in the crystal move towards the positive terminal. This constitutes an electric current, the magnitude of which is proportional to the energy on the photon. Through a series of electronic devices, the current is measured and a count of one added to the relevant channel in a 'histogram' in a multi-channel analyser, which therefore records the arrival of a photon within a specific energy band (usually 20 eV). The spectrum is recorded over the range of either 0–20 keV or 0–40 keV, which includes the K and L emission lines of all the elements of interest. Figure 4.11 is an example of the resulting spectrum, and the K and L energies corresponding to all the elements are tabulated in standard texts (*e.g.*, Jenkins, 1974, Appendices I–III; although this tabulation is by wavelength, energies can be simply obtained by the formula given below). EDXRF is therefore a simultaneous detection technique, in that information from all elements is recorded at the same time. Software in the controlling computer usually allows peaks to be identified and quantified, and often calibration programmes are included which perform absorption corrections using equations of the type given above. The detector has to be maintained under vacuum to ensure its cleanliness, and it is normally separated from the rest of the spectrometer (which may or may not be capable of being evacuated) by a thin beryllium window, which limits the performance of the system at the very light element end of the Periodic Table. In a fully evacuated system (*i.e.*, one in which the sample chamber can be evacuated), an EDXRF spectrometer should be able to detect elements as light as sodium in the Periodic Table, although the sensitivity here is usually relatively poor. Performance at the heavier end of the table is limited by the fact that the more energetic *X*-rays from the heavy elements may pass straight

through the thin detector crystal without being absorbed, thus also reducing analytical sensitivity.

In an EDXRF system, the two tasks of energy measurement and detection are carried out simultaneously. In a WDXRF system, on the other hand, the two processes are separated. The secondary *X*-rays from the sample are regarded as being electromagnetic waves, whose wavelength is characteristic of the atom from whence they came. The WDXRF system operates in a manner analogous to the optical systems described above – a dispersion device is used to separate the radiation into its component wavelengths, and a separate detection system records the intensity of the radiation as a function of wavelength. In this case, however, because of the extremely short wavelength of *X*-rays (typically 0.1–10 Å for the characteristic line spectra of interest) a conventional prism or diffraction grating would not work, since the spacings of the diffracting medium have to be similar to the wavelengths to be separated in order for diffraction to occur. Conveniently, however, nature has provided suitable diffraction gratings for *X*-rays in the form of crystals, whose atomic spacing is similar to the wavelength of *X*-rays. Early work with *X*-rays used crystals of calcite or rocksalt to disperse the beam, but modern spectrometers tend to use LiF for general work. Occasionally, more specialist crystals (such as ammonium dihydrogen phosphate) or different crystal geometries are used for particular applications (Jenkins, 1974; 88). Detection of the dispersed *X*-rays is achieved either with a scintillation counter (comprising a phosphor crystal, which emits light on the impact of an *X*-ray, and a photomultiplier to record the burst of light) for shorter wavelengths, or at longer wavelengths a gas flow proportional counter (a sealed tube containing a gas, which ionizes as the *X*-ray passes through it, and the resulting electron flow is measured via a wire anode passing through the gas). Often these detectors are used in tandem, to cover the entire range of *X*-ray wavelengths. In order to help in relating the energy-dispersive and wavelength-dispersive approaches to XRF, it is most helpful to remember the following simple equation, relating the energy E of an *X*-ray photon (in keV) with its equivalent wavelength λ (in Å):

$$E = 12.4/\lambda$$

As with the ICP spectrometers described above, two modes of operation are possible – sequential or simultaneous. In the simultaneous mode, a bank of *X*-ray detectors is aligned with the dispersive crystal, each making a specified angle with the crystal, so as to detect the characteristic wavelength of a pre-determined element. Up to 20 detectors may be

employed, giving information on up to 20 elements. This mode of operation is suited to a situation where a large number of identical samples have to be analysed as quickly as possible, such as in industrial quality control. In the sequential mode, a single detector is used. Traditionally, the detector was linked to the crystal via a goniometer, allowing *X*-ray intensity to be recorded on a chart recorder as a function of diffraction angle. More modern instruments use a computer controlled detector, which can be programmed to record as many elements as required by moving to the position corresponding to the diffraction angle of the characteristic wavelength of the element of interest, counting, and then moving on to the next.

X-ray fluorescence spectrometers are designed to accommodate solid samples – preferably prepared into a standard shape (usually a disk), and mounted flat in a sample holder. For bulk commercial metal samples, simply cutting to shape and polishing gives a good sample. In routine geological applications, the samples can either be sectioned, cut, mounted, and polished, or powdered and pressed into a disk with a suitable binding medium, or converted into a glass bead by fluxing with excess borax. Solid rock samples are likely to be unsatisfactory, however, if there is significant inhomogeneity, because in commercial XRF spectrometers the primary *X*-ray beam usually has a diameter of several millimetres, going up to more than a centimetre. Pressed disk or glass bead preparation can be time consuming, and a certain amount of care has to go into the choice of the binding medium or flux, but this approach is usually the preferred method for geological analysis. If the object to be analysed cannot be converted into a convenient sample form, as is often the case with archaeological material, specially modified spectrometers can be used which can accommodate large irregular-shaped objects, either in a specially designed sample chamber, or simply by holding the sample in some fixed geometry in front of the spectrometer in air. The inability to evacuate the sample chamber means that information on elements lighter than potassium is usually lost due to air absorption of the characteristic *X*-rays, but the advantages of having a portable machine which can produce qualitative analyses of museum objects have been well documented (*e.g.*, Hall *et al.*, 1973).

Using commercially available systems, it is generally assumed that WDXRF has lower limits of detection than EDXRF, and is capable of higher precision. Detailed comparisons, however, of the trace element analysis of geological material using both methods has shown that this is not necessarily the case (*e.g.*, Potts *et al.*, 1985). This work compared several parameters of interest – the limits of detection of major rock-forming elements as determined by EDXRF and WDXRF on fused glass

beads; the limits of detection for trace elements in pelletized powders as determined by both approaches; the analytical precision of routine EDXRF and WDXRF, and finally a 'blind test' of several WDXRF laboratories against their own EDXRF determinations. The general conclusions were that EDXRF has poorer limits of detection for the major light elements than WDXRF (typically 0.2–1 wt% oxide for Na_2O to SiO_2, compared to better than 0.1% by WDXRF), but that for the trace elements the limits of detection are comparable (3 to 20 ppm for Ni, Cu, Zn, Ga, Rb, Sr, Y, Zr, Nb, Pb, Th, and U). Analytical precisions were found to be comparable for most elements at the levels found in most silicate rocks (coefficients of variation in the range 0.5–6% for major element oxides, 1.5–16% for most trace elements), and the 'blind test' revealed that the EDXRF determinations were statistically indistinguishable from the data produced by a number of laboratories routinely employing WDXRF. The final point of note was that the limit on the accuracy of the data produced was not the analytical technique employed, but the quality of the data on the international rock standards used for calibration.

It must be emphasized that these figures for the performance of EDXRF systems are based on carefully prepared glass beads or pressed pellets, and cannot be applied directly to analyses performed by instruments adapted to work on unprepared samples, where degradation in detection levels and precision of a factor of two at least might be expected. EDXRF is also generally much faster than WDXRF systems, accumulating a usable spectrum in no more than 100 seconds, and the instrumentation is usually much cheaper to buy. For archaeological purposes, many laboratories have equipped themselves with EDXRF systems for rapid identification and semi-quantitative analysis of a wide range of archaeological materials, including metals, ceramics, glasses, jet, *etc.*, whereas WDXRF has had relatively little use on archaeological materials apart from some studies of ceramics, where the material can be treated effectively as rock samples.

One major factor in the archaeological use of EDXRF on essentially unprepared archaeological samples such as coins and glass has been a recognition of the problems of surface sensitivity. Although by no means the most surface sensitive of the analytical techniques (*e.g.*, by comparison with AES or ESCA), EDXRF can be regarded as only giving an analysis of the top fraction of a millimetre of the sample. In the case of metals, where the phenomenon of surface enrichment (either deliberately during manufacture or naturally as a result of the burial environment) has long been known (*e.g.*, Hall, 1961; Cowell and La Niece, 1991), this can be a critical restriction. Similar problems, due to the selective leaching of the

alkali elements, have been noted in glass (Cox and Pollard, 1977). Although careful sample preparation can minimize these difficulties and provide good quantitative data, it is usual to regard EDXRF analyses of unprepared archaeological materials as qualitative, or semi-quantitative. This does not compromise its usefulness in areas such as conservation, where a rapid identification of the material may be all that is required, or in other areas as a preliminary analytical survey technique.

Analytical Electron Microscopy

Although electron microscopy is approached in this chapter initially as a variant of *X*-ray fluorescence as an analytical technique, it is worth stating at the outset that electron microscopy is far more versatile than this. Many standard descriptions of electron microscopy approach the subject from the microscopy end, regarding it as a higher resolution version of optical microscopy. Although equally valid, the emphasis on the analytical capacity given here is justified in the context of a book on archaeological chemistry. Several standard texts, such as Reed (1993) and Joy *et al.* (1986), are devoted to the broad spectrum of analytical electron microscopy.

As noted above, there are several ways of creating an inner shell vacancy which may de-excite via the emission of a characteristic *X*-ray. *X*-Ray fluorescence uses a primary beam of *X*-rays, but suffers from the fact that the characteristic *X*-ray spectrum recorded from a solid sample contains a scattered version of the primary spectrum, increasing the background signal. The use of an electron beam to create inner shell vacancies, and thus stimulate *X*-ray emission, dates back to the early 1950s, and offers a number of major advantages over *X*-ray stimulation:

(i) Electrons, being charged particles, can be focused and steered using relatively simple electrostatic lenses (*X*-rays, being electromagnetic radiation, cannot be focused other than by using curved crystals in a manner similar to focusing mirrors). This means that the primary electron beam can be focused down to a spot diameter of around 1 μm, compared to the millimetre size of *X*-ray beams, allowing small features such as individual mineral inclusions in a rock or ceramic matrix to be chemically analysed. Additional benefits, such as being able to scan the beam across the surface of a sample (described below) allow spatially-resolved analyses to be carried out.

(ii) The optical imaging capability of a microscope using an electron beam can be combined with this analytical facility to allow the

operator to observe clearly the components of the sample which are being analysed.
(iii) The use of electrons to stimulate the characteristic *X*-rays reduces the *X*-ray background in the detector, improving the analytical detection levels.
(iv) The various phenomena associated with electron scattering from solid surfaces can be used to estimate the average atomic weight of different regions within the sample, assisting the characterization of the phase structure of the sample.

Figure 2.11 shows a simplified diagram of the operation of an electron microscope, fitted with a wavelength dispersive *X*-ray detector. The primary beam of electrons is produced in a conventional electron gun, where a heated cathode maintained at ground potential emits electrons which are extracted via a positive potential into the focusing elements of the microscope. This beam is focused into a small cross-sectional area

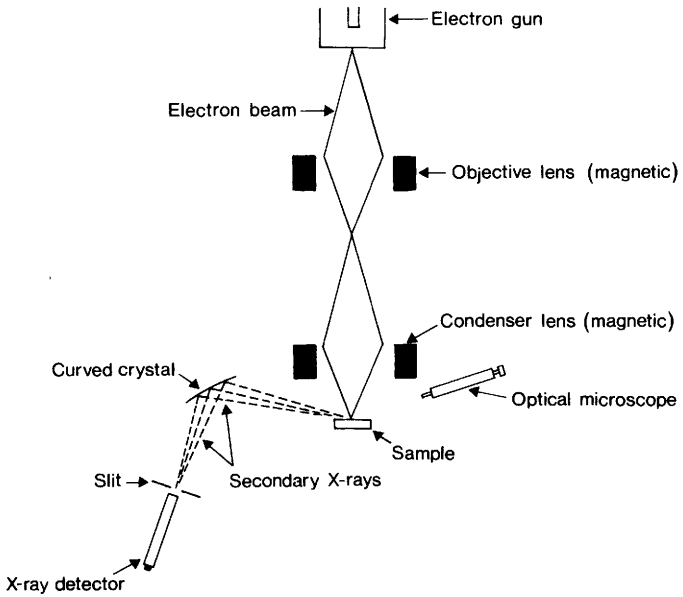


Figure 2.11 *Schematic diagram of an electron microscope with a WD-X-ray detector* (From *Instrumental Methods of Analysis* by H.H. Willard, L.L. Merritt, J.A. Dean and F.A. Settle, Figure 13.18. Copyright © 1988 Wadsworth Publishing Company. By permission of Brooks/Cole Publishing Company, a division of International Thomson Publishing Inc., Pacific Grove, CA 93950)

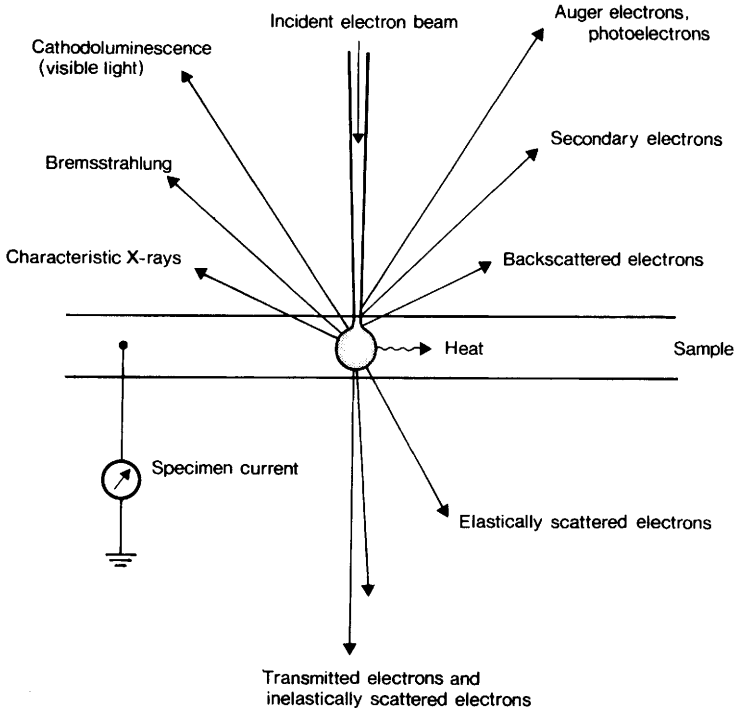


Figure 2.12 *Interaction of primary electrons with a thin solid sample, showing the various processes which can take place*
(After Woldseth, 1973; Figure 4.1)

and can be steered to any point on the sample by a series of magnetic lenses. Once the electron beam strikes the sample, a number of processes take place, as illustrated in Figure 2.12. These include the following:

(i) *Secondary electrons* are very low energy electrons (less than 50 eV) knocked out of the loosely bound outer electronic orbitals of surface atoms. Because of their low energy, they can only escape from atoms in the top few atomic layers, and they are very sensitive to surface topography – protruding surface features are more likely to produce secondary electrons which can escape and be detected than are depressed features. The intensity of secondary electrons across the sample surface therefore accurately reflects topography, and is the basis of the image formation process in electron microscopy.
(ii) *Backscattered electrons* are of higher energy, and result from inter-

actions of the incident beam with the nucleus of the atoms. Their higher energy means that they can escape from deeper within the sample than secondary electrons, so they do not reflect surface topography. Their intensity is, however, proportional to the atomic weight (and therefore atomic number) of the interacting nucleus, and the intensity variation across a surface is therefore proportional to the average atomic number of the surface. This gives what is known as a *'backscattered electron (bse) image'*, and contains useful chemical information.

(iii) Some incident electrons will create inner shell vacancies, in a manner similar to that of *X*-rays described above. The electrons ejected by the primary beam are called *photoelectrons*, and could be used analytically, but are generally neglected. The inner shell vacancy can de-excite via the Auger process (*Auger electrons* are also generally neglected) or via the emission of characteristic *X*-rays, which are the basis of the analytical operation of the electron microscope.

(iv) If the sample is thin enough, incident electrons may go straight through and be detected, as well as elastically and inelastically scattered electrons which are scattered in a forward direction. These form the basis of *transmission electron microscopy* (TEM), which is beyond the scope of this chapter, and has as yet had limited application in archaeology – an exception is the work of Barber and Freestone (1990), who used TEM to identify the nature of the tiny particles involved in the phenomenon of dichroism in the magnificent example of Roman glass known as the Lycurgus Cup.

As an analytical tool, the electron microscope can be operated with either a wavelength-dispersive *X*-ray detector (in which case it is often called an *electron microprobe*), or an energy dispersive detector. Both are essentially as described in the previous sections, with the same advantages and disadvantages, and operational characteristics. Some of the more powerful instruments are fitted with both. Comparative studies have been made of the performance of WD and ED detectors in electron microprobe analysis of silicates (Dunham and Wilkinson, 1978) and, more recently, vitreous material (Verità *et al.*, 1994). Both of these studies agree that the accuracy and precision yielded by both techniques are comparable over the normal ranges found in geological and archaeological material, but ED has considerably poorer limits of detection for all elements, typically by one or two orders of magnitude. The exact figures depend on matrix and counting time, but representative figures for the limit of detection by ED detection are 0.05–0.26 wt% of the element.

These figures are comparable to, or slightly poorer than, the equivalent figures for *X*-ray fluorescence instruments, but the principal advantage given by electron beam stimulation is in the 'steerability' of the primary beam. Not only does this allow the analysis to be carried out on small regions of the sample, identified by either optical or electron microscopy, but it allows line scans and area scans to be carried out, giving spatially resolved chemical information from the surface of the sample. Line scans (involving monitoring the change of elemental composition as the electron beam is moved slowly across the surface) are particularly valuable in the study of the changes caused by corrosion and artificial patination of metal surfaces if the sample has been mounted so as to present a cross-section at its surface. Area scans give rise to *'elemental maps'* of the sample surface, by *rastering* the electron beam across the surface and monitoring the characteristic wavelength of the element of interest. Thus, for example, the distribution of lead can be looked at across the surface of a section through a copper alloy object. This facility is particularly valuable in archaeology, where the samples tend to be inhomogeneous and corroded. Several examples of the applications of this type of work to the study of archaeological metalwork can be found in the work of Cowell and La Niece (1991). It should be pointed out that most routine analyses carried out by electron microscopy are performed on prepared samples – cut and mounted to fit standard sample holders, which typically accommodate samples several millimetres across. Modern machines can be obtained with so-called 'environmental sample chambers' which allow large unprepared samples to be accommodated, which is obviously beneficial for archaeological material. All samples have to be electrically conducting for electron microscopy. Metal samples are no problem, but electrical insulators such as glass and ceramics normally need to be coated with a thin layer of carbon or gold to ensure that they do not charge up and deflect the primary electron beam away from the surface.

Proton-induced *X*-Ray Emission

During the 1980s an alternative approach to *X*-ray analyses of inorganic materials was developed, utilizing the existence of large van de Graaff accelerators (developed for particle beam research), and their ability to produce high intensity, highly focused, beams of particles. These beams can be 'tapped off' from the accelerator and focused onto a sample outside the accelerator (*i.e.*, not in a chamber under high vacuum). This is ideal for archaeological material, since it removes the need for sampling (although prepared sections can of course also be analysed).

The particle beam most usually used for analytical purposes consist of protons, which can be focused and steered just like electrons. The first generation of proton-induced *X*-ray emission (PIXE) machines had beam diameters of the order of half a millimetre (Fleming and Swann, 1986), but more recent machines have μm diameter beams (termed μ-PIXE: Johansson and Campbell, 1988), giving spatial resolution similar to that obtainable by electron microprobe analysis. The proton beam strikes the sample, producing inner shell vacancies, which, as before, may de-excite via the emission of characteristic *X*-rays, which are detected using an energy-dispersive detector as described above. The major advantage offered by this instrument over an electron microprobe is that the use of a primary beam of protons does not give rise to such a high *X*-ray background as experienced by electron stimulation, thus giving improved analytical sensitivity. The reason is that when electrons interact with solids, a relatively high background of *X*-rays is produced (the *bremsstrahlung*), as described above. Protons, being heavier and accelerated to a higher energy, tend to suffer less energy loss on their passage through the sample, producing less *bremsstrahlung*. Detection levels in conventional PIXE may be as low as 0.5–5 ppm (parts per million) for a wide range of elements in thin organic specimens such as biological tissue samples (Johansson and Campbell, 1988).

Archaeological work tends to use the external beam arrangement, whereby the whole object, or a tiny sample removed from it, is situated outside the accelerator, and the passage of the primary beam and (more importantly) the characteristic *X*-rays through air (or sometimes helium to minimize absorption) limits the sensitivity of the method, particularly for the light elements. Even so, lower limits of detection of better than 100 ppm have been reported for elements above calcium in the Periodic Table in archaeological material. Analyses of archaeological and art historical samples ranging from ceramics, metals, paintings, and even postage stamps have been reviewed by Johansson and Campbell (1988; Chapter 14).

NEUTRON ACTIVATION ANALYSIS

Until the advent of ICP and PIXE during the 1980s, the standard analytical method for producing multi-element analyses with detection limits at the ppm level or better was neutron activation analysis (NAA), and the archaeological literature has many examples of its application. A recent review of the analytical methods used in geological publications calculated that 50% of the trace element data reported in the major geochemical journals between 1990 and 1992 had still been obtained by

NAA (Jackson *et al.*, 1993). NAA has been used on archaeological material from almost the inception of the technique in 1950s, particularly for coinage (Kraay, 1958: technique described by Emeleus, 1958), and ceramics (Sayre and Dodson, 1957). It is still a major technique in archaeological chemistry for these materials (*e.g.*, Hughes *et al.*, 1991), but increasing difficulties associated with obtaining irradiation facilities, and increasing competition from ICP-MS, are rapidly challenging this position.

Neutron activation analysis at its simplest is a technique whereby some of the elements in the sample are converted into artificial radioactive elements by irradiation with neutrons. Figure 2.13 shows a schematic diagram of the process. These artificial nuclei decay by one or more of the standard pathways for radioactive decay (normally involving the emission of an α, β, or γ particle). By suitable instrumentation the radioactive decay can be detected, and, by measuring the intensity of the emission, related back to the original concentration of the parent element in the irradiated sample. Most often the radiation detected is γ emission, although β emission can also be used.

A nuclear reactor of the type used for activation analysis is essentially a source of high fluxes of neutrons – typically of the order of 10^{12} neutrons cm^{-2} s^{-1}. Most irradiations are carried out using relatively low energy neutrons, known as *thermal neutrons*, with kinetic energies less than

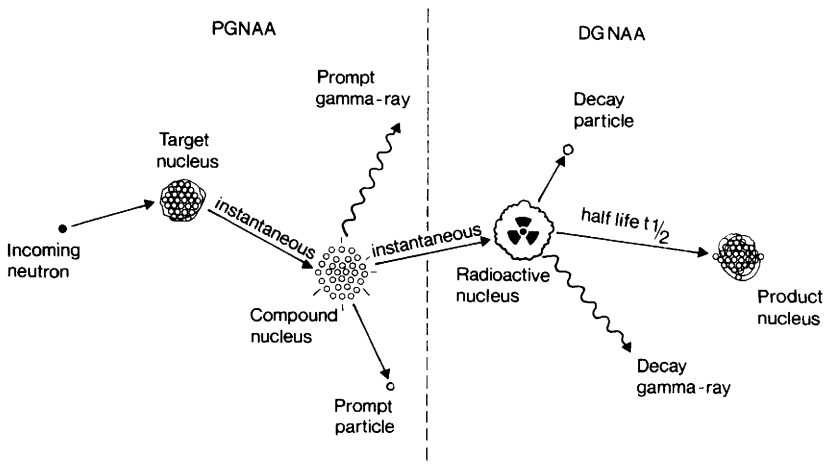


Figure 2.13 *Schematic diagram of the nuclear processes involved in neutron activation analysis. Prompt gamma neutron activation analysis (PGNAA) occurs within the reactor; delayed gamma NAA (DGNAA) occurs at some remote site* (After Glascock, 1994; Figure 1. © John Wiley & Sons, Ltd. Reprinted by permission of the publisher)

0.2 eV. Other applications can use higher energy neutrons, which is known as *fast neutron activation analysis* (FNAA), but for most elements slow neutrons are the most effective at causing nuclear transformations. At these energies, the thermal neutrons can be 'captured' by the nuclei of elements in the sample, causing an instantaneous nuclear transformation to take place. There are several types of possible reactions, but the most relevant in activation analysis are of a type known as (n,p) or, more commonly, (n,γ) reactions. Detailed discussions of all of the nuclear reactions relevant to the common components in archaeological ceramics are given in the appendix to the paper by Perlman and Asaro (1969). A brief explanation of this notation is given here. As noted in Appendix 2, a nucleus can be thought of as being composed of Z protons (the atomic number, which uniquely defines the element, *e.g.*, $Z = 11$ is the element sodium) and N neutrons, where the neutron number can vary, giving rise to different isotopes of the same element. The symbol A is the total number of nucleons (protons plus neutrons) in the nucleus: clearly $A = Z + N$. Isotopes of the same element differ in their A number, and the average atomic weight of the element is the weighted average of the natural abundances of the different isotopes (*e.g.*, chlorine, $Z = 17$, has two stable isotopes of $A = 35$ and 37, with natural abundances 75% and 25% respectively, giving an average atomic weight of 35.5). Symbolically, the notation ${}^{A}_{Z}\mathrm{X}$ is used, where X stands for the chemical symbol of the element defined by atomic number Z. The superscript refers to the number of atomic mass units in the nucleus (with the proton and neutron regarded for these purposes as having the same mass of one unit) and the subscript as the number of positive charges in the nucleus. Appendix 4 lists the atomic numbers and atomic weights of the elements.

When a slow neutron is captured by the nucleus of element X, another isotope of the same element is instantaneously formed, in an excited state because of the impact (labelled 'compound nucleus' in Figure 2.13), which then de-excites by the emission of a gamma particle (and possibly other particles) from the nucleus to produce a radioactive nucleus. For example, when ^{23}Na captures a neutron (signified by ${}^{1}_{0}\mathrm{n}$, since neutrons have a mass of one unit, but no electrical charge), it becomes the radioactive nucleus ^{24}Na, as follows:

$$ {}^{23}_{11}\mathrm{Na} + {}^{1}_{0}\mathrm{n} = {}^{24}_{11}\mathrm{Na} + \gamma $$

The γ particle is emitted virtually instantaneously on the capture of the neutron, and is known as a *prompt* γ – it can be used analytically, in a technique known as *prompt gamma neutron activation analysis* (PGNAA), but under the conditions normally used it would be lost within the nuclear

reactor. In this case, no other prompt particle is emitted. The isotope of sodium formed by this reaction (^{24}Na) is radioactively unstable, decaying by beta emission to the element magnesium (the 'product nucleus' in Figure 2.13), as follows:

$$^{24}_{11}\text{Na} = ^{24}_{12}\text{Mg} + ^{0}_{-1}\beta + \gamma$$

The beta particle emitted by this reaction has a mass of (conventionally) zero mass units, and a charge of -1; hence the notation $^{0}_{-1}\beta$. It should be noted that equations of this type should be balanced top and bottom – *i.e.*, the sum of the superscripts on the right hand side should equal the sum of the superscripts on the left hand side, and the same for the subscripts. The half-life of ^{24}Na is approximately 0.623 days, and so the radioisotope formed decays away relatively quickly following irradiation. The gamma particle has a characteristic energy of 1369 keV, and so the decay of ^{24}Na to ^{24}Mg can be monitored by either measuring the β particle or the γ particle. It is important to appreciate that the γs in the above two equations are quite different, and for the purposes of ordinary (delayed) NAA only that given in the second equation is of any value. In the notation used by radiochemists, all of the above information can be summarized in the following formulation:

$$^{23}\text{Na}(\text{n},\gamma)^{24}\text{Na};\ 0.623\ \text{d},\ 1369\ \text{keV}$$

which means that the atom of ^{23}Na is converted into ^{24}Na via the process called neutron capture, accompanied by the emission of a prompt gamma. The nucleus formed (^{24}Na) is radioactive, decaying by the emission of a gamma particle of energy 1369 keV, and having a half-life of 0.623 days (15 h). This statement contains all the information required by a radiochemist to understand the process, although, as can be seen, some of the detail is omitted.

A slightly more complicated nuclear reaction, but one which is also utilized in the analysis of archaeological material, is the so-called *transmutation*, or (n,p) reaction. In this case, the nucleus captures a neutron, but internal re-arrangements occur, and a proton is immediately ejected from the nucleus (the 'prompt particle'), changing the Z number and hence the chemical identity of the nucleus. For example, neutron irradiation of titanium does not result in any isotopes of titanium which have half-lives suitable for measurement, but the transmutation reaction ^{47}Ti(n,p)^{47}Sc yields the radioactive isotope ^{47}Sc, which decays by β emission, with the emission of a γ particle with an energy of 159 keV and a half-life of 3.43 days. Written out in full, this is:

$$^{47}_{22}\mathrm{Ti} + {}^{1}_{0}\mathrm{n} = {}^{47}_{21}\mathrm{Sc} + {}^{1}_{1}\mathrm{p}$$

and:

$$^{47}_{21}\mathrm{Sc} = {}^{47}_{22}\mathrm{Ti} + {}^{0}_{-1}\beta + \gamma$$

or, more concisely:

$$^{47}\mathrm{Ti(n,p)}^{47}\mathrm{Sc};\ 3.43\ \mathrm{d},\ 159\ \mathrm{keV}$$

Although many other types of nuclear reaction are possible as a result of high neutron fluxes, these two are the ones of prime importance in radioanalytical chemistry. The two principal requirements for a reaction to be useful analytically are that the element of interest must be capable of undergoing a nuclear reaction of some sort, and the product of that reaction must itself be radioactively unstable. Ideally, the daughter nucleus should have a half-life which is in the range of a few days to a few months, and should emit a particle which has a characteristic energy, and is free from interference from other particles which may be produced by other elements within the sample.

The gamma detectors which are used to monitor the γ particles produced by the decay of the irradiated samples are essentially identical to the instruments used in energy dispersive *X*-ray fluorescence (described above), with the exception that the solid state detector is usually a single crystal of germanium rather than silicon. This is because γ rays have energies which are typically 100 times more than those of *X*-rays, and germanium is more efficient at measuring energies in this region. Originally lithium-drifted germanium detectors were used in order to get sufficiently pure crystals, but improvements in manufacturing technology have meant that pure ('*intrinsic*') germanium detectors can now be obtained of sufficient quality. They are still, however, maintained at liquid nitrogen temperatures in order to reduce the electronic noise in the system. The spectra produced are very similar to those shown for EDXRF, but have a higher energy range.

The standard procedure in use at the British Museum for ceramic analysis by NAA has recently been described (Hughes *et al.*, 1991), but is typical of the methods used on a range of materials. Powdered samples (40–80 mg) removed from the ceramic to be analysed (drilled or abraded using a material chosen to minimize contamination) are sealed into silica glass tubes (2 mm internal diameter, length 33 mm). Bundles of five to seven tubes are wrapped in aluminium foil, and then several bundles are packed into an irradiation canister, which can contain up to 70 samples. The canisters are then delivered to the nuclear reactor, where they are subjected to a known neutron flux for a fixed period of time, depending

on the elements to be determined. They are then returned to the laboratory four days after irradiation, for a programme of γ counting using a germanium solid state detector as described above. The British Museum programme allows 23 elements to be determined from the same irradiation, making the whole analytical procedure extremely economical in terms of information obtained for the cost and time involved.

Because the artificially produced radioisotopes of interest can have a wide range of half-lives, it is normal to have a measuring strategy which involves a series of distinctive steps. Some isotopes, such as ^{28}Al (produced by the reaction ^{27}Al(n,γ)^{28}Al) have half-lives so short (in this case 2.3 min) that they can only be measured if they are placed into the gamma counter immediately after removal from the irradiation source. They can in effect only be measured 'on-site' at the reactor. Others, with slightly longer half-lives (such as ^{24}Na, with a half-life of 0.623 days) can be measured if placed into the counter within a few days of irradiation. Those with relatively long half-lives (*e.g.*, ^{46}Sc at 80 days), are better left until the more short-lived isotopes have decayed away, reducing the radiation background and therefore the possible interferences. Measuring laboratories with an on-site reactor clearly have an advantage, in that they can measure the very short lived isotopes immediately. Those laboratories remote from the reactor have devised appropriate measurement programmes, the details of which depend on the nature of the material being analysed. The British Museum, for example, measures pottery samples for a period of 3000 s immediately after receipt of the irradiated samples (which is four days after irradiation) to measure Na, K, Ca, As, Sb, La, Sm, Yb, Lu, and Np (the latter to quantify U). A second measurement run is then carried out 18 days after irradiation, for a period of 6000 s, to measure Sc, Cr, Fe, Co, Rb, Cs, Ba, Ce, Eu, Tb, Hf, Ta, and Pa (to measure Th; Hughes *et al.*, 1991). Other laboratories use different strategies, and clearly the measurement of other materials such as metals requires a quite different strategy.

Detection levels for NAA can be as low as 1.5×10^{-5} ppb for very sensitive elements in a suitable matrix (Glascock, 1994; Table 1 – figure quoted is for Au in semiconductor grade Si). More routinely, one would expect to see figures of the order of 10 ppb or better, up to perhaps 10 ppm, for trace elements in geological or biological material. It is important to realize that not all elements can be analysed by 'normal' neutron activation analysis. As with all analytical techniques, there is a variation in the sensitivity and detection levels from element to element, but additionally with NAA there are some elements which cannot be 'seen' at all, perhaps because neutron irradiation as described does not

produce suitable radioactive nuclei, or possibly because the spectrum has severe spectral interference. Most important archaeologically are elements such as lead and silicon, which means that NAA cannot produce 'total analyses' for some metals, ceramics, and glasses. Some of these problems can be overcome using variations on NAA, such as PGNAA, but routinely they lead the analyst to resort to other methods for some elements.

There have been a limited number of comparisons of NAA (regarded as the established technique) and ICP (either using emission or mass spectrometric detection) on a range of sample materials. Bettinelli *et al.* (1992) compared ICP-AES, graphite furnace AAS, XRF, and NAA for the inorganic analysis of coal fly ash. No simple conclusions were obtained – certain elements (*e.g.*, As, Ca, Cr, Mn, Na, and V) gave good results by NAA, ICP, and XRF, whereas other elements (such as Zn) gave uniformly poor results. This study confirmed the conclusion that the principal limitation on the accuracy of those techniques which require solid reference samples (in this case XRF and NAA) is the quality of the certified Standard Reference Materials. Other, more limited, comparisons have been carried out, such as the study by Rouchaud *et al.* (1993) on impurities in aluminium metal at the 0.1 to 10 ppm level, which concluded that both methods were accurate in this case (apart, obviously, from those elements such as Si, Ca, and Mg which are not easily detected by NAA). Attention has also been paid to comparisons of these techniques in biological materials, such as in the work of Awadallah *et al.* (1986), who compared NAA, ICP-AES, and flameless AAS for the analysis of a range of Egyptian crops and associated soil samples. Again, the general conclusion was that all methods provided reliable results, or at least results in line with those on the certified standards.

More recently, attention has switched to comparing NAA with ICP-MS – acknowledged to be generally more sensitive than ICP-AES, and with LA-ICP-MS, which has the distinct advantage of operating on solid samples. Ward *et al.* (1990) compared NAA and ICP-MS (solution and laser ablation) for the trace element analysis of biological reference materials. They found that both techniques gave good agreement with certified or published values for 18 elements down to the levels of 30 ppb. The comparison between solid and solution ICP-MS work is described as 'fair for most elements'. This comparison has been taken further by Durrant and Ward (1993) who compared LA-ICP-MS with NAA on seven Chinese reference soils. Thirty elements were analysed, and it is stated that 80% of the LA-ICP-MS measurements were within a factor of two of the NAA determinations, although many were considerably closer. Precisions of 2–10% were obtained for most elements by LA-

ICP-MS, although this figure deteriorated at lower concentrations. Clearly, LA-ICP-MS has some way to go to compete with NAA, but most ICP-MS machines are now available with a laser ablation chamber, which should lead rapidly to the development of improved techniques.

MASS SPECTROMETRIC TECHNIQUES

Mass spectrometry is becoming increasingly important in archaeology, both as a technique in its own right, and also as a sensitive detection system in the so-called *hyphenated techniques* as already described, *e.g.*, ICP-MS, GC-MS, *etc.* It is used in its own right for the determination of heavy stable isotope ratios such as lead (Chapter 9), or for light stable isotopes as used in dietary reconstruction (*e.g.*, van der Merwe, 1992). It also forms the basis of several dating techniques involving radioactive isotope determinations (*e.g.*, K–Ar dating) or in accelerator-based methods for radiocarbon dating purposes (Aitken, 1990). In the past ten years, however, it has played an expanding role as a 'bolt-on' detection system offering much greater sensitivity (as well as isotopically resolved measurements, if required) in a wide range of analytical applications. A good example of this is its use as a detector in ICP (inductively coupled plasma, as described on pp. 31 and 33)

Mass spectrometry is based on the principle that electrically charged atoms and molecules can be separated on the basis of their different atomic masses (strictly, mass-to-charge ratio) by controlling their motion through externally imposed electrical and/or magnetic fields. A simple mass spectrometer (of the type first constructed at the beginning of the twentieth century) consists of a source of positively charged ions of the same energy, a magnetic and/or electrostatic deflection system for separating the charged ions, and an ion collector to measure the current flowing in the selected beam. The pressure within the system must be kept low by pumping to ensure that the ions are not absorbed or deflected by passage through residual air. These components are indicated schematically in Figure 2.14. Modern spectrometers are designed to handle either a gaseous sample, such as CO_2, or a solid sample deposited on a wire (usually made of Ta, Re, or W). In the first case, ionization is achieved by injecting the gas into the vacuum system, and bombarding the stream of gas with a beam of electrons, which causes ionization. The positively charged particles are then extracted from the ion source and given a fixed energy by acceleration through fine holes in a series of negatively charged plates which also gives a degree of ion beam focusing. This type of ion source is suitable for the light elements which can be conveniently converted into gases, or for relatively volatile

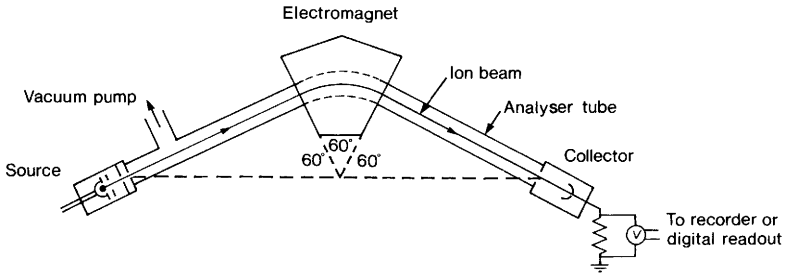


Figure 2.14 *Schematic diagram of a 60° sector mass spectrometer* (After Faure, 1986; Figure 4.9. Copyright 1986 John Wiley & Sons, Inc. Reprinted by permission of the publisher)

organic compounds. For heavier elements, a solid source is needed, and this is the basis of *thermal ionization mass spectrometry* (TIMS) which has become of prime importance in isotope geochemistry. Here the sample is deposited as one of its salts on a refractory metal wire, which is then loaded into the mass spectrometer and electrically heated to cause volatilization. The ions so produced are extracted and accelerated as before. Other types of ion source are available for specialist applications, and modern instrumentation uses more sophisticated systems than those described here, such as the use of a triple filament in TIMS, which allows samples as small as 10^{-9} g to be measured satisfactorily (Duckworth *et al.*, 1986; 43).

Once the positive ions have been injected into the body of the mass spectrometer, the beam is separated into its component masses by one or more deflection devices. In the simpler system, such as that shown in Figure 2.14, this can be an electromagnet designed so that the magnetic field is uniform and perpendicular to the ion beam at the point where it passes the magnet. Under such conditions, the ions experience a force which deflects them in a circular path, the degree of deflection depending on their mass – the heavier ions are deflected less. The equations of motion are as follows. Assume the ion has a mass m and carries a charge of e. If it is extracted from the ion source by an applied voltage V, then it is given a kinetic energy of E, where:

$$E = eV = \tfrac{1}{2}mv^2$$

All ions leaving the ion source have the same kinetic energy E, but the velocity v of a particular ion will depend on its mass. The velocity of an ion is given by rearranging the above:

$$v = (2eV/m)^{1/2}$$

The equation of motion of a charged particle in a magnetic field of strength B is given by the following, where r is the radius of the circular track taken by the ion:

$$Bev = mv^2/r$$

Combining these last two equations to eliminate v gives the following:

$$Be = mv/r = m(2eV/m)^{1/2}/r$$
$$m/e = B^2r^2/2V$$

Giving:

$$r = 1.414(Vm/e)^{1/2}/B$$

Calculations using this equation can be simplified if it is assumed that m is in atomic mass units, e in units of atomic charge (*i.e.*, for a singly ionized atom $e = 1$), B in gauss, and r in cm. It can be seen that the radius taken by an ion is dependent on the square root of the accelerating voltage V and inversely proportional to the magnetic field strength B. For a fixed geometry of mass spectrometer, therefore, the value of these two parameters governs which ion will pass through the magnet and into the detector. If B and V are known, the value of m can be calculated for those ions which are detected. Alternatively, if either B or V is varied systematically, ions of different m will be selected to pass through sequentially. Thus, if one or other is scanned, the mass spectrum of the sample can be obtained. At its simplest, the detector is a Faraday cup-type charge detector which monitors the current of the beam as it is passed to earth – the magnitude of the current flowing is directly proportional to the number of ions being received at the detector.

For mass spectrometry in organic chemistry, where the aim is to identify the mass of the molecular ions present (and therefore identify the chemical nature of the ionized molecules and fragments), the important factor to be able to measure is the precise value of the mass, and a spectrometer of the type shown in Figure 2.14 is adequate. For isotope work in geology and archaeology, what is often more important is the abundance of the ions, or the abundance ratio of two ions, such as $^{13}C/^{12}C$. For this work, it is better to have a dual collector system, so that both ion beams can be monitored at the same time, thus eliminating any fluctuation in the intensity of the beam leaving the ion source and hence giving a more precise ratio estimate. More complex systems have been

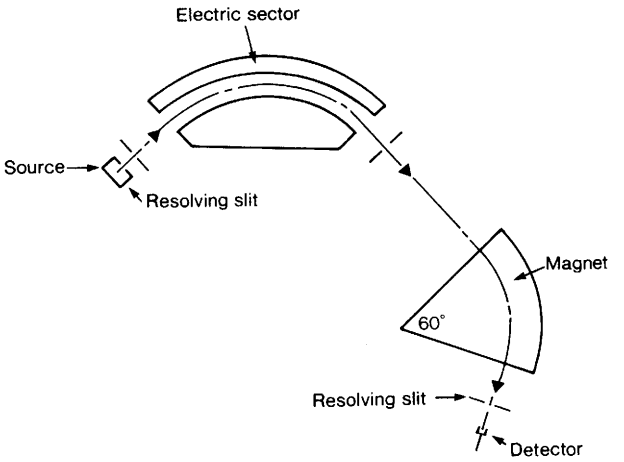


Figure 2.15 *A double focusing mass spectrometer*
(After Beynon and Brenton, 1982; Figure 4.9, by permission of University of Wales Press)

evolved for specific types of measurement, such as the use of the 'double focusing' spectrometer in which enhanced mass resolving power is given by the addition of an electrostatic field via electric sector plates in the path of the ion beam before the electromagnet (Figure 2.15). This vastly improves the mass resolution of the system (Duckworth *et al.*, 1986; 92).

Systems based on the principles outlined above are those most regularly used for 'straight' mass spectrometry. The increased use of MS as a detector in 'hyphenated' techniques has relied on a different approach, using a system known as the *quadrupole mass spectrometer.* This is much more compact than the magnetic sector devices shown above, and allows the mass spectrum to be scanned through very rapidly, but at the expense of poorer mass resolution. It does not require the ions to be accelerated to such high energies as the magnetic sector machines, making it suitable to receive ions from sources other than the conventional ion source, and therefore ideal for a 'bolt-on' detector. A quadrupole mass spectrometer consists of four metal rods arranged in parallel to each other, as shown in Figure 2.16. The ion beam passes down the gap at the centre of the rods. Typically these rods might have a length of around 20 cm, and a diameter of 1 cm, and they are electrically connected together in opposite pairs, as shown in the figure. If one pair carries a small positive potential and the other pair an equal negative potential the path down the centre of the rods is at zero potential. The

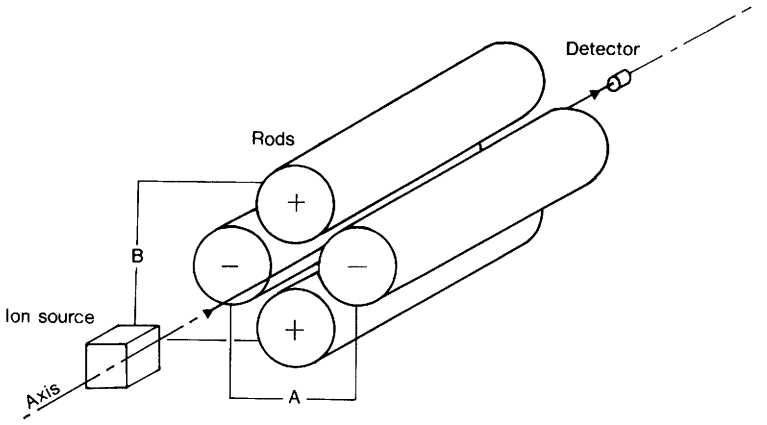


Figure 2.16 *A quadrupole mass spectrometer*
(Adapted from Beynon and Brenton, 1982; Figures 4.6 and 4.7, by permission of University of Wales Press).

ion beam can pass down the centre without deflection, but it is unstable with respect to the negative charges, since a positive beam is attracted to this polarity. In practice, the polarity of the rods is set to be a combination of a DC (steady) voltage and an oscillating voltage of frequency ω. The equations of motion of an ion beam under such conditions are complicated, and involve the solution of differential equations which are beyond the scope of this discussion (see, for example, Duckworth *et al.*, 1986; 124). The mass spectrum can be scanned by varying systematically the magnitude of the DC and oscillating voltages in such a way that their ratio to each other does not change. At any particular value of these voltages, only ions of a certain mass to charge ratio will pass through the quadrupole and be detected (again usually by a simple Faraday cup) at the other end – all other ions will be deflected out of the system and lost. The mass spectrum can be scanned very rapidly – 0 to 800 mass units in a matter of seconds – and the normal mode of operation is to scan many times and accumulate an average spectrum in an attached computer. The mass resolution of the machine depends on the length of the rods and the frequency of oscillation, but is generally less than that achievable by the magnetic sector type instruments. Nevertheless, it is a rapid, relatively cheap, and sensitive detector, and has been increasingly used in ICP and chromatography applications, as well as for applications such as secondary ion mass spectrometry and other surface analytical techniques.

ICP-MS instrumentation of this type has been used in environmental applications to measure isotope ratios of heavy elements such as Pb in body fluids, plant material, and dust samples to determine the source of metal contamination (*e.g.*, Hamester *et al.*, 1994). It is generally accepted that ICP-MS measurements of lead isotope ratios are sufficiently precise to detect differences between sources of pollution, but interlaboratory studies of ICP-MS measurements on a range of environmental samples suggests that realistic precisions are 0.3% for $^{206}Pb/^{207}Pb$, 0.8% for $^{206}Pb/^{204}Pb$, and 1.4% for $^{208}Pb/^{204}Pb$ (Furuta, 1991), which are roughly an order of magnitude worse than can be achieved with 'conventional' lead isotope measurements made by TIMS (see Chapter 9). This is, however, likely to change in the near future, when a new generation of ICP-MS machines becomes available, which have an ICP source linked to a double focusing multiple collector 'conventional' mass spectrometer, giving precisions even better than those achievable by TIMS (Walder and Freedman, 1992; Walder and Furuta, 1993). The first of these machines is now on the market, and is certain to have a tremendous impact over the next few years.

CHROMATOGRAPHIC TECHNIQUES

The term chromatography is used to encompass a wide range of related techniques which enable the separation of the components in a mixture as a result of their distribution between two phases – one *stationary* and one *mobile*. The most widely used chromatographic technique in archaeological applications during the past two decades has been *gas chromatography* (GC), although examples of *liquid chromatography* (LC) are becoming more common. When coupled to a mass spectrometer (MS), combined GC-MS and LC-MS offer powerful tools for analysis of a wide range of biomarkers remnant in the archaeological record (see Table 2.1, page 72).

Superficially, chromatography can be described in terms of placing blotting paper into a bath of ink. The liquid moves up the paper by means of capillary action. After a time, closer observation shows that a separation has occurred – there is a 'front' which marks how far a clear component has moved, and a dark front, lagging behind it, showing the lesser movement of the coloured matter in the ink. This observation shows that the clear liquid – water – has moved slightly faster through the paper than the coloured component (whatever is giving the ink its dark colour). This is the basis of a separation technique called *paper chromatography*, which still has its uses today. In this simple example, the *mobile phase* is the aqueous ink solution and the *stationary phase* is the

paper. The separation is a result of the differing interaction between the various components in the mobile phase and the stationary phase.

The mobile phase can be either a gas or a liquid, and the stationary phase can be either a solid or a liquid, giving rise to the two broad families of instrumental chromatographic techniques, namely GC and LC respectively. In *gas chromatography*, a gaseous mobile phase containing the volatilized mixture to be separated is passed through a column. Gas chromatography is suitable for the separation of thermally stable and volatile organic and inorganic compounds. The volatility of solutes can be promoted by various chemical derivatization procedures. Although GC has wide application and is very sensitive, only about 20% of known compounds can be analysed – the rest are either insufficiently volatile and cannot be made to travel in the gaseous state, or are thermally unstable and decompose under the conditions employed (Willard *et al.*, 1988; Chapters 17–20). Liquid chromatography is not limited by these considerations and can be used for a wide range of materials. The most common form of liquid chromatography is HPLC (*high performance* or *high pressure liquid chromatography*) whereby the liquid mobile phase is forced under pressure through a column containing a solid packing (*liquid-solid* or *adsorption chromatography*) or a solid packing coated with a stationary liquid phase (*liquid-liquid chromatography*).

Both GC and LC operate on essentially the same principle, which can be simply described (for further detail on the theory of chromatographic separation see Braithwaite and Smith, 1985; 11–23). The mobile phase is made up of a *solvent* (either a gas or a liquid) which carries the mixture through the system, and a *solute*, which is the mixture to be separated. The molecules of the solute at any particular time are distributed between those carried by the mobile phase and those in the stationary phase. At any stage there is an equilibrium established between the concentration of the solute distributed in each phase. The competition between the two phases for the solute molecules depends on their physical properties and affinity for the stationary phase (see Braithwaite and Smith, 1985; 13–14 for a discussion of the molecular interactions influencing retention). In column chromatography it is usual to define a *partition* or *capacity ratio* (k') which is a measure of how long a time the molecules of a given species spend in the stationary phase relative to their time in the mobile phase. If a component has a low partition ratio it will pass through the column relatively quickly. At the extreme, if k' is zero (*i.e.*, the compound spends no time in the stationary phase) it will pass through the column at the same rate as the mobile phase. If k' is very high, the 'compound' spends a long time in the stationary phase, and may never emerge from the column! The time taken for each

molecular species to emerge or elute from a column is termed the *retention time.* The *relative retention* (α) of two compounds is a measure of how well they can be separated (*i.e.*, the resolution of the system), and is the ratio of the two *partition ratios.* For any compound, the value of k' depends on the combination of stationary/mobile phase chosen, and on the temperature of the system. The retention time depends on these and other factors such as the length of the column and the flow rate of the mobile phase. Complex mixtures can be separated by careful selection of these variables.

A gas chromatograph system requires a means of introducing a sample onto the column where separation is achieved. The column feeds into a detector which acts as a monitor of the column effluent. A basic GC system is shown in Figure 2.17. The mobile phase is an inert gas (also known as the carrier gas) such as helium, hydrogen, or nitrogen; the rate of flow of which must be accurately controlled for the purposes of reproducibility. A number of injection systems are available (Tipler, 1993; 20–30), including automated samplers which give more reproducible injections than can be achieved by a human operator. The column is housed in an oven to allow the column temperature to be controlled. In GC, two basic types of column are employed; the *packed column* and the *wall-coated open tubular* (WCOT) or *capillary column.* In the majority of cases the latter have replaced the former. Packed columns are generally made of stainless steel or glass tubing, with an internal diameter of 2–6 mm, and a length of up to 5 m. They are packed with a fine grained inert support material which is coated with the stationary

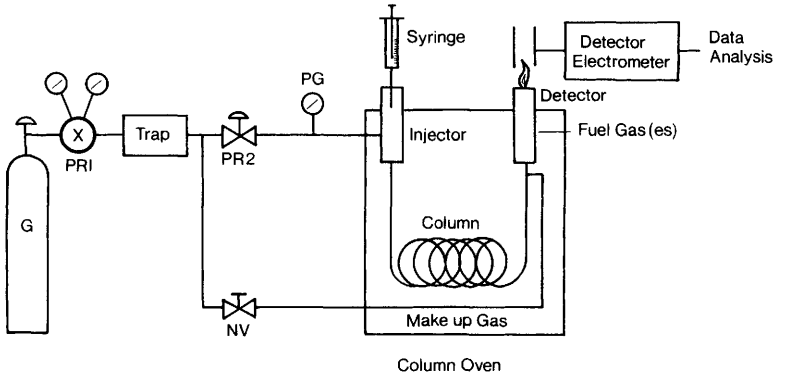


Figure 2.17 *Schematic diagram of a gas chromatography system*
Adapted from Bartle, 1993; Figure 1, by permission of Oxford University Press)

liquid phase. A capillary column has an internal diameter of less than 1 mm, and may have a length of 10–30 m although longer columns are possible. Normally the stationary phase is coated and bonded to the internal wall of the column. The column is housed in an oven which, through the use of incrementing temperatures, means that the least volatile components in a mixture can be eluted more rapidly. The majority of archaeological applications have made use of capillary columns employing low polarity, bonded stationary phases.

The most common detector used in GC is the *flame ionization detector* (FID), in which solutes separated during the chromatographic process are burned in a hydrogen-air flame. Designs vary, but, in the simplest form, just above the tip of the flame are a pair of metal plates which carry a potential difference of around 400 V. The current flowing across the flame between the electrodes is measured. Normally it is very low, but as charged particles enter the gap between the plates from the flame, the resistance between the plates drops and the current flow increases markedly. If only the carrier gas is present in the flame, the current flow will be a fixed baseline value, but as the separated components elute from the column and into the flame, the current will rise and fall as each separated component passes through. The peaks are recorded on a chart recorder, or digitally in a computer, and the output takes the form of a plot of peak height as a function of time, with the signal produced proportional to the amount of each component in the mixture. The FID is very sensitive, responds to the vast majority of organic compounds and retains good linearity over a wide sample concentration range. Other detectors (such as *thermionic emission*, or *thermal conductivity*) can be used for specific application. A more complete description of GC can be found in the recent text edited by Baugh (1993).

Equipment for HPLC is slightly more complicated (see Figure 2.18), due to the need to handle a liquid mobile phase. This requires the presence of one or more reservoirs for the solvent which forms the mobile phase – many systems have the facility for changing the composition of the mobile phase as the run progresses, by mixing solvents from two or more reservoirs (*gradient elution*). The mobile phase is delivered to the column by a pumping system, which can give a high pressure liquid and control the flow rate without undue variation in the pressure or flow. The most common type of pump is a reciprocating piston, often coupled with a pulse damper to ensure smooth flow. The sample (dissolved, if possible, in the solvent which comprises the mobile phase) is injected into a microsampling injector valve, which holds the sample in an injection loop, and releases a fixed volume (usually either 10 or 20 μl) into the liquid flow at the top of the column. Again, autosamplers are used to

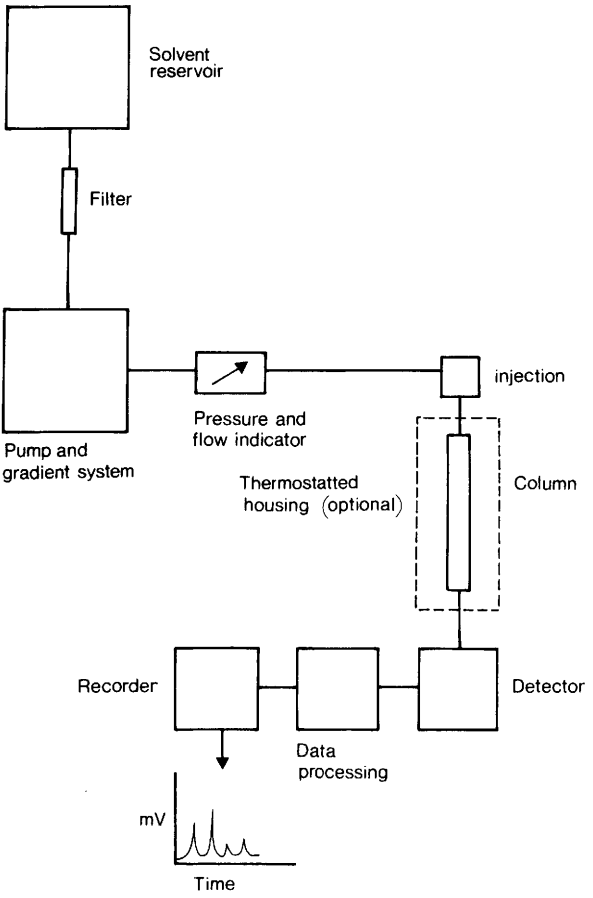


Figure 2.18 *Schematic diagram of an HPLC system*
(Reproduced from *High Performance Liquid Chromatography*, 2nd Edition (S. Lindsay, 1992) with the permission of the University of Greenwich)

relieve the tedium for the operator, and ensure better reproducibility of injection. Columns for HPLC need to be of sturdy construction to withstand the high pressures involved (up to several hundred atmospheres). Typically they are made of stainless steel, or glass-lined metal for chemical inertness, and are straight, with a length of 10–30 cm depending on application. They are usually fitted with a guard column or an in-line filter to prevent unwanted particles entering the column, and are housed in an oven to allow temperature control.

Once the liquid sample has passed through the column (and hopefully

been separated) it is fed directly into a detector which detects and quantifies the eluting compounds. Several types of detector are available, depending on the application, but the two most common types are *ultraviolet/visible spectrophotometers* and *fluorescence* detectors. The UV/visible (UV/vis) type of detector comes in a variety of forms, but again the most common type is a fixed wavelength detector. The mobile phase plus separated compound passes through a detector cell (typically of 2–8 μl capacity) through which is focused radiation of the selected wavelength, and the mobile phase only passes through a reference cell. The absorbance at this wavelength is measured as the difference between the absorbance of the two cells, as determined by a photometer. Many compounds absorb around the 254 nm wavelength, which is one of the wavelengths emitted from a medium pressure mercury vapour lamp, and so this is usually the wavelength chosen. A fluorescence detector is similar to a UV/vis detector, but the irradiating light illuminates the sample cell and the resulting fluorescence is detected at right angles to this. Since not all compounds exhibit fluorescence to UV/vis radiation, a fluorophor may have to be added to the sample by derivatization prior to injection. The fluorescence detector is generally considerably more sensitive (perhaps by two or three orders of magnitude), but the need to sensitize the sample to the fluorescing radiation may make the sample preparation more complex.

Chromatographic techniques have developed rapidly over the last few years, resulting in a wide range of 'hyphenated techniques'. For GC, one of the main drawbacks is the requirement that the compounds to be separated should be volatile but stable. This can be overcome in certain cases by using a *pyrolysis injection system* (Willard *et al.*, 1988; 453), in which a solid sample (*e.g.*, rubber, resin, soil, or coal and so on) is heated rapidly (by Curie point or ohmic heating) to cleave certain bonds producing lower molecular weight fragments which then enter the chromatographic system. This gives a technique called *pyrolysis-GC* (Py-GC). Commonly, the detector of either a GC (or less frequently an HPLC) is interfaced to a quadrupole mass spectrometer. The combination of a separation technique coupled with a versatile detection system, which results in the generation of a mass spectrum of each component separated, allows much greater opportunities of structure identification. Increasingly, combined GC-MS is finding applications in archaeology (see Table 2.1).

In view of the fact that the major emphasis of this book, largely for historical reasons, is the analysis of inorganic materials, we have assembled a table summarizing the applications of chromatographic techniques to archaeology (Table 2.1) and associated references. Although this is undoubtedly not exaustive, it gives a good perspective

Table 2.1 *Some recent combined chromatographic/spectroscopic (GC-MS, LC-MS) applications in archaeology (see also Mills and White, 1994)*

Sample	*Principal associations*	*References*
Food constituents	Pottery vessels, charred food deposits	Heron and Evershed (1993); Evershed *et al.* (1992); Oudemans and Boon (1991)
Resins, tars, and waxes	Pottery vessels, stone tools	Beck *et al.* (1994); Heron *et al.* (1994); Mills and White (1989)
Dyestuffs	Textiles, pottery vessels	Wouters and Verhecken (1989)
Psychoactive substances	Pottery vessels and other receptacles	Hocart *et al.* (1993); Torres *et al.* (1991); Hurst *et al.* (1989)
Asphalt/Bitumen	Various artefacts, mummified bodies	Connan *et al.* (1992)
Lipids preserved in human soft tissue and bone	Bog bodies, mummified bodies	Evershed *et al.* (1995) Evershed and Connolly (1994); Gülaçar *et al.* (1990)
Biomarkers in soils	Anthropogenic soil horizons	Bethell *et al.* (1994); Pepe and Dizabo (1990)

on the range of material and information now accessible to modern chromatographic methods.

OTHER TECHNIQUES

A wide range of other methods from analytical chemistry have been applied to archaeological samples, but space precludes detailed descriptions of them all. Some, such as *X*-ray photoelectron spectroscopy, have only been employed sporadically because of the specialized nature of the technique. Others are increasing in application as their archaeological potential is explored. Foremost of these techniques is probably infrared spectroscopy, which has been extensively used to characterize the sources of European amber (Beck, 1986; Beck and Shennan, 1991), and has recently been applied to the study of the differential preservation of bone (Weiner *et al.*, 1993). *Infrared microspectroscopy* (the use of a microscope with IR capability) has found application for the identification of mineralized fibres from archaeological contexts (Gillard *et al.*, 1994). As

shown in Figure A1.2 in Appendix 1 (p. 353), the infrared region of the electromagnetic spectrum carries information about the vibrational states of the chemical bonds. Thus, peaks in IR spectra are often labelled 'C–H stretch', 'C–H bend', *etc.* (for a typical table of IR peak identifications, see Ewing, 1985; Table 4-2). Thus it gives details of the types of chemical bonds present in a sample, and detailed study of the spectra can reveal details about the chemical environment of particular functional groups. IR should be regarded as a 'chemical fingerprinting' technique – it does not provide precise chemical characterization of a sample, and quite different compounds may have very similar IR spectra. Further details of IR are given in most textbooks on analytical chemistry (*e.g.*, Ewing, 1985; Chapter 4; Willard *et al.*, 1988; Chapter 11): undoubtedly, because of its relative cheapness and versatility, IR will become more widely used in archaeology.

Another class of methods which has had some application are *resonance techniques* (*e.g.*, Ewing, 1985; Chapter 13). These are based on another aspect of the interaction between matter and electromagnetic radiation, in which a sample is subjected simultaneously to two magnetic fields – one stationary, the other varying at a known frequency (usually a radiofrequency, or RF). Under these conditions, sub-atomic particles in the sample are aligned in a particular orientation by the steady field, and, as the frequency of the RF is varied, a resonance condition is met when the 'spin' of the particle is caused to 'flip' causing an absorption in the frequency spectrum. *Electron spin resonance* (ESR) is now best known archaeologically for its application as a dating technique (Aitken, 1990; Schwarcz and Grün, 1992), but it has been used for a variety of other applications, such as the elucidation of the thermal history of flint (Robins *et al.*, 1978). *Nuclear magnetic resonance* (NMR) has also been used for the characterization of amber (Lambert *et al.*, 1988) and coal-like material (Lambert *et al.*, 1992). Although the instrumentation for resonance measurement is relatively expensive, it is likely that selected archaeological problems, particularly those involving the solution of structural organic problems, may well benefit from further applications of such approaches. Further details on NMR can be found in Willard *et al.* (1988; Chapter 15), and Ewing (1985; 255).

One final area worthy of brief mention because of its great potential in archaeology is the thermal analysis of solid materials (*e.g.*, Ewing, 1985; Chapter 23) A range of techniques are included under this heading, but the basic principle is the quantitative measurement of the weight change in a solid sample under controlled heating conditions. In its simplest form (*thermogravimetry*, or TGA) the weight of the sample is recorded as the temperature is raised, thus recording the loss of components such as

water and CO_2, allowing the composition to be elucidated. In *differential thermal analysis* (DTA) the heat evolved or absorbed by the sample as it is heated is measured by comparison with an inert reference material, which allows phase changes to be monitored as well as chemical changes. The related technique of *differential scanning calorimetry* (DSC) allows more quantitative estimates to be made of the thermodynamic heats of transition. Thermal analytical techniques are widely used in materials science to identify and characterize mineral species, polymers, *etc.*, and to study the high temperature thermodynamic properties of compounds. An example of the use of TGA on archaeological material can be found in the work of Hunter *et al.* (1993), in which ESR, TGA, Raman spectroscopy, IR, and *X*-radiography were used in combination with the more usual methods to characterize the various black lithic materials (jet, shale, lignite, *etc.*) used in antiquity. Clearly, archaeological chemists are becoming conversant with a much wider range of techniques than has been the case hitherto to characterize and study a wide range of materials from the past.

REFERENCES

Aitken, M.J. (1990). *Science-based Dating in Archaeology*. Longman, London.

Awadallah, R.M., Sherif, M.K., Amrallah, A.H. and Grass, F. (1986). Determination of trace elements of some Egyptian crops by instrumental neutron activation analysis, inductively coupled plasma atomic emission spectrometric and flameless atomic absorption spectrophotometric analysis. *Journal of Radioanalytical and Nuclear Chemistry Articles* **98** 235–246.

Barber, D.J. and Freestone, I.C. (1990). An investigation of the origin of the colour of the Lycurgus Cup by analytical transmission electron microscopy. *Archaeometry* **32** 33–45.

Bartle, K.D. (1993). Introduction to the theory of chromatographic separations with reference to gas chromatography. In *Gas Chromatography: A Practical Approach*, ed. Baugh, P.J., Oxford University Press, Oxford, pp. 1–14.

Baugh, P.J. (ed.) (1993). *Gas Chromatography: A Practical Approach*. Oxford University Press, Oxford.

Beck, C. (1986). Spectroscopic investigations of amber. *Applied Spectroscopy Reviews* **22** 57–110.

Beck, C.W. and Shennan, S. (1991). *Amber in Prehistoric Britain*. Oxbow, Oxford.

Beck, C.W., Stewart, D.R. and Stout, E.C. (1994). Appendix D: Analysis of Naval stores from the Late Roman ship. In *Deep Water Archaeology: A*

Late Roman Ship from Carthage and an Ancient Trade Route near Skerki Bank off Northwest Sicily, eds. McCann, A.M. and Freed, J., *Journal of Roman Archaeology Supplementary Series* **13** pp. 109–121.

Bethell, P.H., Goad, L.J., Evershed, R.P. and Ottaway, J. (1994). The study of molecular markers of human activity: The use of coprostanol in the soil as an indicator of human faecal material. *Journal of Archaeological Science* **21** 619–632.

Bettinelli, M., Baroni, U., Pastorelli, N. and Bizzarri, G. (1992). ICP-AES, GFAAS, XRF and NAA coal fly-ash analysis – comparison of different analytical techniques. In *Elemental Analysis of Coal and Its By-Products,* Proceedings of 2nd International Conference, Barren River Resort, Kentucky, pp. 372–394.

Beynon, J.H. and Brenton, A.G. (1982). *An Introduction to Mass Spectrometry*. University of Wales Press, Cardiff.

Brady, J.E. (1990). *General Chemistry*. John Wiley, New York (5th edn.).

Braithwaite, A. and Smith, F.J. (1985). *Chromatographic Methods*. Chapman and Hall, London.

Briggs, D. and Seah, M.P (eds.) (1990). *Practical Surface Analysis. Vol. 1. Auger and X-Ray Photoelectron Spectroscopy*. John Wiley, Chichester.

Britton, D. and Richards, E.E. (1969). Optical emission spectroscopy and the study of metallurgy in the European Bronze Age. In *Science in Archaeology*, eds. Brothwell, D. and Higgs, E., Thames and Hudson, London, pp. 603–613 (2nd edn.).

Caro, D.E., McDonell, J.A. and Spicer, B.M. (1978). *Modern Physics*. Edward Arnold, London (3rd edn.).

Christian, G. (1994). *Analytical Chemistry*. John Wiley, New York (5th edn.).

Connan, J., Nissenbaum, A. and Dessort, D. (1992). Molecular archaeology: Export of Dead Sea asphalt to Canaan and Egypt in the Chalcolithic-Early Bronze Age (4th–3rd Millennium BC). *Geochimica et Cosmochimica Acta* **56** 2743–2759.

Cotton, F.A. and Wilkinson, G. (1988). *Advanced Inorganic Chemistry*, Wiley-Interscience, New York (5th edn.).

Cotton, F.A., Wilkinson, G. and Gaus, P.L. (1995). *Basic Inorganic Chemistry*. John Wiley, New York (3rd edn.).

Cowell, M. and La Niece, S. (1991). Metalwork: artifice and artistry. In *Science and the Past*, ed. Bowman, S., British Museum Publications, London, pp. 74–98.

Cox, G.A. and Pollard, A.M. (1977). *X*-ray fluorescence of ancient glass: the importance of sample preparation. *Archaeometry* **19** 45–54.

Dawson, P.T., Heavens, O.S. and Pollard, A.M. (1978). Glass surface analysis by Auger electron spectroscopy. *Journal of Physics C* **11** 2183–2193.

Duckworth, H.E., Barber, R.C. and Venkatasubramanian, V.S. (1986). *Mass Spectroscopy*. Cambridge University Press, Cambridge (2nd edn.).

Dunham, A.C. and Wilkinson, F.C.F. (1978). Accuracy, precision and detection limits of energy-dispersive electron-microprobe analyses of silicates. *X-Ray Spectrometry* **7** 50–56.

Durrant, S.F. and Ward, N.I. (1993). Rapid multielemental analysis of Chinese reference soils by laser ablation inductively coupled plasma mass spectrometry. *Fresenius Journal of Analytical Chemistry* **345** 512–517.

Emeleus, V.M. (1958). The technique of neutron activation analysis as applied to trace element determination in pottery and coins. *Archaeometry* **1** 6–15.

Evershed, R.P. and Connolly, R.C. (1994). Post-mortem transformations of sterols in bog body tissues. *Journal of Archaeological Science* **21** 577–583.

Evershed, R.P., Heron, C., Charters, S. and Goad, L.J. (1992). The survival of food residues: New methods of analysis, interpretation and application. In *New Developments in Archaeological Science*, ed. Pollard, A.M., Proceedings of the British Academy **77**, Oxford University Press, Oxford, pp. 187–208.

Evershed, R.P., Turner-Walker, G., Hedges, R.E.M., Tuross, N. and Leyden, A. (1995). Preliminary results for the analysis of lipids in ancient bone. *Journal of Archaeological Science* **22** 277–290.

Ewing, G.W. (1985). *Instrumental Methods of Chemical Analysis*. McGraw-Hill, New York (5th edn.).

Faure, G. (1986). *Principles of Isotope Geology*. John Wiley, Chichester (2nd edn.).

Fleming, S.J. and Swann, C.P. (1986). PIXE spectrometry as an archaeometric tool. *Nuclear Instruments and Methods in Physics Research* **A242** 626–631.

Furuta, N. (1991). Interlaboratory comparison study on lead isotope ratios determined by inductively coupled plasma mass spectrometry. *Analytical Sciences* **7** 823–826.

Gillard, R.D., Hardman, S.M., Thomas, R.G. and Watkinson, D.E. (1994). The mineralization of fibres in burial environments. *Studies in Conservation* **39** 132–140.

Gillies, K.J.S. and Urch, D.S. (1983). Spectroscopic studies of iron and carbon in black surfaced wares. *Archaeometry* **25** 29–44.

Glascock, M.D. (1994). Nuclear reaction chemical analysis: prompt and delayed measurements. In *Chemical Analysis by Nuclear Methods*, ed. Alfassi, Z.B., John Wiley, Chichester, pp. 75–99.

Gülaçar, F.O., Susini, A. and Klohn, M. (1990). Preservation and post-

mortem transformation of lipids in samples from a 4000-year-old Nubian mummy. *Journal of Archaeological Science* **17** 691–705.

Hall, E.T. (1961). Surface enrichment of buried metals. *Archaeometry* **4** 62–66.

Hall, E.T., Schweizer, F. and Toller, P.A. (1973). *X*-ray fluorescence analysis of museum objects: a new instrument. *Archaeometry* **15** 53–78.

Hamester, M., Stechmann, H., Steiger M. and Dannecker, W. (1994). The origin of lead in urban aerosols – a lead isotopic ratio study. *Science of the Total Environment* **146/7** 321–323.

Hatcher, H., Tite, M.S. and Walsh, J.N. (1995). A comparison of inductively-coupled plasma emission spectrometry and atomic absorption spectrometry analysis on standard reference silicate materials and ceramics. *Archaeometry* **37** 83–94.

Heron, C. and Evershed, R.P. (1993). The analysis of organic residues and the study of pottery use. In *Archaeological Method and Theory,* Volume 5, ed. Schiffer, M.B., University of Arizona Press, Arizona, pp. 247–286.

Heron, C., Nemcek, N., Bonfield, K.M., Dixon, J. and Ottaway, B.S. (1994). The chemistry of Neolithic beeswax. *Naturwissenschaften* **81** 266–269.

Hocart, C.H., Fankhauser, B. and Buckle, D.W. (1993). Chemical archaeology of kava, a potent brew. *Rapid Communications in Mass Spectrometry* **7** 219–224.

Hurst, J., Martin, R. Jr., Tarka, S. Jr., and Hall, G. (1989). Authentication of cocoa in Maya vessels using high-performance liquid chromatographic techniques. *Journal of Chromatography* **466** 279–289.

Hughes, M.J., Cowell, M.R. and Craddock, P.T. (1976). Atomic absorption techniques in archaeology. *Archaeometry* **18** 19–37.

Hughes, M.J., Cowell, M.R. and Hook, D.R. (1991). *Neutron Activation and Plasma Emission Spectrometric Analysis in Archaeology.* British Museum Occasional Paper 82, London.

Hunter, F.J., McDonnell, J.G., Pollard, A.M., Morris, C.R. and Rowlands, C.C. (1993). The scientific identification of archaeological jet-like artefacts. *Archaeometry* **35** 69–89.

Jackson, L.L., Baedecker, P.A., Fries, T.L. and Lamothe, P.J. (1993). Geological and inorganic materials. *Analytical Chemistry* **65** 13R–28R.

Jarvis, K.E., Gray, A.L. and Houk, R.S. (1992). *Handbook of Inductively Coupled Plasma Mass Spectrometry.* Blackie, Glasgow.

Jenkins, F.A. and White, H.E. (1976). *Fundamentals of Optics.* McGraw-Hill, New York (4th edn.).

Jenkins, R. (1974). *An Introduction to X-Ray Spectrometry.* John Wiley, Chichester.

Jenkins, R. (1988). X-*ray Fluorescence Spectrometry*. Wiley-Interscience, Chichester.

Johansson, S.A.E. and Campbell, J.L. (1988). *PIXE: A Novel Technique for Elemental Analysis*. John Wiley, Chichester.

Jones, R. E. (1986). *Greek and Cypriot Pottery*. British School at Athens Fitch Laboratory Occasional Paper 1, Athens.

Joy, D.C., Romig, A.D. Jr. and Goldstein, J.I. (1986). *Principles of Analytical Electron Microscopy*. Plenum, New York.

Kraay, C.M. (1958). Gold and copper traces in early Greek silver. *Archaeometry* **1** 1–5.

Lambert, J.B., Beck, C.W. and Frye, J.S. (1988). Analysis of European amber by carbon-13 nuclear magnetic resonance spectroscopy. *Archaeometry* **30** 248–263.

Lambert, J.B., Frye, J.S. and Jurkiewicz, A. (1992). The provenance and coal rank of jet by carbon-13 nuclear magnetic resonance spectroscopy. *Archaeometry* **34** 121–128.

Lindsay, S. (1992). *High Performance Liquid Chromatography*. Wiley, Chichester (2nd edn.).

Mills, J.S. and White, R. (1989). The identity of the resins from the Late Bronze Age shipwreck at Ulu Burun (Kaş). *Archaeometry* **31** 37–44.

Mills, J.S. and White, R. (1994). *The Organic Chemistry of Museum Objects*. Butterworth-Heinemann, London (2nd edn.).

Oudemans, T.F.M. and Boon, J.J. (1991). Molecular archaeology: analysis of charred (food) remains from prehistoric pottery by pyrolysis-gas chromatography/mass spectrometry. *Journal of Analytical and Applied Pyrolysis* **20** 197–227.

Pepe, C. and Dizabo, P. (1990). Étude d'une fosse du 13ème Siècle par les marqueurs biogeochimiques: Chantier archéologique du Louvre (Paris). *Revue d'Archéometrie* **13** 1–11.

Perlman, I. and Asaro, F. (1969). Pottery analysis by neutron activation analysis. *Archaeometry* **11** 21–52.

Physical Science Study Committee (1960). *Physics*. D.C. Heath, Boston.

Potts, P.J., Webb, P.C. and Watson, J.S. (1985). Energy-dispersive X-ray fluorescence analysis of silicate rocks: comparisons with wavelength-dispersive performance. *Analyst* **110** 507–513.

Reed, S.J.B. (1993). *Electron Microprobe Analysis*. Cambridge University Press, Cambridge (2nd edn.).

Robins, G.V., Seeley, N.J., McNeil, D.A.C. and Symons, M.C.R. (1978). Identification of ancient heat treatment in flint artifacts by ESR spectroscopy. *Nature* **276** 703–704.

Rouchaud, J.D., Boisseau, N. and Federoff, M. (1993). Multielement analysis of aluminium by NAA and ICP/AES. *Journal of Radioanalytical and Nuclear Chemistry Letters* **175** 25–31.

Sayre, E.V. and Dodson, R.W. (1957). Neutron activation study of Mediterranean potsherds. *American Journal of Archaeology* **61** 35–41.

Schwarcz, H.P. and Grün, R. (1992). Electron spin resonance (ESR) dating of the origin of modern man. *Philosophical Transactions of the Royal Society of London B* **337** 145–148.

Sibilia, J.P. (1988). *A Guide to Materials Characterization and Chemical Analysis.* VCH, New York.

Skoog, D.A. and Leary, J.J. (1992). *Principles of Instrumental Analysis.* Saunders College Publishing, Fort Worth (4th edn.).

Slavin, W. (1992). A comparison of atomic spectroscopic analytical techniques. *Spectroscopy International* **4** 22–27.

Thompson, M. and Walsh, J.N. (1989). *A Handbook of ICP Spectrometry.* Blackie, Glasgow (2nd edn.).

Tipler, A. (1993). Gas chromatography instrumentation, operation, and experimental considerations. In *Gas Chromatography: A Practical Approach*, ed. Baugh, P.J., Oxford University Press, Oxford, pp. 15–70.

Torres, C.M., Repke, D.B., Chan, K., McKenna, D., Llagtostera, A. and Schultes, R.E. (1991). Snuff powders from Pre-Hispanic San Pedro de Atacama: Chemical and contextual analysis. *Current Anthropology* **32** 639–649.

van der Merwe, N.J. (1992). Light stable isotopes and the reconstruction of prehistoric diets. In *New Developments in Archaeological Science*, ed. Pollard, A.M., Proceedings of the British Academy 77, Oxford University Press, Oxford, pp. 247–264.

van Loon, J.C. (1980). *Analytical Atomic Absorption Spectroscopy: Selected Methods.* Academic Press, New York.

van Loon, J. C. (1985). *Selected Methods of Trace Metal Analysis: Biological and Environmental Samples.* John Wiley, Chichester.

Verità, M., Basso, R., Wypyski, M.T. and Koestler, R.J. (1994). *X*-ray microanalysis of ancient glassy materials: a comparative study of wavelength dispersive and energy dispersive techniques. *Archaeometry* **36** 241–251.

Walder, A.J. and Freedman, P.A. (1992). Isotopic ratio measurement using a double focusing magnetic sector mass analyser with an inductively coupled plasma as an ion source. *Journal of Analytical Atomic Spectrometry* **7** 571–575.

Walder, A.J. and Furuta, N. (1993). High-precision lead isotope ratio measurements by inductively coupled plasma multiple collector mass spectrometry. *Analytical Sciences* **9** 675–680.

Walsh, J.N. and Howie, R.A. (1986). Recent developments in analytical methods: uses of inductively coupled plasma source spectrometry in applied geology and geochemistry. *Applied Geochemistry* **1** 161–171.

Ward, N.I., Abou-Shakra, F.R. and Durrant, S.F. (1990). Trace element content of biological-materials – a comparison of NAA and ICP-MS analysis. *Biological Trace Element Research* **26–7** 177–187.

Weiner, S., Goldberg, P. and Bar-Yosef, O. (1993). Bone preservation in Kebara Cave, Israel, using on-site Fourier transform infrared spectrometry. *Journal of Archaeological Science* **20** 613–627.

Whiffen, D.H. (1972). *Spectroscopy*. Longman, London (2nd edn.).

Willard, H.H., Merritt, L.L. Jr., Dean, J.A. and Settle, F.A. Jr. (1988). *Instrumental Methods of Analysis*. Wadsworth, Belmont, California (7th edn.).

Woldseth, R. (1973). *X-Ray Energy Spectrometry*. Kevex Corporation, Burlingame, California.

Wouters, J. and Verhecken, A. (1989). The coccid insect dyes: HPLC and computerized diode-array analysis of dyed yarns. *Studies in Conservation* **34** 189–200.

Chapter 3

Obsidian Characterization in the Eastern Mediterranean

INTRODUCTION

The best known explosive eruption in the Aegean took place on the island of Santorini (Thera) sometime during the Second Millennium BC. It is also the youngest. Volcanic activity of much greater antiquity in other areas of the Aegean and beyond gave rise to a valuable artefactual material known as *obsidian.* Obsidian is a volcanic glass, formed when lava is cooled rapidly, often at the margins of a flow. It is normally shiny in appearance, and dark in colour (black or grey), but may be colorless, red, green, or brown, depending on the composition and circumstances of formation. The characteristic conchoidal fracture of obsidian ensured its early use in the manufacture of stone tools, particularly for sharp flakes and blades. Obsidian was also valued as a decorative item – small vessels, statuettes, and mirrors could be formed from nodules of the glassy rock. Obsidian has been used in the Aegean for more than 10 000 years. When seashells modelled from a form of obsidian with distinctive white inclusions, or *spherulites*, were found at the Late Bronze Age Minoan palace of Knossos on the island of Crete, Arthur Evans, excavator of the site, considered the obsidian to have come from the western Mediterranean island of Lipari where natural deposits of similar spotted obsidian are known to occur (Evans, 1921; 87; 412). He referred to these examples as 'Liparite'. Scientific study carried out several decades later demonstrated that the obsidian came from a source in the Aegean. This chapter examines how suggestions regarding source attribution of archaeological materials might be confirmed on the basis of chemical composition. Furthermore, models developed to explain the patterns underlying the distribution of obsidian away from the natural source are reviewed briefly.

Chapter 1 identified source attribution of archaeological materials as one of the most important areas of scientific analysis. This chapter focuses on one of the most successful applications of archaeological chemistry: the provenancing of obsidian (see also the recent review by Williams-Thorpe, 1995). With hindsight, it is possible to conclude that it was successful for two reasons – the relative ease by which compositional data can be used to characterize obsidian, and the fact that the vast majority of studies have been embedded within the clear archaeological objective of establishing exchange mechanisms. The ability to source this durable volcanic glass has provided archaeologists with a material record of cultural contact over wide areas, and, indirectly, with the earliest evidence for seafaring (and therefore the existence of boats), since the exploitation of island sources would have necessitated marine travel.

Obsidian became widely used in parts of the Aegean in the Neolithic (beginning around 7 000 BC) and Early Bronze Age (beginning around 2 500 BC). By the 6th Millennium BC, obsidian could be found at sites from Crete to Macedonia. Every Early Neolithic site so far excavated in southern Greece has yielded at least some obsidian. Systematic and intensive field surveys of the modern land surface in these regions, a powerful tool for assessing the development of settlement patterns and land use, recover finds of obsidian (in addition to large quantities of pottery) scattered across the landscape. These finds are indicators of prehistoric activity which, taken together with finds of pottery, enable the location of settlements and other activity areas from the surface record alone. The earliest use of obsidian has been traced back to the Upper Palaeolithic levels at Franchthi Cave in the Argolid, Greece (Perlès, 1987) prompting the discussions noted above concerning the earliest seafaring and deep sea fishing. In the later Mesolithic, significant quantities of large bluefin tunny (*Thunnus thynnus*) bones appear in the faunal assemblage. According to Perlès (1990a; 46–47), Melian obsidian becomes abundant soon after. The link between obsidian procurement and fishing expeditions for tunny has been disputed by some (*e.g.*, Bloedow, 1987). It is generally assumed that the use of obsidian declined (around 2 300 BC) with the increasing availability of metals, although there is evidence for continued use of obsidian throughout the Bronze Age. As a footnote to the history of human use of obsidian, very recently surgeons have been investigating the use of obsidian blades instead of steel for surgical operations – it is said that obsidian gives a cleaner wound which heals faster (Disa *et al.*, 1993). At his own request, the archaeologist and lithic specialist François Bordes underwent major surgery using obsidian blades.

ORIGIN AND FORMATION OF OBSIDIAN

Igneous rocks (*i.e.*, rocks of magmatic origin) are classified chemically according to their percentage weight of silica (SiO_2). Those with more than 66% are generally termed *acidic*, between 66% and 52% *intermediate*, between 52% and 45% *basic*, and less than 45% *ultrabasic* (see Read, 1970; 204). They may additionally be classified according to their grain size, loosely falling into two categories, the *coarse-grained* (which includes granites) and *fine-grained.* Since many igneous rocks are rich in quartz (crystalline SiO_2), they are also traditionally classified according to their relative proportions of three families of minerals quartz (Q), alkali-feldspar (A), and plagioclase (P). These can be plotted on the so-called *QAP triangular diagrams* (Figure 3.1), which shows how the relative abundance of these three minerals relate approximately to the chemical classifications quoted above. These QAP diagrams are divided into sub-regions within the acidic and intermediate categories. In coarse-grained rocks, these include *granite* (1), *adamellite* (2), *granodiorite* (3), *syenite* (4), *diorite* (5), and *gabbro* (6), according to their chemical composition. Similar chemical categories for fine-grained rocks would be labelled *rhyolite* (1), *rhyo-dacite* (2), *dacite* (3), *trachyte* (4), *andesite* (5), and *basalt* (6).

Obsidian is an acidic volcanic glass formed when silica-rich (typically 65–75% SiO_2) magma cools quickly so that little or no crystallization

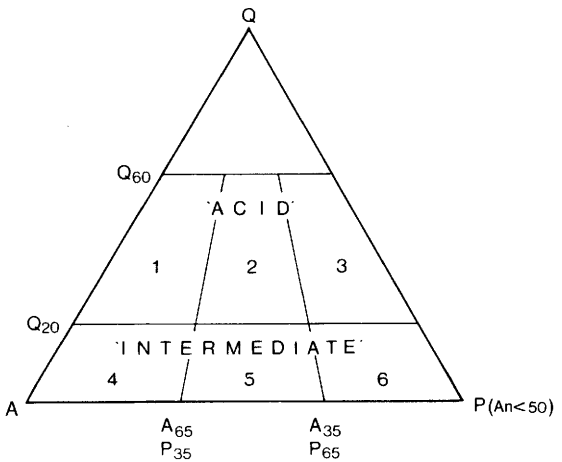


Figure 3.1 *A QAP triangular diagram for the classification of quartz-bearing rocks, showing the fields of composition of the six major families of coarse-grained rocks and their fine-grained equivalents*
(Redrawn from Hatch *et al.*, 1972; Figure 76).

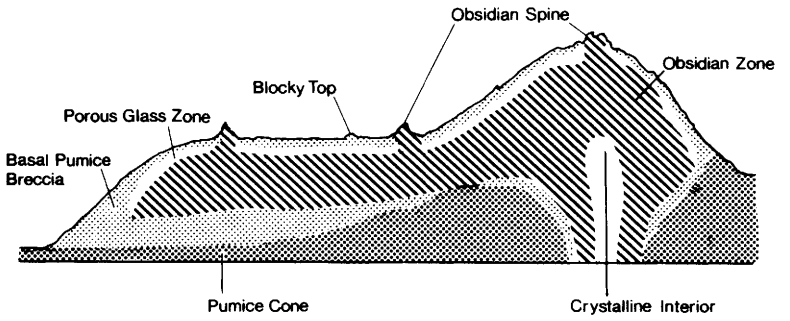


Figure 3.2 *A zone of obsidian in a rhyolite flow-dome structure* (Redrawn with permission from Hughes and Smith, 1993; Figure 2A)

occurs. The abundance of silica confers a high viscosity on the lava. This combination of high viscosity and rapid chilling, either at the thin margins of a flow, or as a result of extrusion under water, causes the molten siliceous material to vitrify into a glass rather than crystallize into a rock. A detailed description of the process of vitrification, and the structure of silicate glasses, is given in Chapter 5. If the magma which solidifies in this way were to crystallize completely, it would almost always be classified on the above definitions as a *rhyolite*, and hence obsidians are often described as being *rhyolitic*. Figure 3.2 shows a cross-section through a typical obsidian zone of artefact-quality obsidian in a rhyolite flow-dome structure. The outer shell of the structure is highly porous (*pumiceous*) glass except where spines of dense obsidian protrude through the porous top. The interior of the structure is crystalline. In some cases, artefact-quality obsidian may not form. In reality a combination of artefact-quality and lesser quality obsidian will normally be present in the same structure (see Hughes and Smith (1993; 81–82), for a more detailed discussion of the formation and geology of obsidian). Obsidian exhibits an excellent conchoidal fracture and a vitreous lustre, and hence its value as a raw material for tools. It has a hardness of 6 on the Moh scale, but is brittle, and has a specific gravity of around 2.4, and a refractive index of about 1.5 (Thorpe, 1978; 93). True obsidians contain little or no original crystalline inclusions, but variations in the magma chemistry and cooling conditions can give rise to a range of vitreous materials, some of which contain appreciable levels of crystalline material. Vitreous material which contains more crystallites is often called *pitchstone*, which can usually be recognized in hand specimens from its dull resinous lustre. Chemically, pitchstone is also characterized by

having a much higher water content than obsidian – levels of up to 10% by weight are not uncommon, whereas obsidian usually contains less than 4%, and often below 1%. During the extrusion of the magma, a fraction may become highly aerated as a result of the rapid expansion (exsolution) of dissolved gases. The resulting solidified material is called *pumice*. This is a highly vesicular version of obsidian, which is unusual and because of its porous structure it is less dense than water, and thus may float for long periods of time, to be washed up on a beach far distant from the source of the magma (Whitham and Sparks, 1986).

Obsidians formed during the Quaternary Period (within the last two million years) possess excellent conchoidal fracture, as do certain obsidians which formed during the Tertiary Period (up to 65 million years old). However, many pre-Tertiary obsidians have lost their ability to fracture conchoidally and have undergone spontaneous re-crystallization, forming a rhyolitic acid rock (Cann and Renfrew, 1964). This is because obsidian formed at high-quench temperatures is not stable over geological time. Generally, vitreous silicates are thermodynamically less stable than their crystalline counterparts, and most glasses will crystallize spontaneously if sufficient energy is available to the atoms within the structure to enable them to re-organize themselves into a crystalline form. In areas of volcanic activity, such energy is often available in the form of heat as a result of further lava flows, or from hot water percolating through the surface rocks. Therefore, on a geological scale obsidians have a short lifetime, and few workable deposits are older than 10 million years (Cann, 1983). This restricts the number of potential sources for obsidian in antiquity, and can give a good indication as to the geographical location of sources on the basis of their geological age.

Another type of alteration which obsidian can undergo forms the basis of a dating technique for tools, called *obsidian hydration dating*. Once a fresh surface of obsidian is formed, such as by knapping to make a tool, the surface takes up water from its surroundings to form what is known as a *hydration layer* on the surface. As noted above, obsidians are characterized by a low original water content, largely because the high temperatures of formation drive off most of the water at the time of solidification. Typical values for water content are 0.1–0.3% by weight, depending in detail on the water pressure in the parent magma and the temperature of extrusion. This water is thermally stable within the obsidian up to temperatures in the region of 800–1 000 °C, again because of the high initial temperatures (Ericson *et al.*, 1976). As the magma flow cools, and after cooling, the surface of the obsidian reacts with water vapour in the atmosphere to form a hydration product which is termed *perlite*, which has a characteristic appearance under the microscope as a result of strain

birefringence. The reaction is essentially one of ion exchange, whereby ions of the alkali metals are removed and replaced by hydroxyl ions, in exactly the same manner as described in Chapter 5 for the corrosion of glass. The birefringence seen optically is a result of changing the chemistry of the layer without a change in volume to accommodate the hydroxyl ion. The result is a clear hydrated layer which can be seen under a polarizing microscope. If a fresh surface is created as a result of tool manufacture, then any geologically accumulated hydration layer is removed, and the rate of build up of the new layer as a result of environmental exposure can be used to estimate the archaeological age of the tool. For archaeological dating purposes, this hydration 'rim' is typically of the order of 1 to 50 μm thick. It is assumed that hydration approximately follows *Fick's Law* of diffusion:

$$x = D\,t^{1/2}$$

where x is the thickness of the layer, D is the hydration rate, and t is the time since exposure of the fresh surface. Complications arise due to the influence of the variations in chemistry between different obsidians and the original water content, as well as expected variations due to moisture content in the environment and temperature. The result is that calibration calculations must be done to accommodate different obsidians in different burial environments. Despite these problems, obsidian hydration dating is still claimed to be capable of producing accurate dates for the manufacture of these artefacts, in the range of 200 to 100 000 years ago (Aitken, 1990; 217).

Chemically, obsidian is composed principally of silica (SiO_2, typically around 74%), together with a number of other major elements, such as aluminium (Al_2O_3, average 13%), sodium (4% as oxide), potassium (4% as oxide), and calcium (0.5–1.5% as oxide). Iron oxides vary from around 0.5% to several percent, and the ratio of Fe^{2+}/Fe^{3+} can vary considerably (Longworth and Warren, 1979). Obsidians may be classified on the basis of their major element composition into two groups: *peralkaline*, with higher combined levels of sodium and potassium than aluminium, and *subalkaline*, in which the reverse is true. Obsidians are often further divided into the following sub-groups: *calcic*, with high levels of calcium and typically low alkalis; *calc-alkaline*, with high levels of both calcium and alkalis; and *alkaline*, with high alkali but low calcium (Williams-Thorpe, 1995; 219). A wide range of trace elements are also present, either dispersed in the glassy matrix or associated with the crystalline phases present, which are normally either quartz or feldspar. Geochemical associations between major and trace elements mean that,

for example, calc-alkaline obsidians are also likely to be richer in other alkaline earth trace elements, such as barium and strontium. The partitioning of trace elements between crystalline and glassy phases gives rise to inhomogeneity on a fine scale, but since solidification from the parent magma is by definition rapid, obsidian is normally said to be chemically relatively homogeneous on a large scale within a single flow. Different flows in the same geographical area may well be different, reflecting changes in composition within the magma chamber with time. Extensive geochemical prospecting of obsidian flows is needed to substantiate this hypothesis, since it is crucial to the chemical characterization of obsidian sources, but in general this has not been done adequately. For example, one of the major publications relating to the sources of artefact grade obsidians in California characterized most of the major obsidian sources with less than five analyses, on the assumption of large scale homogeneity (Jack, 1976). Some of the comparative geochemical work that has been carried out has been directed towards comparing the compositions of obsidian, perlite, and rhyolite from the same source (*e.g.*, Thorpe, 1978, studying three outcrops in the Guadalajara region of Mexico). Recently, Hughes (1994) has investigated the geochemical variability between quarry areas within an outcrop of obsidian in California previously believed to represent one chemically defined 'source'. By analysing 200 samples from 20 distinct geographical locations within the Casa Diablo region, he demonstrated that two or possibly three chemically distinct sources of obsidian exist within this area. It has been appreciated for some time that certain flows, such as the Borax Lake flow in California's North Coast Ranges, exhibit systematic variations in composition, as evidenced by a coherent linear variation between six trace elements and iron content (Bowman *et al.*, 1973), perhaps reflecting the mixing of two different magmas prior to eruption. The implications of the work by Hughes (1994) in California are that more detailed geochemical mapping of obsidian sources may be necessary before homogeneity can be accepted, even for single flows within the same volcanic field or dome complex (see also Hughes and Smith, 1993, for consideration of this and other factors relating to the formation and natural distribution of obsidian).

SOURCES OF OBSIDIAN IN THE EASTERN MEDITERRANEAN AND NEIGHBOURING REGIONS

For the reasons outlined above, outcrops of workable obsidian are relatively few in number and are restricted to areas of geologically recent lava flows. Most sources are therefore reasonably well known, and,

because of these constraints, recognition of new sources in the eastern Mediterranean region becomes ever more unlikely. This makes the exercise of characterizing archaeological obsidians an attractive proposition, since, unlike potential clay sources for pottery provenance, the existence of completely unknown sources can be (cautiously) ignored. This is, of course, subject to the requirement noted above for more detailed geochemical characterization of existing sources.

Systematic exploration of the natural sources of obsidian in the eastern Mediterranean began towards the end of the 19th Century. In 1897, Duncan Mackenzie published the results of a survey of the sources of obsidian on the Cycladic island of Melos which lies around 160 km north of Crete (see Figure 3.3). This was followed shortly after in 1904 by Bosanquet's study of the prehistoric supply of obsidian from Melos to the Aegean. Two sources were identified on the island at Sta Nychia or Adhamas and Demenghaki (Shelford *et al.*, 1982). Other sources in the Eastern Mediterranean include the tiny island of Giali (Yali) in the Dodecanese some 240 km to the east of Melos (see Figure 3.3). Here, the obsidian contains white spherulites (clusters of needle-shaped crystals) and was used to make the small stone vases and other items found on Minoan Crete, which the excavator (Evans) originally ascribed to a source on the island of Lipari on the basis of visual appearance. There have been suggestions that 'pure' obsidian may also have been available from this source (Torrence and Cherry 1976, quoted in Shelford *et al.*, 1982; 190). Small nodules of obsidian on the island of Antiparos, 60 km northeast of Melos, are thought to be of negligible significance, although two unworked pieces were found nearby at the Neolithic site of Saliagos near Antiparos (Cann *et al.*, 1968; 106).

Any provenance investigation of obsidians used in the eastern Mediterranean obviously must also take account of sources of workable obsidian outside the region, although a judgement has to be made about how far this consideration must extend. A number of sources are known in the western Mediterranean, namely the islands of Sardinia, Lipari, Pantelleria, and the Pontine islands. These sources have been investigated in detail (*e.g.*, Hallam *et al.*, 1976; Tykot, in press), enabling valuable discussions to take place regarding the nature of Neolithic exchange (Ammerman *et al.*, 1990). The rich sources in the Carpathian mountains in southeast Slovakia (Zemplén mountains) and the Tokaj mountains of northeast Hungary some 1 200 km to the north-northwest of Melos have been studied in detail (Thorpe, 1978). Some 900 km to the east of Melos, major sources are known at Acigöl and Çiftlik in central Anatolia (Turkey) and further east at Bingöl and Nemrut Dag in eastern Anatolia. Newly discovered sources to the north and northwest

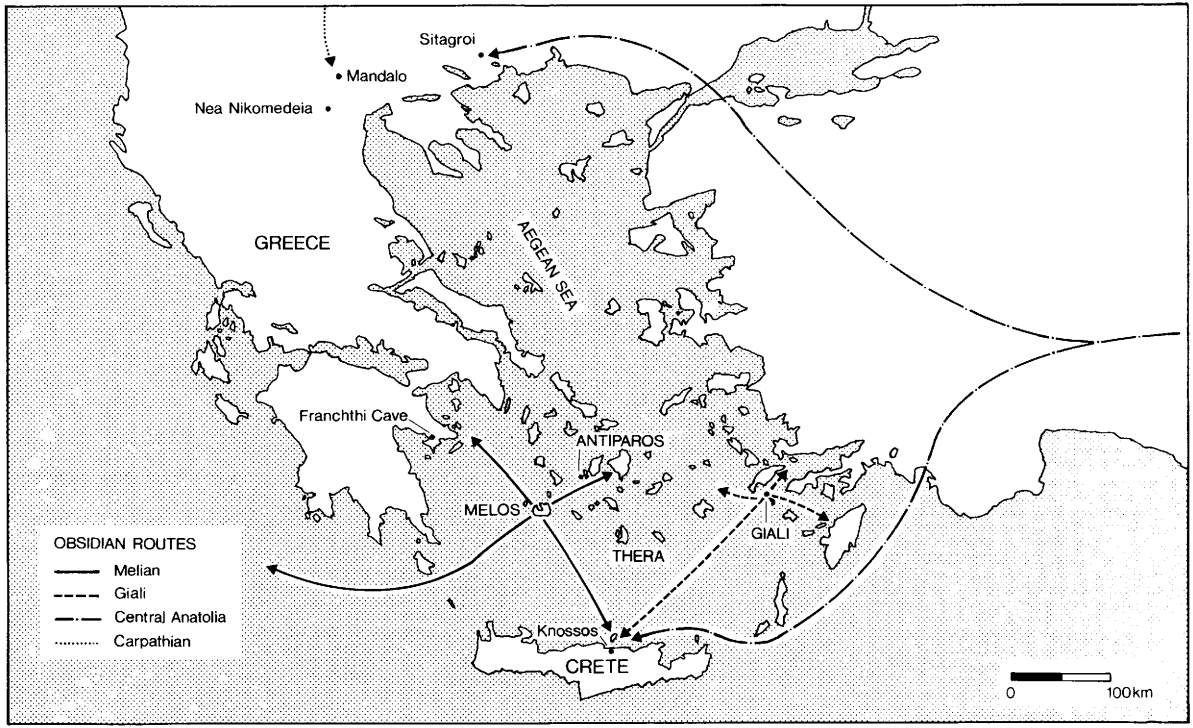


Figure 3.3 *Map of the Aegean showing the location of obsidian sources (Melos, Antiparos, and Giali) and sites referred to in the text. Melian obsidian is found in abundance on many Aegean islands and on the southern Greek mainland. It is also found in northern Greece, western Anatolia, and, apparently, mainland Italy in much smaller quantities. Other arrows represent the most southerly location of Carpathian obsidian (Mandalo) and the presence of central Anatolian obsidian at Sitagroi and Knossos (see text for details)*

of Ankara in northern Anatolia have recently been mapped and must also be taken into account (Keller and Seifried, 1992; 61–62). Further afield, sources in Armenia, the Caucasus, Ethiopia, and Kenya are also known, but are generally discounted from any discussion of Neolithic/Bronze Age supply to the Eastern Mediterranean.

REVIEW OF ANALYTICAL WORK

Early scientific investigations of obsidian examined the usefulness of a number of properties based on appearance. Variations in colour (in transmitted and reflected light), fracture, and translucency did not provide a reliable method of characterization, nor did variations in specific gravity, density, and refractive index. However, some of these properties have proved useful when used in conjunction with the results of chemical analysis. Because of the largely glassy nature of the material, thin (petrographic) sections of obsidian are too isotropic to be used as a means of characterization. Analysts subsequently turned to methods of chemical characterization. Because only a restricted range of magma compositions can satisfy the conditions for vitrification, analyses of the major elements (*e.g.*, Si, Al, Na, K, and Ca) are mainly helpful in only characterizing whether obsidian belongs to per-alkaline, alkaline, and calc-alkaline types, so trace element analysis was pursued as a geographical characterization technique. Trace elements are those which occur at the level of parts per million by weight in the sample, and this restricts the choice of analytical techniques available.

The results of a preliminary investigation of obsidian from the western Mediterranean using wet chemical determinations of elements present in the glass were published by Cornaggia Castiglioni and coworkers in 1963. This was followed shortly after by the first application of optical emission spectroscopy (OES; see Chapter 2 for details of the techniques described in this section; Cann and Renfrew, 1964; Renfrew *et al.*, 1966, 1968). Sample preparation involved crushing small samples of obsidian (*ca.* 60 mg). The technique proved able to determine the amounts of trace elements present in obsidian in proportions between one part per million and around one per cent. The number of samples of obsidian studied was initially rather low, particularly in authentic samples collected at the obsidian source. Quantitative values for 16 elements were determined, although Ba, Zr, Nb, and Y appeared best to differentiate many of the Mediterranean sources. Simple bivariate displays of element concentrations plotted on a logarithmic scale (for example, Ba versus Zr) underlined the potential value of trace element analysis for separating the relevant sources (Figure 3.4). Other pairs of

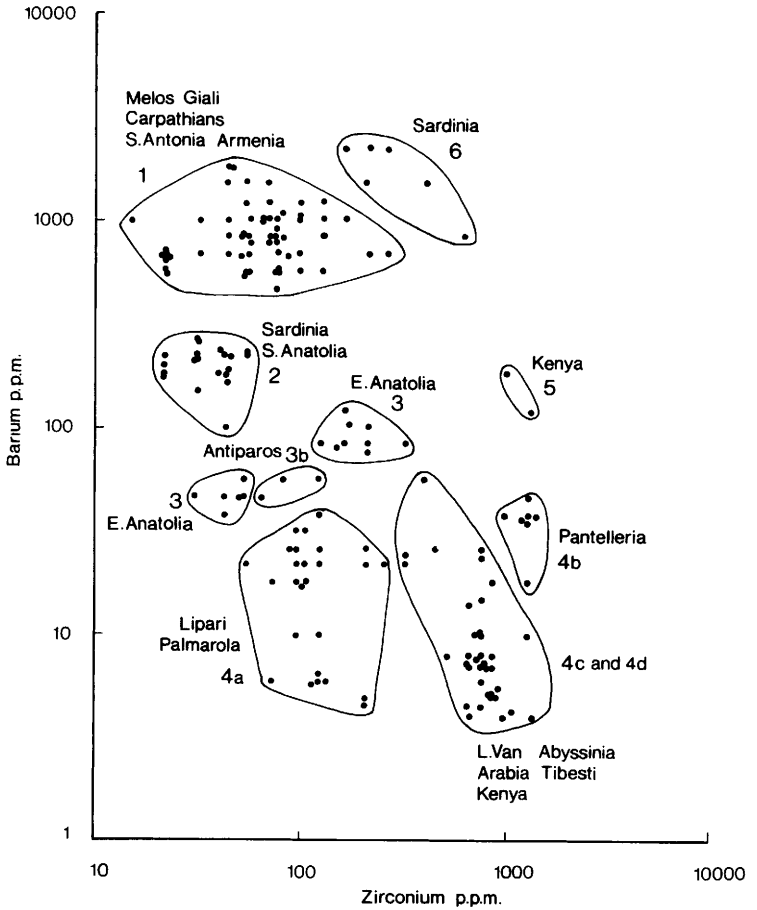


Figure 3.4 *The content of Ba and Zr in a number of Old World obsidians. Each dot represents a single specimen and divisions into discrete groups are indicated* (Redrawn with permission from Cann *et al.*, 1969; Figure 106)

elements could be used to improve definition. However, overlap between certain sources was evident: for example, the so-called Group 1 included obsidians from Giali, Acigöl in central Anatolia, both sources on Melos, and two sources in the Carpathian mountains. It was evident that the technique of OES did not have sufficient precision or the ability to analyse a large enough range of elements to discriminate between these sources.

It was not until the application of neutron activation analysis (NAA) that the problem of overlapping sources could be resolved. NAA is a

highly sensitive and essentially non-destructive technique, although samples have to be taken which remain radioactive for some time after analysis (see also Williams-Thorpe, 1995; 240). The use of NAA in characterizing obsidian from the eastern Mediterranean was first demonstrated in the early 1970s (Aspinall *et al.*, 1972). Sample preparation for NAA is straightforward; samples of obsidian weighing from around 300 to 1000 mg could be analysed intact after washing the tool surface to remove surface impurities. By suitable choice of analytical strategy, a wide range of elements can be detected, in some cases at the sub-ppm level. However, the determination of some elements is problematic. For example, although the aluminium nucleus captures a neutron, the unstable isotope formed has a very short half-life of 2.3 min and can easily be missed if counting takes place some days after irradiation. It is also not possible to determine light elements such as silicon and oxygen, since they present a very low cross-section for neutron capture. In the case of obsidian, the initial gamma ray spectrum is dominated by ^{24}Na (with a half-life of 15 h), so first examination normally takes place after several days has elapsed. In neutron activation analysis, direct comparison is made between the unknown sample and a set of standards of known composition which is irradiated together with the samples under consideration.

The original technique employed allowed quantitative measurement of 16 elements with a much greater precision than OES. In contrast to the OES data, plots of simple elemental ratios allowed a clear distinction to be made between sources. A procedure of internal normalization of a number of elements (*e.g.*, Cs, Ta, Rb, Th, Tb, Ce, and Fe) to Sc (an element determined in obsidian with high precision) was devised to mitigate the effects of precision and accuracy loss arising from neutron flux variations during sample irradiation and variations in geometry during counting due to the widely varying shape and mass of the samples. The normalization procedure led to an improvement in intra-source homogeneity, although at the cost of a loss of accessibility of the data itself. Although the number of samples investigated in the preliminary study was small, these NAA data served to separate all the obsidians in Group 1 as defined by the OES data, including the two sources on Melos only ten kilometres apart.

Although historically NAA has been the method of choice for this work, more recently, because of the increasing difficulties associated with obtaining irradiation facilities for NAA, a growing number of studies have been undertaken using *X*-ray fluorescence (XRF; *e.g.*, Giaque *et al.*, 1993). In such a situation, the question of intercomparability between analytical techniques is bound to become significant – otherwise, all the

data obtained by one method becomes inaccessible to studies using an alternative approach. Shelford *et al.* (1982) report a study of compositional variation within the two sources on Melos using both XRF and NAA. Although the analyses were carried out on different specimens, values could be compared between six elements common to both analyses, and were found to be in good agreement. In this case, the combination of the XRF and NAA data underlined the homogeneity of individual flows on Melos. For most chemical elements each source is homogenous to better than 7% and there are significant differences between the sources for major, minor, and trace element compositions.

Neutron activation analysis and XRF remain the most widely used techniques for obsidian analysis. However, a wide array of other analytical, geochemical, dating, and magnetic approaches have been utilized during the past 20 years. These include fission track dating, thermoluminescence dating, potassium–argon dating, electron spin resonance, strontium isotope ratios, inductively coupled plasma mass spectrometry, back-scattered electron imaging, Mössbauer spectroscopy, magnetic properties, and so on (Williams-Thorpe, 1995; Table 3). These techniques have been applied variously in order to verify conclusions based on earlier analyses, to explore rapid, non-destructive, or minimally-destructive, approaches and, perhaps more cynically, simply to parade new techniques to an analytical audience. One of the most interesting complementary techniques has been the use of *fission track dating* to obtain the geological age of the obsidian flows, originally proposed by Durrani *et al.* (1971). Fission track dating is conceptually simple, in that a freshly polished surface is etched with acid to reveal the structural damage caused by the fission of an uranium nucleus. The 'tracks' are the result of the energetic recoil of the two halves of the uranium nucleus. If the number of tracks can be counted, and the uranium content of the obsidian measured, it is possible in principle from a knowledge of the half-life for fission of the uranium nucleus to calculate the time elapsed since the solidification of the obsidian (Aitken, 1990; 132). The date obtained is generally the geological age of the flow, unless re-heating of the sample has taken place to a temperature sufficient to anneal out the tracks accumulated over geological time, in which case the age may be the time since the last heating event. The obsidian sources in central Anatolia, Giali, and Melos appeared to have broadly similar volcanic histories (two general periods of tectonic activity) on the basis of the fission track dating of the eruptions in these areas (see also Dixon, 1976; 299–300). Keller and co-workers (1990; 24) report much younger dates by fission track methods. The earliest dates for Melos are 3.15–2.7 million years ago. By contrast, Giali [at 24 000 BP (uncalibrated radiocarbon years before

present time)] is among the youngest events in eastern Aegean volcanism. In the western Mediterranean, although obsidian from Sardinia, Lipari, and Pantelleria can be distinguished by fission track dating, the combination of this method and potassium–argon dating appear to give similar dates to the various flows on Sardinia (see review by Tykot, in press). Fission track dating has been used in a combined investigation with NAA to examine obsidian from Anatolian sites near Istanbul in an effort to evaluate the possibility of Aegean or Carpathian obsidian in the Bosphorus region (Bigazzi *et al.*, 1993).

Using another approach, Gale (1981) investigated the use of strontium isotope ratios ($^{87}Sr/^{86}Sr$) to characterize the different sources. This ratio depends on the strontium isotope ratio in the initial magma and the time elapsed since formation, since the decay of the radioactively unstable isotope ^{87}Rb into ^{87}Sr contributes to the abundance of ^{87}Sr. Strontium-86 is stable so elevated ratios of $^{87}Sr/^{86}Sr$ should correlate with older rocks. Whilst the measurement of this parameter proved insufficient by itself to discriminate between all sources, plotting the strontium isotopic composition against the rubidium concentration (as determined by XRF) enabled separation of the major groups, as did plotting the actual concentrations of Sr *versus* Rb.

The usefulness of the magnetic properties of obsidian for characterization purposes has been considered in a pilot study by McDougall *et al.* (1983). On cooling, the magma can be regarded as acquiring a permanent record of the strength and direction of the Earth's magnetic field appropriate to the place and time. If subsequent alteration to these parameters can be ruled out, then this too should provide a 'fingerprint' of the parent obsidian flow. However, although the original magnetic direction can be measured in the parent flow, this cannot subsequently be determined from pieces of obsidian removed from the flow, since the original orientation of the piece will be unknown. Nevertheless, measurements of the intensity of magnetization, saturation magnetization, and low field susceptibility (properties which are independent of original orientation) have proved to be a rapid and useful means for discriminating between many Mediterranean, central European, and Near Eastern sources. Somewhat surprisingly, this promising line of research does not seem to have been followed up.

ARCHAEOLOGICAL IMPLICATIONS

The principal aim of any archaeological provenance study should be an assessment of the economic and social factors which underlie the movement of materials. Few provenance programmes have actually gone as

far as contributing directly to quantitative models of exchange (Cherry and Knapp, 1991). The early success of NAA and XRF in providing an elemental 'fingerprint' for each known source in Europe and the Near East has enabled the definition of a series of obsidian '*interaction zones*' in the western Mediterranean, Aegean, central Europe, and the Near East. An interaction zone has been defined as '*the area within which sites, within the time-range considered, derived thirty per cent or more of their obsidian from the same specific source*' (Renfrew and Dixon, 1976; 147). Since the figure of 30% was arbitrary, Torrence (1986; 13) suggests that all sites comprising obsidian should be included in any study of exchange mechanisms.

Both sources on Melos have proved to be the source for the great majority of the obsidian used in the Aegean during the Neolithic and Early Bronze Age (*i.e.*, in the above terminology, much of the Aegean falls into the single interaction zone of obsidian from Melos). Melian obsidian served much of the Greek mainland as far north as the sites at Nea Nikomedeia and Servia in Macedonia and Sitagroi in Thrace, although very few pieces are found in this region. At Sitagroi, Melian obsidian occurs with Anatolian obsidian from Çiftlik (one piece). The latter piece was not well stratified but probably dates from the Early Bronze Age. Renfrew and co-workers (1966) found that three pieces from later Neolithic levels at Knossos on Crete are also from Çiftlik. This suggests relatively long range contacts at a very early date. The characteristic obsidian from Giali in the Dodecanese is found in Crete, Kos, Rhodes, Kalymnos, and Saliagos (Cherry, 1985; 15).

The obsidian in the Upper Palaeolithic and Mesolithic levels at Franchthi Cave in the Argolid was also found to be Melian (Renfrew and Aspinall, 1990). Obsidian exploitation at this time pre-dates settlement evidence on Melos, which may not have been inhabited until the Late Neolithic. Melian obsidian is also found in western Anatolia, albeit in low abundance. For example, Renfrew *et al.*, (1965; 238) characterized two pieces of Melian obsidian at the Late Neolithic/Early Chalcolithic site of Morali. In the study by Gale (1981), one find from Haçilar in Turkey, previously examined by Renfrew and co-workers (1966) and assigned to the Acigöl source in central Turkey on the basis of trace element composition, was considered to originate from the Demenghaki source on Melos using the isotopic and chemical data. However, this reinterpretation has subsequently been regarded as unlikely, on the grounds that insufficient samples were used to define the characteristics of each source on the basis of the $^{87}Sr/^{86}Sr$ ratio and Sr/Rb concentrations (Renfrew and Aspinall, 1990; 270). Adequate characterization of sources is discussed in detail by Hughes and Smith (1993). In their study of obsidian in the Bosphorus region, Bigazzi *et al.* (1993) concluded that

most of the obsidian derived from central Anatolia, northern Anatolia, as well as some minor local occurrences, and/or eastern Anatolia sources. Recent evidence has come to light of Carpathian obsidian at the Neolithic/Bronze Age site at Mandalo in Macedonia, Greece where the unusually large obsidian assemblage for this area also includes obsidian from Melos (Kilikoglou *et al.*, in press). Neutron activation analysis was carried out on twelve samples; ten from the Late Neolithic Phase II (5th Millennium BC) and two from the Bronze Age Phase III (4th–3rd Millennium BC). One of the latter samples corresponded with the Demenghaki source on Melos, while the other eleven fell within the expected range for the Carpathian 1 source in southeastern Slovakia. This finding extends the known distribution of Carpathian obsidian for another 400 km to the south into modern Greece where it also occurs with obsidian originating from Melos. Obsidian from the Carpathian 1 source (a grey transparent obsidian from Szöllöske and Malá Toroňa in Slovakia) was widely used and distributed, and a single piece of central European obsidian from this source has been found at the Neolithic cave site of Grotta Tartaruga in northeast Italy, where obsidian from the island of Lipari predominates (Williams-Thorpe *et al.*, 1984). Lipari lies off the coast of Sicily and supplied large quantities of obsidian mainly to the southern Italy peninsula. More recently, Randle *et al.* (1993) have identified a single piece of Carpathian obsidian at the Early Neolithic site at Sammardenchia in northeast Italy. Thus the western Mediterranean shows at least some overlap with the central European (Carpathian 1) source which in turn overlaps with Melian obsidian at Mandalo in Macedonia.

Suggestions made in the early 20th Century that obsidian on the island of Malta, where no natural sources occur, might have been brought there from Melos by Minoan traders was given some support by the findings of Cornaggia Castiglioni and co-workers (1963). However, the analytical programme of trace element analyses carried out by Renfrew, Cann, and Dixon suggested strongly that the obsidian derived from Lipari and Pantelleria, a tiny island 240 km northwest of Malta. More recently, the case for Melian obsidian in mainland Italy has been made on the basis of fission track dates on obsidian from Grotta del Leone and Capraia Island (both in Tuscany) and Ponte Peschio in Abruzzi (Bigazzi and Radi, 1981; Arias *et al.*, 1984). Obsidian from Grotta del Leone was subsequently analysed by NAA (Bigazzi *et al.*, 1986; 1992). The data appear to reinforce the suggestion that Melos is the source of the obsidian. Unfortunately the obsidian is from archaeologically undatable contexts.

The underlying aim of obsidian provenance has always been to

understand the mechanisms which account for its distribution over large areas. A quantitative approach to obsidian distribution in the Neolithic of the Near East was proposed by Renfrew and co-workers (Renfrew *et al.*, 1968). This study proceeded by noting the 'fall-off' with distance of obsidian as a proportion of the obsidian from that source in the total worked stone assemblage. Up to around 300 km from each source, the proportion of obsidian in the worked stone assemblage remained close to 100% (*i.e.*, the slope of the line is nearly flat). At greater distances, the proportion of obsidian declines dramatically. The concept of a '*supply zone*' and a '*contact zone*' was developed to explain this pattern of obsidian distribution around the Near Eastern sources (Renfrew *et al.*, 1968). The supply zone encompassed the area around each source, within which groups travelled to the source directly in order to procure their own supplies of obsidian. Quantitatively the supply zone incorporated those sites where the lithic assemblage contained more than 80% obsidian. The contact zone comprised assemblages on the sloping line. Within this zone obsidian from the source is still found in the worked stone assemblage, but settlements obtained obsidian, not through direct access to the source, but through regular reciprocal exchanges with neighbours located nearer the source. The system of communities passing on some of their obsidian to adjacent villages was termed '*down-the-line exchange*'. Since each village would retain a proportion of the obsidian, the fall-off with distance in the contact zone approaches an exponential. Organized trade in obsidian, involving profit motives and fixed rates of exchange has been rejected (see Renfrew, 1973; 180) in favour of these anthropological models of so-called '*primitive*' exchange (see Torrence, 1986; 14).

In the Aegean, the application of such models has been made difficult by the absence of quantitative data on the proportion of obsidian in stone tool assemblages and the fact that in some areas obsidian is the only stone type in the worked stone assemblage (see Torrence, 1986; 97–98). However, in the Neolithic (and earlier periods), Renfrew suggested a *direct access model*, whereby most consumers travelled to Melos and obtained their own supplies from the quarries (Renfrew, 1972; 442–3). On the Greek mainland, down-the-line exchange could account for an expansive distribution network. However, the sea precludes the establishing of straightforward supply and contact zones with plots of the fall-off in obsidian with distance from Melos. The direct access model may have continued in the Bronze Age (Torrence, 1986; 105). A comprehensive review and analysis of the theories of obsidian exchange applied in the eastern Mediterranean has been carried out by Torrence, (1986). Many of Renfrew's suggestions have been supported (and

augmented) by detailed study of the quarry sites on Melos. For example, Torrence suggests that direct access could have included skilled knappers visiting the island intentionally to collect suitable obsidian as well as visits made as part of some other activity, such as fishing, when unmodified nodules were collected for use back home. In contrast, Perlès (1990b) argues that in the Early to Middle Neolithic, the movement of obsidian was the result of specialized trading whereas in the Late Neolithic, this system becomes 'despecialized' with nearer groups obtaining their obsidian by direct access (see Phillips, 1992 for a review). The current debate relies on the nature of obsidian assemblages recovered from excavation and survey of both sites and sources, such as the presence or absence of unworked nodules, preformed cores, debitage, and so on. More generally, Perlès (1992) has considered the evidence for the movement of a number of rock types in Greece, including andesite, emery, and honey flint, and emphasized the importance of viewing a distribution network of a specific material in the context of the production and distribution of other materials.

Moving away from the Aegean, it is important to note that these models might not be universally applicable. In the western Mediterranean, obsidian could be obtained from the islands of Sardinia, Lipari, Palmarola (Pontine Islands), and Pantelleria. Recent work by Ammerman *et al.* (1990) suggests that in northern Italy, large distances to the sources precludes direct access. In this region certain sites, where obsidian is found in large quantities, may have served as nodes of exchange, and certain types of obsidian (such as the translucent blades from Lipari) may have been more valued as prestige items compared to the utilitarian tools fabricated from Sardinian obsidian. The implication is that a number of different exchange mechanisms (both synchronic and diachronic) may be needed to account for obsidian originating from different sources (see also Tykot, in press).

SUMMARY

As long ago as 1982, it was possible to state that '*we can now be reasonably confident that the vast majority of the obsidians found on other Cycladic Islands and on the Greek mainland are Melian in origin*' (Shelford *et al.*, 1982; 120). For Cherry and Knapp (1991; 95), the attribution to source obsidian in the Aegean represents '*an unusually simple provenance problem*'. Without doubt they are correct. For some, early successes with obsidian may have raised 'unrealistic expectations' about the ability of analytical techniques to locate the source of other materials, particularly ceramics (*ibid.*; 95). Whether further large-scale analytical programmes in the Aegean are

justified is questionable although the quantities of Melian obsidian reaching western Anatolia and the source and availability of obsidian reaching the fringes of the Carpathian and Melian 'spheres of influence' are still to be elucidated.

In contrast to Aegean obsidian, large scale research programmes continue in the western Mediterranean (Tykot, 1992; Tykot, in press). In the east, the volcanic sources of western Anatolia and adjacent areas have not been surveyed fully. In addition to the known sources in central Anatolia (Acigöl and Çiftlik) and around Lake Van in eastern Anatolia, new important sources may yet come to light (Keller and Seifried, 1992). In the latter region, a large number of obsidian occurrences (currently about 25) are known stretching over very substantial areas. The approach by Keller and Seifried has been to use XRF combined with discriminant function analysis. New sources in northern Anatolia have been added to the corpus. In other areas of Anatolia, some sources may have been covered by alluvium several metres thick. The natural movement and redeposition of obsidian must also be taken into account. Undoubtedly the interaction zones will continue to be refined, but the general patterns of distribution are unlikely to be greatly changed by future chemical analysis. The study of exchange mechanisms will continue to be debated. As a case study in archaeological science, the characterization of obsidian exemplifies the successful integration of chemistry and physics within archaeological goals.

Not all rocks can be characterized as successfully as obsidian. The success of obsidian provenance is due to the limited number of workable sources and the fact that each source, while being relatively homogenous, is sufficiently different from the other sources to enable elemental 'fingerprinting'. This picture contrasts markedly from flint; a stone not restricted to a few sources and which is inhomogeneous in composition.

REFERENCES

Aitken, M.J. (1990). *Science-based Dating in Archaeology*. Longman, London.

Ammerman, A., Cesana, A., Polglase, C. and Terrani, M. (1990). Neutron activation analysis of obsidian from two Neolithic sites in Italy. *Journal of Archaeological Science* **17** 209–220.

Arias, C., Bigazzi, G., Bonadonna, F.P., Cipolloni, M., Hadler, J.C., Lattes, C.M.G. and Radi, G. (1984). Fission track dating in archaeology: A useful application. In *Scientific Methodologies Applied to Works of Art*, ed. Parrini, P.L., Montedison Progetto Cultura, pp. 151–159.

Aspinall, A., Feather, S.W. and Renfrew, C. (1972). Neutron activation analysis of Aegean obsidians. *Nature* **237** 333–334.

Bigazzi, G. and Radi, G. (1981). Datazione con le tracce di fissione per l'identificazione della provenienza dei manufatti di ossidiana. *Rivista di Scienza Preistoriche* **36** 223–250.

Bigazzi, G., Meloni, S., Oddone, M. and Radi, G. (1986). Provenance studies of obsidian artefacts: Trace elements and data reduction. *Journal of Radioanalytical and Nuclear Chemistry Articles* **98** 353–363.

Bigazzi, G., Meloni, S., Oddone, M. and Radi, G. (1992). Study on the diffusion of Italian obsidian in the Neolithic settlements. In *Atti del VIII Convegno Nazionale sulla Attività di Ricerca nei Settori della Radiochimica e della Chimica Nucleare, delle Radiazioni e dei Radioelementi, Torino, 16–19 Giugno 1992*, CNR and Università degli Studi di Torino, pp. 243–247.

Bigazzi, G., Ercan, T., Oddone, M., Ozdogan, M. and Yegingil, Z. (1993). Application of fission track dating to archaeometry: Provenance studies of prehistoric obsidian artifacts. *Nuclear Tracks and Radiation Measurements* **22** 757–762.

Bloedow, E.F. (1987). Aspects of ancient trade in the Mediterranean. *Studi Micenei ed Egeo-Anatolici* **26** 59–124.

Bowman, H.R., Asaro, F. and Perlman, I. (1973). Composition variations in obsidian sources and the archaeological implications. *Archaeometry* **15** 123–127.

Cann, J.R. (1983). Petrology of obsidian artefacts. In *The Petrology of Archaeological Artefacts*, eds. Kempe, D.R.C. and Harvey, A.P., Clarendon Press, Oxford, pp. 227–255.

Cann, J.R., Dixon, J.E. and Renfrew, C. (1968). Appendix IV: The sources of the Saliagos obsidian. In *Excavations at Saliagos near Antiparos*, eds. Evans, J.D. and Renfrew C., British School of Archaeology at Athens, Supplementary Volume 5, Thames and Hudson, London, pp. 105–107.

Cann, J.R., Dixon, J.E. and Renfrew, C. (1969). Obsidian analysis and the obsidian trade. In *Science in Archaeology*, eds. Brothwell, D. and Higgs, E., Thames and Hudson, London, pp. 578–591 (2nd edn.).

Cann, J.R. and Renfrew, C. (1964). The characterisation of obsidian and its application to the Mediterranean region. *Proceedings of the Prehistoric Society* **30** 111–133.

Cherry, J.F. (1985). Islands out of the stream. In *Prehistoric Production and Exchange*, eds. Knapp, A.B. and Stech, T., UCLA Institute of Archaeology Monograph 25, Institute of Archaeology, UCLA, Los Angeles, pp. 12–29.

Cherry, J.F. and Knapp, A.B. (1991). Quantitative provenance studies and Bronze Age trade in the Mediterranean: Some preliminary reflections. In *Bronze Age Trade in the Mediterranean*, ed. Gale, N.H.,

Studies in Mediterranean Archaeology XC, Paul Åströms Förlag, Jonsered, pp. 92–119.

Cornaggia Castiglioni, C.O., Fussi, F. and D'Agnolo, M. (1963). Indagini sulla provenienza dell'ossidiana utilizzata nelle industrie preistoriche del Mediterraneo occidentale. *Atti della Societa Italiana di Scienz Naturali e del Museo Civico di Stroia Naturale in Milano* **102** 310–322.

Disa, J.J., Vossoughi, J. and Goldberg, N.H. (1993). A comparison of obsidian and surgical steel scalpel wound-healing in rats. *Plastic and Reconstructive Surgery* **92** 884–887.

Dixon, J.E. (1976). Obsidian characterization studies in the Mediterranean and Near East. In *Advances in Obsidian Glass Studies*, ed. Taylor, R.E., Noyes Press, Park Ridge, New Jersey, pp. 288–333.

Durrani, S.A., Khan, H.A., Taj, M. and Renfrew, C. (1971). Obsidian source identification by fission track analysis. *Nature* **233** 242–245.

Ericson, J.E., Mackenzie, J.D. and Berger, R. (1976). Physics and chemistry of the hydration process in obsidians I. Theoretical implications. In *Advances in Obsidian Glass Studies*, ed. Taylor, R.E., Noyes Press, Park Ridge, New Jersey, pp. 25–45.

Evans, A. (1921). *The Palace of Minos*, Volume 1, Macmillan, London.

Gale, N.H. (1981). Mediterranean obsidian source characterisation by strontium isotope analysis. *Archaeometry* **23** 41–52.

Giaque, R.D., Asaro, F., Stross, F.H. and Hester, T.R. (1993). High-precision nondestructive *X*-ray fluorescence method applicable to establishing the provenance of obsidian artefacts. *X-Ray Spectrometry* **22** 44–53.

Hallam, B.R., Warren, S.E. and Renfrew, C. (1976). Obsidian in the western Mediterranean. *Proceedings of the Prehistoric Society* **42** 85–110.

Hatch, F.H., Wells, A.K. and Wells, M.K. (1972). *Petrology of the Igneous Rocks*. Thomas Murby, London (13th edn.).

Hughes, R.E. (1994). Intrasource chemical variability of artefact-quality obsidians from the Casa Diablo area, California. *Journal of Archaeological Science* **21** 263–271.

Hughes, R.E. and Smith, R.L. (1993). Archaeology, geology and geochemistry in obsidian provenance studies. In *Effects of Scale on Archaeological and Geoscientific Perspectives*, eds. Stein, J.K. and Linse, A.R., Geological Society of America Special Papers **283** 79–91.

Jack, R.N. (1976). Prehistoric obsidian in California I. Geochemical aspects. In *Advances in Obsidian Glass Studies*, ed. Taylor, R.E., Noyes Press, Park Ridge, New Jersey, pp. 183–217.

Keller, J. and Seifried, C. (1992). The present status of obsidian identification in Anatolia and the Near East. In *Volcanologie et Archeologie*, eds. Livadie, C.A. and Widemann, F., PACT **25** 57–87.

Keller, J., Rehren, T. and Stadlbauer, E. (1990). Explosive volcanism in the Hellenic arc. In *Thera and the Aegean World III, Volume 2, Earth Sciences*, ed. Hardy, D.A., Thera Foundation, London, pp. 13–26.

Kilikoglou, V., Bassiakos, Y., Grimanis, A.P., Souvatzis, K., Pilali-Papasteriou, A. and Papanthimou-Papaefthimiou, A. (in press). Carpathian obsidian in Macedonia, Greece. *Journal of Archaeological Science.*

Longworth, G. and Warren, S.E. (1979). The application of Mössbauer spectroscopy to the characterisation of western Mediterranean obsidian. *Journal of Archaeological Science* **6** 179–193.

McDougall, J.M., Tarling, D.H. and Warren, S.E. (1983). The magnetic sourcing of obsidian samples from Mediterranean and Near Eastern sources. *Journal of Archaeological Science* **10** 441–452.

Perlès, C. (1987). *Les Industries Lithiques Taillées de Franchthi (Argolide, Grèce). Tome 1, Présentation générale et industries paléolithiques, Excavations at Franchthi Greece, Fasicule 3*, Indiana University Press, Bloomington and Indianapolis.

Perlès, C. ed. (1990a). *Les Industries Lithiques de Franchthi (Argolide, Grèce), Tome II: Les Industries de Mésolithique et du Néolithique Initial, Fasicule 5*, Indiana University Press, Bloomington and Indianapolis.

Perlès, C. (1990b). L'outillage de pierre taillée néolithique en Grèce: Approvisionnement et exploitation des premieres materiéres. *Bulletin de Correspondance Hellenique* **114** 1–42.

Perlès, C. (1992). Systems of exchange and organization of production in Neolithic Greece. *Journal of Mediterranean Archaeology* **5** 115–164.

Philipps, P. (1992). Western Mediterranean obsidian distribution and the European Neolithic. In *Sardinia in the Mediterranean: A Footprint in the Sea*, eds. Tykot, R.H. and Andrews, T.K., Sheffield Academic Press, Sheffield, pp. 71–82.

Randle, K., Barfield, L.H. and Bagolini, B. (1993). Recent Italian obsidian analyses. *Journal of Archaeological Science* **20** 503–509.

Read, H.H. (1970). *Rutley's Elements of Mineralogy*, Thomas Murby & Co., London (26th edn.).

Renfrew, C. (1972). *The Emergence of Civilisation: The Cyclades and the Aegean in the Third Millennium B.C.*, Methuen, London.

Renfrew, C. (1973). Trade and craft specialisation. In *Neolithic Greece*, ed. Theocharis, D.R., National Bank of Greece, Athens, pp. 179–200.

Renfrew, C. and Aspinall, A. (1990). Aegean obsidian and Franchthi Cave. In *Les Industries Lithiques de Franchthi (Argolide, Grèce), Tome II: Les Industries de Mésolithique et du Néolithique Initial, Fasicule 5*, ed. Perlès, C., Indiana University Press, Bloomington and Indianapolis, pp. 258–70.

Renfrew, C. and Dixon, J.E. (1976). Obsidian in western Asia: A review.

In *Problems in Economic and Social Archaeology*, eds. Longworth, I.H. and Sieveking, G., Duckworth, London, pp. 137–150.

Renfrew, C., Cann, J.R. and Dixon, J.E. (1965). Obsidian in the Aegean. *Annual of the British School of Archaeology at Athens* **60** 225–247.

Renfrew, C., Dixon, J.E. and Cann, J.R. (1966). Obsidian and early culture contact in the Near East. *Proceedings of the Prehistoric Society* **32** 30–72.

Renfrew, C., Dixon, J.E. and Cann, J.R. (1968). Further analysis of Near Eastern obsidians. *Proceedings of the Prehistoric Society* **34** 319–331.

Shelford, P., Hodson, F., Cosgrove, M.E., Warren, S.E. and Renfrew, C. (1982). The sources and characterisation of Melian obsidian. In *An Island Polity: The Archaeology of Exploitation in Melos*, eds. Renfrew, C. and Wagstaff, J.M., Cambridge University Press, Cambridge, pp. 182–92.

Thorpe, O.W. (1978). *A Study of Obsidian in Prehistoric Central and Eastern Europe, and Its Trace Element Characterization.* Unpublished Ph.D. Thesis, Postgraduate School of Studies in Physics, University of Bradford.

Torrence, R. (1986). *Production and Exchange of Stone Tools.* Cambridge University Press, Cambridge.

Tykot, R.H. (1992). The sources and distribution of Sardinian obsidian. In *Sardinia in the Mediterranean: A Footprint in the Sea*, eds. Tykot, R.H. and Andrews, T.K., Sheffield Academic Press, Sheffield, pp. 57–70.

Tykot, R.H. (in press). Mediterranean islands and multiple flows: The sources and exploitation of Sardinian obsidian. In *Method and Theory in Archaeological Obsidian Studies*, ed. Steven Shackley, M., Advances in Archaeology and Museum Science Series, Plenum Press, New York.

Whitham, A.G. and Sparks, R.S.J. (1986). Pumice. *Bulletin of Vulcanology* **48** 209–223.

Williams-Thorpe, O. (1995). Obsidian in the Mediterranean and the Near East: a provenancing success story. *Archaeolometry* **37** 217–248.

Williams-Thorpe, O., Warren, S.E. and Nandris, J.G. (1984). The distribution and provenance of archaeological obsidian in central and Eastern Europe. *Journal of Archaeological Science* **11** 183–212.

Chapter 4

The Geochemistry of Clays and the Provenance of Ceramics

INTRODUCTION

If *clay* is heated to a sufficient temperature, an irreversible chemical change takes place. The product, *pottery* or *ceramic*, is in most instances a durable material although there are examples of the comminution of low fired, friable ceramics leading to complete mechanical destruction as a result of repeated freeze-thaw cycles or by cryoturbation. The earliest 'ceramics' derive from the assemblage of fired loess figurines (including the famous 10 cm Venus figure) and other objects from the Palaeolithic site cluster at Dolní Věstonice and Pavlov in the Czech Republic (Vandiver *et al.*, 1989). Some 15 000 years later, at around 10–12 000 years ago, the first pottery vessels were produced, although precise dates are open to debate. Generally, pottery combines durability with ubiquity. Pottery-using communities are to be found in many areas of the world. The reasons underlying the adoption and subsequent trajectory of use are not straightforward. For example, equating pottery with the first farming communities is anachronistic since in some regions hunter-gatherer populations made and used pottery, whilst in others early farming communities were aceramic (hence the term '*pre-pottery Neolithic*'). In some areas (such as the Viking period on Orkney) pottery manufacture became redundant. The causes of adoption and use (or disuse) of ceramics may vary according to local social, economic, or environmental factors.

Some years ago, Renfrew (1977; 3) considered that the study of pottery had '*won itself a bad name in some archaeological circles.*' Archaeologists had become obsessed with typologies of pottery styles. For many years, sites and regions could only be dated through comparison of characteristic pottery types with an established pottery typology, based on the evolution of form or decoration to construct a sequence. In some

periods, dating by pottery is far more precise than any scientific dating technique and is only surpassed if identifiable coins are found in stratified contexts. In the Roman period, it is not unusual to date pottery to within 25 years. Renfrew identified a number of possible research areas for ceramic studies, including the study of the origins of pottery, the wider implications for form and decoration of pottery, manufacturing technology, provenance, function, and the influence of formation processes on ceramic evidence. The considerable literature focusing on at least some of these topics and the deployment of a battery of scientific, experimental, and ethnoarchaeological approaches attests to a healthy perspective for ceramic studies in general, although problems remain in the lack of interdisciplinary overlap (Bishop and Lange, 1991). Recent general textbooks covering pottery analysis and interpretation include Rice (1987) and Orton *et al.* (1993).

Identification of trade in pottery is also important for dating purposes, since the presence of imported goods can be used to construct temporal linkages between undated cultures and those with either a calendrical chronology (such as Egypt) or with a well-established chronology such as in the prehistoric Aegean. However, the movement of pottery is of significance in its own right for examining cultural relations and the economic and social factors responsible for promoting trade. In the ancient world, the significance of pottery must, however, be kept in perspective. It was almost always only the container for the item of real trade significance (*e.g.*, the use of *amphorae* to transport wine around the Roman empire) and it only represents a fraction of the trade economy – much of it would have been perishable, and therefore largely invisible today. The problem of ceramic provenance is one of immense subtlety, since even today the supply of local (or cheaper imported) imitations of recognized 'brand names' can be of major economic significance. So it was in antiquity – for example, large pottery workshops were set up in Britain after the Roman conquest to produce vessels similar to those made on the Continent. The identification of the place of manufacture of some ceramic vessels is therefore not always straightforward. Consequently, one of the major applications of analytical chemistry to archaeology has for some time been to measure the major, minor, or trace element composition of archaeological objects for the purposes of assigning *provenance* – the geographical origin of the raw materials used. Given the importance of traded pottery, it is not surprising that one of the biggest single classes of material to be studied has been pottery. A good example of the use of chemical techniques of ceramic provenance is the work carried out on pottery from the prehistoric Mediterranean, summarized by Jones (1986).

The theory behind this work is extremely simple – perhaps to the point of being naive. It is assumed that the chemical composition of the fired ceramic is indicative of the chemical composition of the principal raw material – clay. It has always been acknowledged that the transformation of raw clay into fired pottery is potentially a complex process, which might involve a number of factors, all of which could influence the final composition of the product. These include:

(i) the natural variability of the clay beds themselves;
(ii) selection and mixing of clays from different sources to give the correct colour and working properties;
(iii) *levigation* and/or processing of the raw clay to remove unwanted material and give the desired texture;
(iv) addition of *temper* (non-clay filler), usually to modify the thermal properties of the body;
(v) the firing cycle itself, which might affect the composition via the volatility of some components.

An additional problem in the study of archaeological ceramics is the possibility of post-depositional alteration.

In order to compensate for these problems, the normal procedure has been to compare the finished pottery not with the raw clay from which it was produced, but with fired pottery of certain or assumed provenance. Most commonly, 'control groups' are established from either kiln '*wasters*' (distorted or broken vessels found by archaeological investigation in the vicinity of the kilns themselves), or by comparison with material of impeccable provenance – examples certified by experts as being 'type specimens', or, occasionally, bearing a makers stamp (although even this can occasionally be deceptive). Very few chemical studies have attempted to relate a fired piece of pottery to the actual clay bed from which it was made.

There has been endless (and ultimately rather pointless) debate in the archaeological literature over the past 20 years about the best procedure for carrying out ceramic provenance work. Much of it has centred around the 'best' instrumentation to use, which relates to deciding which elements are the 'best'. Neutron activation analysis (NAA), which is capable of measuring more than 20 trace or ultra-trace elements has had a great deal of support, although the proponents of the various types of inductively-coupled plasma spectroscopies with similar analytical capabilities have become more vocal in recent years. Other techniques, such as atomic absorption spectrometry or *X*-ray fluorescence, which can produce a 'total' analysis of major and minor constituents, have been

used for particular studies, particularly those with an eye on the relationship between original clay mineralogy and the composition of the finished product. There has even been a 'reductionist' (but to our mind misleading) view which believes that chemical provenancing of ceramics can be reduced to a routine procedure whereby the same limited number of elements are measured by the same method, regardless of the mineralogy or geochemistry of the clays involved, or the manufacturing process used. In this latter context, the contribution of *ceramic petrology* has been of paramount importance in the re-construction of production techniques, and in relating the ethnographic record of 'primitive' pottery production with that observed in archaeological examples.

Much of the debate about the 'best' analytical method, and also that which ensues about the 'best' multivariate analysis techniques to be used on the data, is largely fruitless. A much greater understanding of the analytical data is obtained if it is considered in the light of the known geochemical behaviour of trace elements (or minor elements, depending on the techniques employed) in clays, and in the context of the archaeological question to be answered. This chapter therefore reviews the basic structure of silicate minerals, leading to an overview of the chemistry of the *aluminosilicate (clay) minerals*, followed by a discussion of the relevant trace element geochemistry of clays and the effect of firing on the composition of clay. It concludes with a case study of the provenance of a particular class of Roman fineware known as '*Rhenish*' ware.

THE STRUCTURE OF CLAY MINERALS

The term 'clay' is a somewhat ambiguous one. To a soil scientist, it implies a particular size fraction of the fine portion of the soil – one whose size is such that its properties can be regarded as those of a *colloid* – a large surface area per unit weight, with the consequent presence of surface electrical charges which attract ionic species and water (Brady, 1974). The term 'clay' is usually used for those particles with a diameter of less than 0.002 mm, with 'silt' referring to diameters between 0.002 and 0.05 mm, and 'sand' up to 2 mm. Soil classifications depend on the exact particle size distribution, with regions defined on a triangular diagram with axes labelled 'clay', 'silt', and 'sand' (Figure 4.1). To a chemical mineralogist, clay minerals are part of the large family of silicates, which form the majority of the rocks of the Earth's surface. Specifically, clay minerals are related to the *micas*, which are *sheet silicates* (see below), and chemically they are generally hydrated silicates of aluminium. The structural chemistry of silicates in general and clay minerals in particular is discussed below. To a geologist, clays are the weathering products of rocks. They have been

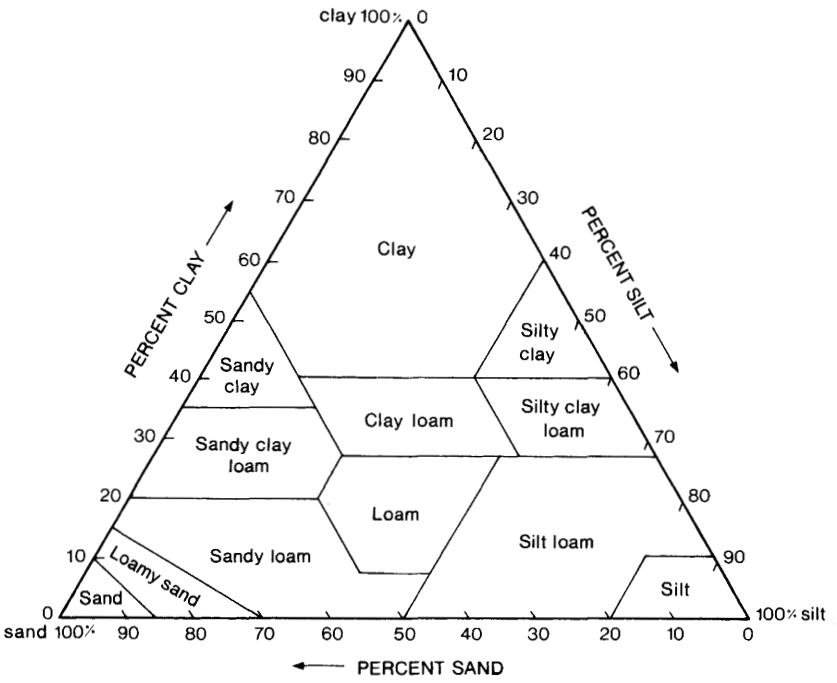


Figure 4.1 *Relationship between class name of a soil and its particle size distribution* (After Brady/Buckman, *The Nature and Property of Soils*, 10th edn., © 1990, p. 99. Reprinted by permission of Prentice Hall, Upper Saddle River, New Jersey)

classified by Ries (1927) according to factors such as the nature of their parent rock (*e.g.*, granite, gneiss, *etc.*), the weathering process involved and their position with respect to their parent rock (*e.g.*, residual, colluvial, or transported). To a ceramic technologist, clays have been classified according to the system devised by Norton (1952), which is based on their properties, such as white-burning, refractory, brick clays, *etc.* The geological and ceramic classifications of clays are discussed in more detail by Singer and Singer (1963); Table 4.1 reproduces the latter classification, since it is of great relevance to archaeological ceramicists.

Chemical mineralogy has a long history in western science, stretching as far back as the observation by Libavius in 1597 that certain salts crystallized during alchemical processes could be identified by their *crystal habit* (Evans, 1966). Habit is a crystallographic term which refers to the characteristic shape of crystals, which often reflects the internal structure. Observations were made throughout the 17th and 18th Centuries of the crystal habits of many common inorganic minerals, including precise

measurements of crystallographic angles after the invention of the contact goniometer in 1780. The classification of minerals as stoichiometric compounds is attributed to Berzelius in 1824 (von Meyer, 1891), but it was not until the development of *X*-ray crystallography in 1912 that serious structural analysis became possible (Bragg, 1937).

Despite the fact that 95% of the Earth's crust is made up of only five silicate mineral groups – *feldspars, quartz, amphiboles, pyroxenes*, and *micas* (Putnis, 1992), the silicates proved particularly intractable to classify before the advent of structural analysis by *X*-ray crystallography. Initially classification was attempted on the basis of regarding them as salts of fictitious silicic acids, but this was unsuccessful. There are many reasons for this failure, largely relating to the wide variety of chemical compositions of the silicate minerals and the resulting lack of any apparent simple stoichiometric formulae. When this is combined with the fundamental difficulty of studying silicates – they are virtually insoluble in anything other than hydrofluoric acid, and their thermal behaviour lacks any sharply-defined features – it is hardly surprising that virtually nothing was known about their classification until *X*-ray analysis became routine during the first half of the 20th Century. It is only when the full complexity of the possible structures of silicates, and their associated

Table 4.1 *Classification of clays by their uses and properties* (From Singer and Singer, 1963; 26, by permission of Chapman and Hall)

White-burning clays (used in whitewares)
(1) Kaolins: (a) residual; (b) sedimentary.
(2) Ball clays.
Refractory clays (fusion point above 1 600 °C but not necessarily white burning)
(1) Kaolins (sedimentary).
(2) Fire clays: (a) flint; (b) plastic.
(3) High alumina clays: (a) gibbsite; (b) diaspore.
Heavy clay products (low plasticity but containing fluxes)
(1) Paving brick clay and shales.
(2) Sewer-pipe clays and shales.
(3) Brick and hollow tile clays and shales.
Stoneware clays (plastic, containing fluxes)
Brick clays (plastic, containing iron oxide)
(1) Terra-cotta clays.
(2) Face and common brick.
Slip clays (containing more iron oxide)

propensity for substitution and the formation of solid solution series, has been understood, that some insight can be gained into the properties and behaviour of these minerals.

In order to study in more detail the clay minerals, it is first helpful to review briefly the basic structural classification of the silicates in general. Although ultimately complicated, the general progression is logical, and is based on the degree of 'polymerization' of the basic structural unit which is the SiO_4 tetrahedron (see below). The sequence runs as follows:

(i) minerals containing isolated SiO_4 tetrahedra (or isolated groups of SiO_4 tetrahedra, including SiO_4 rings), which are bonded only to other cations;
(ii) minerals consisting of single infinite chains of linked SiO_4 tetrahedra;
(iii) minerals consisting of double infinite chains of linked SiO_4 tetrahedra;
(iv) minerals consisting of infinite sheets of SiO_4 tetrahedra (which includes most of the clay minerals); and
(v) minerals made up of infinite three dimensional frameworks of SiO_4 tetrahedra, the simplest of which is crystalline quartz.

Table 4.2 summarizes this basic classification and terminology of silicates. It is worth remembering here that the term 'infinite' refers to an atomic scale of measurement, and does not imply that the chains extend to infinity in the mathematical sense of the word! The terminology surrounding the silicates and clay minerals, in common with other mineralogical terms, has never been fully systematized, and so the names given are often ill-defined, or not unique to a structural class, and should be regarded as the mineralogical equivalent of common names in chemistry.

The basic structural unit of all silicate rocks is the tetrahedron formed by the tetrahedral co-ordination of four oxygen atoms around a single silicon ion. The Si–O bond is the most stable bond formed between any other element and oxygen, and it is therefore unsurprising that elemental silicon is not found naturally in the presence of oxygen. It is conventional to assume that the tetrahedron is held together by purely electrostatic forces as a result of the attraction between the Si^{4+} ion and the O^{2-} ions. Whilst this is not entirely true – the Si–O bond is thought to have about a 50% covalent character (Putnis, 1992), otherwise the tetrahedron would not have the shape it does – it is adequate to consider the bonding in silicates as an ionic phenomenon, and to regard the SiO_4 tetrahedron as being $[SiO_4]^{4-}$. The huge difference in ionic radii between the O^{2-}

Table 4.2 *Summary of silicate mineral classification*

Structure	*Formula of Silicate Group*	*Common Names*	*Si:O Ratio*	*Examples of Mineral Groups*
(i) Separate SiO_4 tetrahedra	$[SiO_4]^{4-}$	orthosilicates or nesosilicates	1:4 (0.25)	Olivines Garnets
Separate Si_2O_7 groups	$[Si_2O_7]^{6-}$	sorosilicates	2:7 (0.29)	Melilite $Ca_2MgSi_2O_7$
Separate silicate rings	$[(SiO_3)_n]^{2n-}$	metasilicates or cyclosilicates	1:3 (0.33)	Beryls $(Si_6O_{18})^{12-}$
(ii) Single chain silicates	$[(SiO_3)_n]^{2n-}$	inosilicates	1:3 (0.33)	Pyroxenes
(iii) Double chain silicates	$[(Si_4O_{11})_n]^{6n-}$	inosilicates	4:11 (0.36)	Amphiboles
(iv) Layer (sheet) silicates	$[(Si_2O_5)_n]^{2n-}$	phyllosilicates	2:5 (0.4)	Micas Clay minerals
(v) Framework silicates	SiO_2	tectosilicates	1:2 (0.5)	Quartz (feldspars)

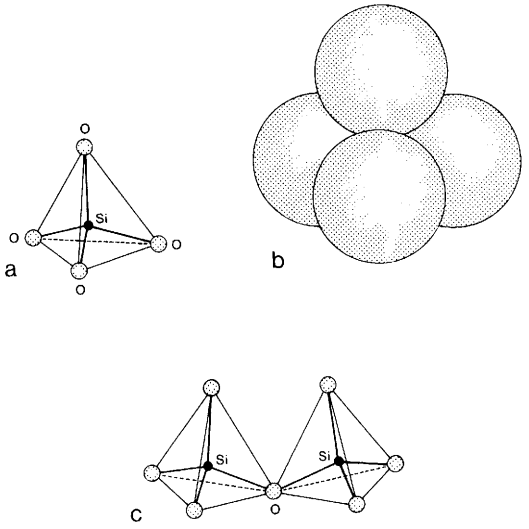


Figure 4.2 *Tetrahedral co-ordination of silica,* (a) *conventional model of [SiO_4] tetrahedra,* (b) *space-filling model of [SiO_4] tetrehedra;* (c) *model of linked tetrahedra [Si_2O_7]*
(From Putnis, 1992; Figure 6.1, by permission of Cambridge University Press)

(1.40 Å or 0.14 nm) and the Si^{4+} (0.42 Å, 0.042 nm) means that on the atomic scale the central silicon ion is virtually lost in the centre of the four co-ordinating oxygens – in fact, it sits comfortably in the tetrahedral hole formed in the centre of the four oxygen spheres (Figure 4.2). The sequence of silicate structures given above can therefore be interpreted as an increasing number of direct linkages between adjacent tetrahedra via a bridging (Si–O–Si) bond. Isolated tetrahedra have no shared corners, single chains share two, double chains alternately share two or three, sheets share three, and frameworks share four.

Minerals containing isolated single tetrahedra are referred to as *orthosilicates* or *nesosilicates.* Simple examples are the *olivine* and *garnet* minerals. The general formula for the olivines is M_2SiO_4, where M can either be Mg^{2+}, Fe^{2+}, Ca^{2+}, or Mn^{2+}. The structure of the olivine group of minerals is shown in Figure 4.3, where the SiO_4 tetrahedra are shown as triangles – those with full internal lines represent tetrahedra with the apex pointing up, those with dashed lines point down. Even with this simple silicate the actual unit cell structure of the mineral is complicated, but it is easy to see that the SiO_4 tetrahedra are not linked together other than by the intervening M^{2+} cations. The pure mineral Mg_2SiO_4 is

called *forsterite*, whilst Fe_2SiO_4 is called *fayalite*, which is an extremely important mineral archaeologically since it is one of the major constituents of iron smelting slag. *Forsterite* and *fayalite* form the end members of a complete solid substitutional series, with the stoichiometry $(Mg,Fe)_2SiO_4$, which implies that the cationic site can be randomly occupied by either Mg^{2+} or Fe^{2+}. This is the formula for the natural mineral *olivine*, after which the family takes its name. It should not be surprising, therefore to find that the fayalite phase of an ironworking slag contains a proportion of magnesium on analysis. This illustrates the facility with which silicate minerals can form solid substitution series, which is one of the chief difficulties to be faced if they are studied using only chemical information, and not structural data. Many other orthosilicates are known, but all have a silicon:oxygen ratio of 1:4. Thus garnets, with the general formula $A_3^{2+}B_2^{3+}Si_3O_{12}$ (where A^{2+} can be any divalent ion such as Ca^{2+}, Mg^{2+}, or Fe^{2+}, and B^{3+} is any trivalent

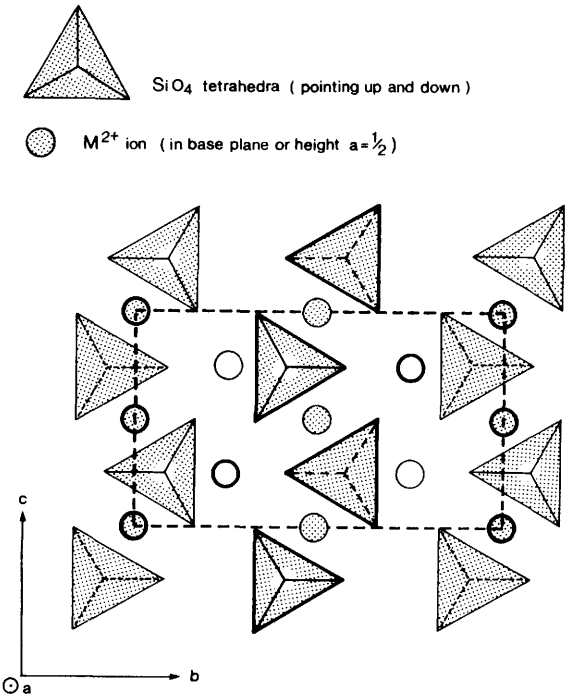


Figure 4.3 *Structure of the olivine family of minerals down the a axis of the unit cell* (After Putnis, 1992; Figure 6.4a, by permission of Cambridge University Press)

ion, such as Al^{3+}, Fe^{3+}, *etc.*) is still classified as an orthosilicate despite the fact that there are three silicons in the stoichiometric formula (and hence in the unit cell).

More complex isolated units can be formed by two tetrahedra linked together (*i.e.*, one corner shared, conventionally written as $[Si_2O_7]^{6-}$ and referred to mineralogically as *sorosilicates*). This dimeric tetrahedral structure is also shown in Figure 4.2. An example of this is the mineral *melilite* ($Ca_2MgSi_2O_7$), which is also a common constituent of vitreous slag. Minerals containing three-membered rings are known, with the structural unit $[Si_3O_9]^{6-}$, and the general formula for isolated rings of linked SiO_4 tetrahedra is $[(SiO_3)_n]^{2n-}$. These ring structures are known as *metasilicates* or *cyclosilicates*, and are characterized by a silicon:oxygen ratio of 1:3. An example of this type of mineral is *beryl*, $Be_3Al_2Si_6O_{18}$, where the large isolated six-membered Si_6O_{18} rings are only linked together by either Be or Al cations, thus defining the mineral as a metasilicate.

All of the above therefore, despite the increasing size of the silicate group, are examples of silicate minerals containing isolated SiO_4 tetrahedra or groups of tetrahedra. The general rules observed in the polymerization of silicates are that linking of the tetrahedra only occurs by sharing corners (*i.e.*, *via* an Si–O–Si bridging bond), and that only two tetrahedra may share a corner (this follows from the divalency of the oxygen). Figure 4.4 shows how individual SiO_4 tetrahedra can link together to form infinite chains if these rules are followed. If the tetrahedra share two corners (Figure 4.4a), the result is an infinite single chain, termed an *inosilicate*. Each tetrahedron has two bridging and two non-bridging corners (oxygens), giving an Si:O ratio of 1:3 since each shared oxygen contributes half towards the count for each tetrahedron. The resulting general formula is $[(SiO_3)_n]^{2n-}$, the same as for metasilicates (ring silicates), and the net charge imbalance per tetrahedron is 2^- (equivalent to two unsatisfied valencies, one on each of the non-bridging oxygen ions). On average, therefore, in order to maintain electrical neutrality (which is a requirement for all crystal structures) each tetrahedron should be associated with two positive charges (*i.e.*, a single divalent cation such as Ca^{2+}, or two monovalent cations such as Na^+). This leads to minerals of the *pyroxene* group, such as *enstatite* ($MgSiO_3$), in which the single silicate chains stack together lengthwise with alternately the apices and the bases pointing towards each other, and the magnesium cations joining pairs of chains together (Figure 4.5). The resulting crystal habit, as might be expected from minerals with an infinite chain length in one direction, is often *acicular* or *needle-like*, and in fact the mineralogical term *inosilicate* is derived from the Greek *inos*, meaning fibre.

The double chain silicates (Figure 4.4b) are also termed *inosilicates*

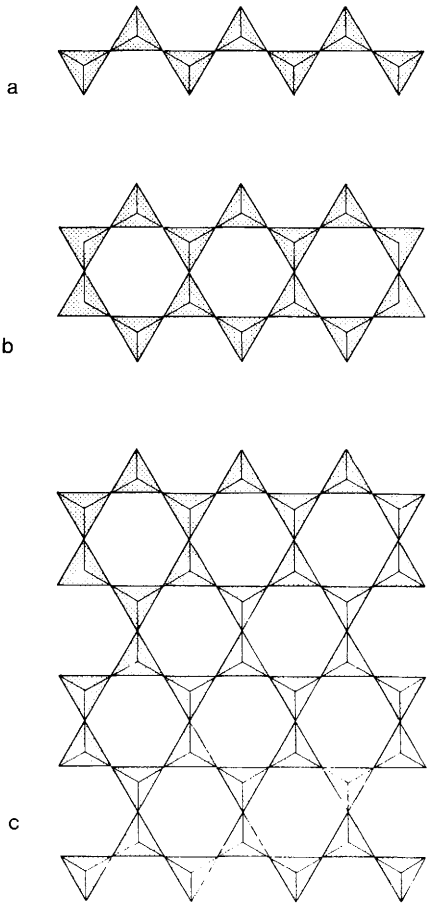


Figure 4.4 *Infinite chain silicates (single, double, and sheet):* (a) *infinite single chain silicate with two corners shared per tetrahedron (*pyroxene *structure);* (b) *infinite double chain, with alternately two and three corners shared (*amphibole *structure);* (c) *infinite sheet structure, with each tetrahedron sharing three corners (sheet silicates)*
(From Putnis 1992; Figure 6.3, by permission of Cambridge University Press)

because of their fibrous or needle-like qualities. The double chains can be thought of as two cross-linked single chains, with the link being made between the basal plane of each chain – in other words, as with the single chain, the apices of each individual tetrahedron are all pointing in the same direction. The net result is that alternate tetrahedra have two or three bridging corners, depending on whether they are on the

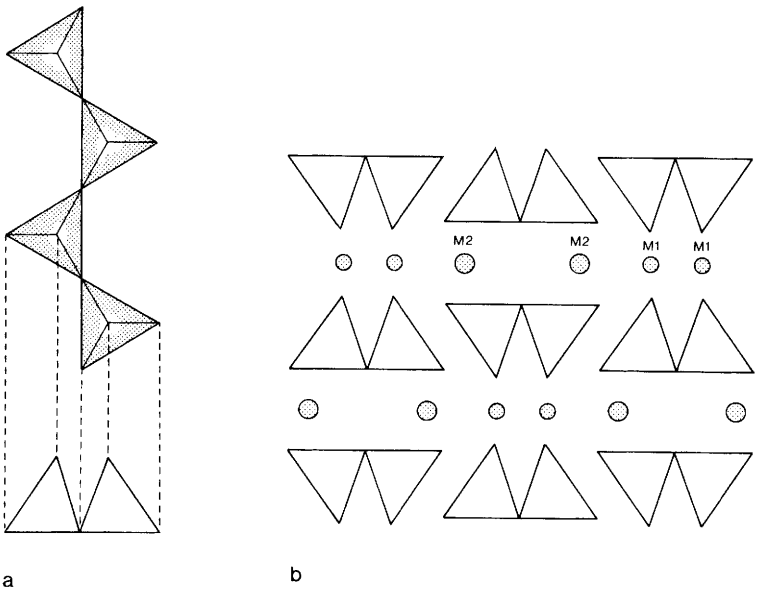


Figure 4.5 *Structure of pyroxene minerals:* (a) *demonstration of the end view of the single silicate chain;* (b) *end view of the stacking arrangement of single chains, showing the position of the metal cations. There are two different cationic environments, M1 and M2*
(After Putnis, 1992; Figure 6.11, by permission of Cambridge University Press)

'outside' of the double chain (with two links) or on the 'inside' (with three). The general formula is $[(Si_4O_{11})_n]^{6n-}$, and is the structure found in the *amphibole* group of minerals, which makes up a large group of rock-forming minerals. The chains are stacked together in a manner very similar to that described for the pyroxenes, with two double chains bridged apex to apex by cations and hydroxyl (OH^-) groups, and then the pairs linked together base-to-base by more cations.

The next step in this logical progression (Figure 4.4c) is to link together an infinite number of these double chains to produce a continuous sheet of SiO_4 tetrahedra, all linked in the basal plane and all sharing three corners, and all with their apices pointing in the same direction. These minerals are known as layer silicates, or *phyllosilicates* (from the Greek *phylon*, meaning leaf), and they have the general formula $[(Si_2O_5)_n]^{2n-}$. The sheets are stacked one on top of the other, alternately 'up' and 'down' (*i.e.*, with the apices pointing up or down). The two layers facing each other are held together by cations and hydroxyl groups, and the

resulting 'sandwiches' are held together by other cations. In general, the phyllosilicates are *laminar* minerals, and are characterized by being flaky or plate-like; they are analogous to graphite in carbon chemistry. Aluminium substitution for silicon is common, and these minerals are the starting point for a study of the aluminosilicate clay minerals, discussed below.

The final group of silicate structures are the framework silicates, in which all four corners of the tetrahedra are linked to another tetrahedron, to form an infinite three-dimensional network, analogous to the diamond structure in carbon chemistry. With no substitution, the general formula for these minerals is SiO_2, which is polymorphic (*i.e.*, has several crystalline forms), but is generally known as *quartz*. Aluminium substitution for silicon within the framework (which demands charge compensation by interstitial cations to counteract the resulting charge deficit) gives the large family of *feldspar* minerals, which are also very important rock-forming minerals.

As alluded to above, the chemistry of the silicate minerals is made infinitely more complex by the ability of certain elements to substitute for silicon within the structure. The most important substitution is Al^{3+} for the Si^{4+}; since Al^{3+} is only slightly larger than Si^{4+} (0.51 Å, compared to 0.42) and the tetrahedral co-ordination of the four oxygens is maintained (with some distortion of the bond angles), but obviously for each ion substituted there is a net charge deficit equivalent to one additional negative charge per substituted tetrahedron. This has to be compensated by an additional positive charge, supplied by a cation which has to be accommodated *interstitially* (*i.e.*, in the holes in the structure). This is most easily illustrated by considering the framework silicates, as indicated above. A pure quartz structure has an infinite framework of SiO_4 tetrahedra, each linked to each other via a corner (Si–O–Si) bond. All bonds are bridging, and the Si:O ratio is 1:2. If one silicon ion in four is substituted by an aluminium ion, the framework structure can be symbolized on average as being $MAlSi_3O_8$, where M is a monovalent cation such as K^+ (in which case the mineral becomes *orthoclase*). If one in two silicon ions are substituted, the general formula becomes $NAl_2Si_2O_8$, where N could be a divalent cation such as Ca^{2+}, in which case the mineral becomes *anorthite*. These aluminium-substituted framework silicates are called *feldspars*. Most naturally occurring feldspars have between 25 and 50% aluminium substitution, and their composition can be usually plotted on a ternary $KAlSi_3O_8$–$NaAlSi_3O_8$–$CaAl_2Si_2O_8$ (*orthoclase-albite-anorthite*) diagram. These general rules for silicon substitution by aluminium (and other small highly charged cations) apply to all the structural categories of silicate minerals

described above, and give rise to the full complexity of the subject. The standard text on the structure and properties of the rock-forming minerals (in five volumes) is that of Deer *et al.* (1962, 1963). One point to note is that the Al^{3+} ion can also enter the mineral as an interstitial cation as well as a substituent for silicon in the network. The aluminium ions in these two environments are not equivalent in terms of oxygen co-ordination, *etc.*, and it is important to distinguish between the two. For example, the mineral *muscovite* is one of the mica family of phyllosilicates, with aluminium substituted for about one atom in four of the silicons, but it also includes potassium and aluminium as cations between the aluminosilicate sheets. It is given the general formula $KAl_2(AlSi_3O_{10})(OH,F)_2$ to signify that some of the aluminium (that within the brackets) is substituted for one in four of the silicon ions in the sheets, and also that some of the aluminium exists between the sheets, together with potassium. The existence of both OH and F in the second set of brackets indicates that the hydroxyl group between the silicate layers can also be substituted at random by the fluoride ion. This example indicates the difficulty of classifying the aluminosilicate minerals from chemical analytical data alone, since a traditional chemical analysis of the mineral would give only the total amount of aluminium present, but would not identify the extent of the partitioning of the aluminium between the two sites, and it would therefore be impossible to deduce the stoichiometry of the compound from the analytical data without the additional structural information.

The clay minerals can now be discussed in terms of their relationship with the phyllosilicates (sheet silicates). It is important to keep clearly in mind here the difference between 'clay' – the substance which is dug out of the ground, and which may be a mixture of different clay minerals, together with various non-clay minerals (such as quartz, pyrite, *etc.*, which may also be the product of the weathering of parent rock), as well as unaltered rock fragments, and incorporated organic material (Grim, 1968) – and the clay minerals themselves, which are crystalline compounds of specified stoichiometry and structure. At this stage, we are only considering the structure of the clay minerals.

It is instructive to begin with the mineral *kaolinite* $Al_2(Si_2O_5)(OH)_4$, which does not have aluminium substituted into the silicon sheets, and is therefore relatively simple. It used to be thought that kaolinite was the only true clay mineral, and that natural clay deposits differed in having different proportions of kaolinite mixed with other finely-divided minerals, but this is not the case. Figure 4.4c shows part of the infinite sheet of SiO_4 tetrahedra which form the basis of the phyllosilicates and clay minerals. Since all the tetrahedra point in the same direction, the

result is a sheet with 'mountain peaks' (the non-bridging oxygens) rising above the basal plane, and these peaks make hexagonal patterns of oxygens. A hydroxyl ion sits in this hexagonal cradle like an egg in an egg box, forming a close-packed raft of oxygen (from the silica) and hydroxyl ions. On top of this layer lies an aluminium ion, with a further close-packed layer of hydroxyls above this (Figure 4.6). The aluminium ion, being small, in fact sits in the hollow formed between three touching spheres (two oxygen from the SiO_4 tetrahedra, and a hydroxyl) in the middle of the sandwich, with the top layer of hydroxyls sitting above it. Its nearest neighbours are four hydroxyls and two oxygens, and it is therefore in *octahedral co-ordination.* The simplest way of thinking about this structure is to imagine it as a tetrahedral layer of SiO_4 bonded to an octahedral layer of aluminium co-ordinated by (O + OH). The complete

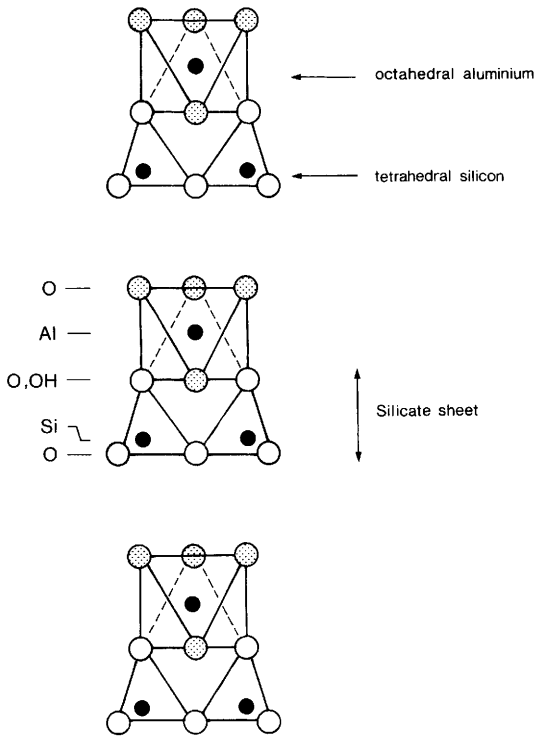


Figure 4.6 *Layer structure of kaolinite. Edge view showing the aluminosilicate sheets, and the stacking arrangement of these sheets*
(After Evans, 1966; Figure 11.11, by permission of Cambridge University Press and the author)

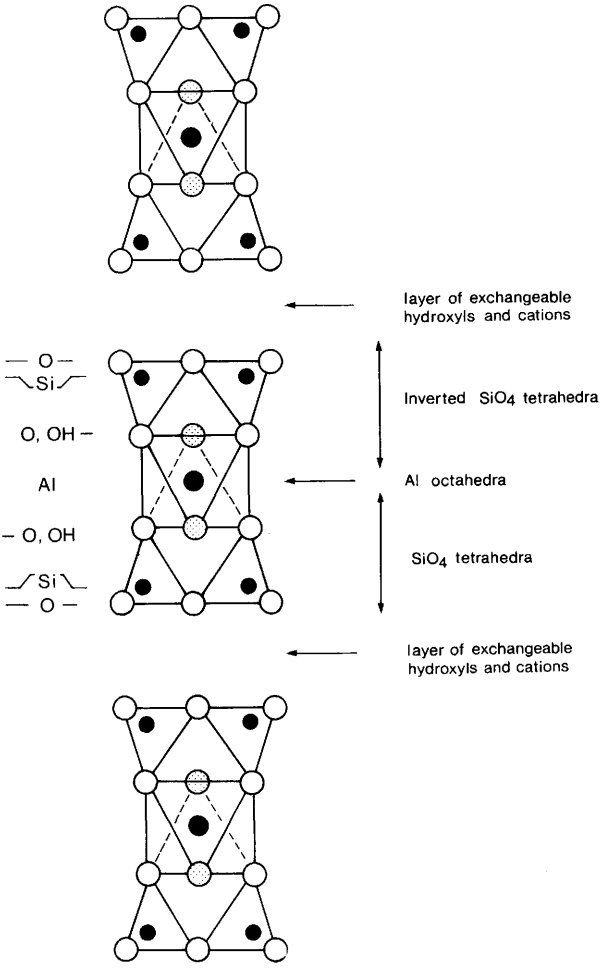


Figure 4.7 *Structure of one of the smectite group of clay minerals: montmorillonite,* $Al_2Si_4O_{10}(OH)_2.nH_2O$
(After Evans, 1966, Figure 11.08b, by permission of Cambridge University Press and the author)

structure is made up of stacks of these double layers but (unusually for phyllosilicates) they are all oriented in the same way, with the double layers held together by weak van der Waals (London) forces – hence the plasticity of such minerals.

In phyllosilicate terminology, kaolinite is called a 1:1 structure type, referring to the ratio of tetrahedral to octahedral layers (Brown, 1984). Most phyllosilicates and clay minerals are 2:1 structures, with the

sequence tetrahedral-octahedral-inverted tetrahedral (*i.e.*, kaolinite with an additional inverted layer of SiO_4 tetrahedra on top). This is illustrated by Figure 4.7, which shows an edge view of the structure of one of the *smectite* group of clay minerals, specifically *montmorillonite*, $Al_2Si_4O_{10}(OH)_2.nH_2O$ (which again, for simplicity, has no Al substitution into the SiO_4 tetrahedra). The Al ion sits in the octahedral hole in the (O + OH) layer, as described above, but the upper layer, instead of being simply a layer of hydroxyls, is now an inverted version of the basal layer of SiO_4 tetrahedra and associated OH groups. The Al is still in octahedral co-ordination, this time by four oxygens (two from each of the SiO_4 tetrahedral layers) and two hydroxyls (one associated with each layer). These triple layers are stacked upon each other, but now a variable number of water molecules and loosely-bound (*exchangeable*) cations may now occupy this interlayer plane. This water is easily lost on heating, but can be regained by absorbing water, thus explaining one of the familiar properties of clays – shrinkage on drying, and swelling on rehydration. In montmorillonite, as in other clay minerals, the full complexity of the family is developed (i) by allowing Al to replace Si in the lattice (with the attendant necessary charge-balancing cations occupying holes in the structure, or the interlayer gap), and (ii) by replacing Al in the octahedral layer with other cations, commonly Fe^{2+} or Mg^{2+}, again with the necessity for an additional cation somewhere (usually the interplanar layer) to maintain electrical neutrality. It can be seen that the detailed structure becomes very complicated as more and more substitutions occur because of the demands of electrical neutrality, but the overall systematic structural classification of the clay minerals remains valid. Variations on this 2:1 structure account for the majority of phyllosilicate and clay mineral structures, including *talcs, vermiculites, micas*, and *chlorites*. For a full description of these structures see Deer *et al.* (particularly Volume 3, 1962), Grim (1968), or Brown (1984).

THE FIRING OF CLAYS AND THE MINERALOGICAL COMPOSITION OF CERAMICS

The discovery of the working properties of clays must have resulted in one of humankind's first expressions of representational art, roughly contemporaneous with the discovery of the colouring properties of natural pigments and their use in cave art. The additional discovery that the result of the manipulation of this art form could be rendered permanent by the use of fire must indeed have been a source of wonder. The earliest fired ceramic so far known is a small moulded figurine from Dolní Věstonice in what was Czechoslovakia, dated to approximately 24 000

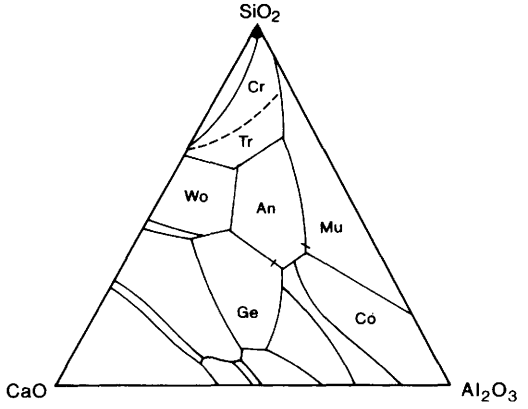
years BP (Vandiver *et al.*, 1989). By approximately 10 000 years ago, simple utilitarian vessels were being produced in the Near and Far East.

The value of clay as a raw material for the manufacture of vessels stems from two properties – its plasticity when wet, and its hardness when heated to a moderately high temperature. Most clays suitable for pottery making will fire at a temperature achievable in a small bonfire (very approximately 800 °C), although specialist products such as porcelain require temperatures in excess of 1200 °C to mature. Plasticity is the property whereby sufficiently wet clay can be deformed by modest pressure, and the clay will retain the given form when the pressure is removed. It stems largely from the sheet-like structure of the clay minerals, as described above. Most clay minerals incorporate water molecules as part of their structure between the composite layers of aluminosilicates, and it is the presence of these which allows the individual crystals to slide over each other and renders clay plastic. Plasticity varies from clay mineral to clay mineral – montmorillonite and kaolinite are highly plastic, whereas illite and dickite are considerably less so. The plasticity of a lump of clay therefore depends partially on the exact mixture and particle size distribution of the clay minerals within it, but also on the proportion of non-plastic minerals present, and the pH of the water (Singer and Singer, 1963).

The firing behaviour of clay minerals and commercially valuable clay deposits has of course been extensively studied (Singer and Singer, 1963; Kingery *et al.*, 1976; Rice, 1987). The key factor is the irreversible removal of the interplanar water molecules, but the study of this process is complicated by the multiplicity of sites which water may occupy within the structure – *adsorbed water*, weakly held on the surface of the crystals themselves, which is probably lost below 100 °C, *interplanar water* between the layers, relatively weakly held, and lost at a few hundred degrees, and *structural OH* groups which may require up to 1000 °C to remove them. The phenomenon of water loss is best studied using the various techniques of thermal analysis (Grim, 1968; Rice, 1987; see also Chapter 2), but is complicated by the fact that results may vary as a function of the degree of crystallinity of the mineral under study, and also depending on the rate of heating employed. The most important clay mineral thermal decomposition commercially is that of *kaolinite* to *mullite* ($3Al_2O_3.2SiO_2$), and consequently this is the best studied. In well-crystallized samples of kaolinite little happens up to 400 °C, but above that, differential thermal analysis shows a sharp endothermic peak beginning at about 400 °C and usually reaching a maximum at about 600 °C, which corresponds to the onset of the loss of OH groups from the lattice. By about 800 °C, virtually all of the water has gone. The resulting

mineral has been called *metakaolin*, and there is some discussion about its true nature, since it appears to have a slightly disordered structure (Murad and Wagner, 1991), but it is generally thought to have a similar structure to kaolinite, with the loss of OH causing the aluminium octahedral layer to become reordered into a tetrahedral (oxygen) co-ordination (Grim, 1968). Rehydration studies have shown that this phase can revert to kaolinite, depending to some extent on the degree of crystallinity of the kaolinite from which it was formed, so that this reaction cannot be regarded as truly irreversible. At higher temperatures, irreversible changes do occur. Differential thermal analysis shows an exothermic reaction occurring between 900 and 1000 °C, attributed to the formation of high temperature phases of silica (SiO_2) and alumina (γ-Al_2O_3), although there has been considerable debate about the nature of these phases, with some suggestion that minerals with a *spinel* structure can form (Brindley and Lemaitre, 1987). Above 1200 °C *mullite* ($3Al_2O_3.2SiO_2$) forms, together with *cristobalite*, which is the high temperature form of SiO_2 (stable above 1470 °C).

All the various clay minerals undergo a similar chain of dehydration reactions, ultimately ending in a number of high temperature aluminosilicate phases, usually with excess silica in the form of cristobalite (or some other high temperature silica phase). The complication arises from the presence of other cations in most clay minerals, which can result, for example, in phases such as *enstatite* ($MgSiO_3$) if large proportions of magnesium are present. Other substitutions will result in other mineral phases being formed, as documented by Grim (1968) or Brindley and Lemaitre (1987). From the point of view of a ceramicist in general (and an archaeological ceramicist in particular), what becomes important for practical purposes is not the high temperature thermal behaviour of individual clay minerals, but the interactions at high temperatures between the various phases present resulting from the decomposition of the various clay minerals present, and also the non-plastic inclusions which are likely to be present, such as carbonates or even organic material. In particular, the melting points of various phases become critical in high temperature reactions, since impurities in the clay usually mean that sufficient fluxes are present to melt (or sinter) at least some of the phases. In order to consider these, it is usual to consider triangular phase diagrams, the most important of which for archaeologists is either the CaO–Al_2O_3–SiO_2 system for earthenware, or the K_2O–Al_2O_3–SiO_2 system for higher fired stonewares and porcelains (Heimann, 1989). Examples of these diagrams are shown in Figure 4.8. Although in reality the composition of ancient ceramics is always more complex than the three oxides specified, it is usually sufficiently accurate to sum elements



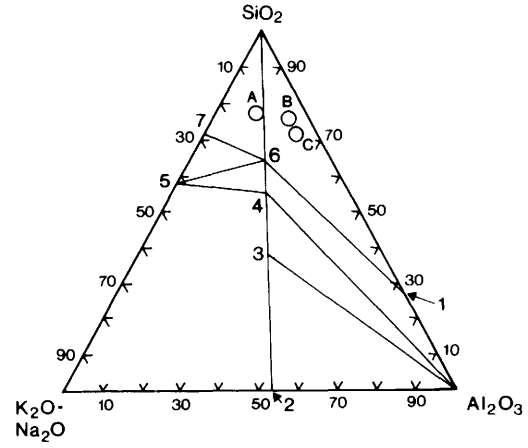


Figure 4.8 *Triangular diagrams of the systems CaO-Al_2O_3-SiO_2 and K_2O-Al_2O_3-SiO_2 (a) simplified liquidus surface in the system CaO-Al_2O_3-SiO_2 (the 'ceramic triangle'). Symbols: Cr = cristobalite; Tr = tridymite; Wo = wollastonite; An = anorthite: Mu = mullite; Ge = gehlenite; Co = corundum. (b) phase diagram K_2O-Al_2O_3-SiO_2. Symbols: 1 = $3Al_2O_3.2SiO_2$; 2 = $K_2O.Al_2O_3$; 3 = $K_2O.Al_2O_3.2SiO_2$; 4 = $K_2O.Al_2O_3.4SiO_2$; 5 = $K_2O.SiO_2$; 6 = $K_2O.Al_2O_3.6SiO_2$; 7 = $K_2O.4SiO_2$; A = Medici 'porcelain'; B = Thai stoneware; C = Rhenish stoneware*
(After Heimann, 1989; Figures 8 and 12, by permission of the author)

which behave similarly into one of the terms, in order to reduce the dimensionality of the problem to three, which can then be displayed graphically. For example, with earthenware (essentially any archaeological ceramic which is not stoneware or porcelain can be regarded as earthenware) it is usual to construct the 'CaO' value from the total 'fluxes' present – CaO + MgO + Fe_2O_3 + K_2O, which will then allow the composition to be plotted on the 'CaO'–Al_2O_3–SiO_2 ternary diagram, with any other unaccounted oxides contributing little to the total. The resulting plots allow the composition and proportion of phases present to be predicted with reasonable confidence, given that there is already an inbuilt assumption which in some circumstances can limit the usefulness of such an approach – the assumption of equilibrium conditions, which is inherent in the use of all phase diagrams. Nevertheless, using the usual techniques for interpreting phase diagrams such as the *Phase Rule* and the *Lever Principle* (described by Heimann, 1989), it is often possible to obtain a reasonably realistic picture of the phase structure and composition of fired ceramics.

It is always important when working with phase diagrams to be certain of the identity of the units used in the axes of the diagram, and the temperature at which it is applicable. Standard multicomponent phase diagrams used in ceramic technology usually deal with concentrations expressed as *weight percent oxide*, in which cases routine analytical data (which is normally quoted as 'weight percent oxide') can be used directly to calculate the coordinates of a point. Occasionally axes are labelled as *molar percent* (or *molar fraction*) of oxides – conversion between weight and molar units is relatively straightforward, involving dividing the weight percentage by the stoichiometric formula molecular weight, and re-normalizing to 100%. Slightly more complicated are those diagrams labelled not as weight percent oxide, but as weight (and sometimes molar) percentages of other stoichiometric compounds, usually using theoretical mineralogical formulae. This can be particularly useful if the mineralogy of the clay is known (or can reasonably be assumed). For example, in the study of the chemical compositions of the bodies of Chinese porcelains, Pollard and Wood (1984) demonstrated the value of converting the analytical data (expressed as weight percent oxides) into ternary co-ordinates using the assumed mineralogy of the raw materials – in this case quartz (SiO_2), kaolinite ($Al_2(Si_2O_5)(OH)_4$), potash mica (*muscovite* or *sericite*, $KAl_2(AlSi_3)O_{10}(OH)_2$), and soda feldspar (*albite*, $NaAlSi_3O_8$). In order to plot the data onto a triangular diagram, it is necessary to recalculate the weight percentages into these four components and then to normalize the three non-quartz components to 100% which then gives the triangular coordinates (see Figure 4.9). A sample

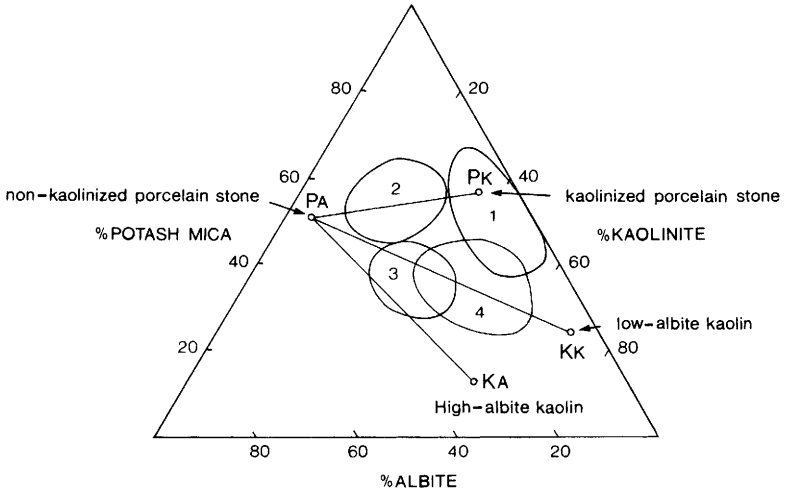


Figure 4.9 *Possible mineralogical constitution of Jingdezhen porcelain bodies. Average composition of four types of raw material are marked (P_A, P_K, K_A, K_K), and the approximate compositions of Chinese porcelains from four periods: 1 = predominantly 10th–12th C AD; 2 = 12th–13th C AD; 3 = predominantly 14th C AD; 4 = post-17th C AD*
(Pollard and Wood, 1984; Figure 10.7)

calculation, using the average body composition of 18 samples of Tianqi porcelain is given in Table 4.3 (see pp. 128–129). (Tianqi porcelain is a 17th C. AD oriental porcelain of disputed origin, but thought on the basis of these analyses to have been made at Jingdezhen in China rather than in Japan, as was originally postulated by Impey *et al.* (1983). In this way, it proved possible not only to predict the raw material mixtures used to produce the porcelain by comparison with the known mineralogies of the raw material sources (and changes in this pattern with time), but also to suggest a provenance using the major and minor element analyses, rather than the more usual trace element patterns (see below).

TRACE ELEMENT GEOCHEMISTRY IN CLAYS

Before turning to study the use of trace element data in ceramics for provenance studies, it is important to consider briefly some of the general principles of trace element geochemistry, so that we may look at the data from a more informed viewpoint. The literature on the geochemistry of the elements is vast, and much of our knowledge stems directly from the pioneering work of Goldschmidt (1954). Our theoretical understanding

of the behaviour of trace elements in the geochemical environment has progressed considerably since then, but the majority of this effort has gone into understanding the processes involved in the partitioning of trace elements between fluid and rock during the crystallization of the rock minerals, and only relatively recently has the behaviour of the trace elements in the sedimentary environment become of central interest, with the rising importance of environmental geochemistry. It is nevertheless useful to review briefly those factors which control the distribution of trace elements in geochemistry, since the general principles will apply to the sedimentary environment of clays, and allow us to predict some of the features we would expect in the chemistry of clays. Of particular interest in the context of clay is the *cation exchange capacity* (CEC), which is of great importance in governing the degree to which clays can pick up soluble elements from the surrounding aqueous environment. Finally, some attention must be given to the vastly-underesearched area of post-depositional alteration to fired archaeological ceramics.

The factors which control the distribution of trace elements (defined arbitrarily in geochemistry as those elements present at less than 0.1 wt%) can be discussed under a number of headings – *structural, thermodynamic, kinetic*, and, in the sedimentary environment, *solubility* and *speciation*. At the crudest level, the chemical elements can be divided up into three categories, from a consideration of the free energies of formation of the relevant oxides and sulfides. It is generally assumed that as it condensed the Earth separated into three liquid phases – metallic iron, molten iron sulfide, and a molten silicate phase – which rapidly solidified. The distribution of the elements between these phases is largely governed first of all by the stability of the appropriate oxides relative to that of iron oxide (FeO). All those elements which form more stable oxides were oxidized preferentially, whilst those forming less stable oxides remained in the reduced metallic state and were therefore concentrated in the molten iron core of the Earth – the so-called *siderophil* elements (principally Ni, Co, and the platinum and palladium group metals, as well as P, but this for other reasons). All of the other elements concentrated in the sulfidic or silicate regions of the Earth's crust, and are known as *chalcophil* elements. Some of these, essentially those with a greater affinity for oxygen than sulfur, concentrated in the *lithosphere*, and became the rock-forming (*lithophil*) elements – largely Mg, Al, Na, and Si. The remainder – Pb, Zn, Cd, Hg, Cu, Ag – concentrated in the sulfidic ores and have become the familiar primary ore metals from near-surface deposits.

Beyond this crude level of partitioning of the elements during the formation of the Earth, it is possible to consider the factors which control the partitioning of the elements during mineralization processes, such as

Table 4.3 *Sample of calculation of theoretical mineralogy of Tianqi porcelain*

Theoretical composition of minerals involved:

Sericite $KAl_2(AlSi_3)O_{10}(OH)_2$ or $K_2O.3Al_2O_3.6SiO_2.2H_2O$

Albite $NaAlSi_3O_8$ or $Na_2O.Al_2O_3.6SiO_2$

Kaolinite $Al_2(Si_2O_5)(OH)_4$ or $Al_2O_3.2SiO_2.2H_2O$

Quartz SiO_2

Calculate wt % of oxides in minerals:

	$\%SiO_2$	$\%Al_2O_3$	$\%K_2O$	$\%MgO$	$\%Na_2O$	$\%CaO$
Sericite	45.2	38.4	11.8			
Albite	68.8	19.4			11.8	
Kaolinite	46.6	39.5				
Quartz	100.0					

Average Composition of Tianqi Porcelain Body:

$\%SiO_2$	$\%Al_2O_3$	$\%K_2O$	$\%MgO$	$\%Na_2O$	$\%CaO$
67.2	25.5	3.47	0.25	1.7	0.94

Step 1. Calculate proportion of sericite in porcelain:
Assuming all K_2O comes from sericite: **% sericite** = 3.47/11.8 * 100 = **29.4%**
Calculate how much $SiO_2$and Al_2O_3 are contributed by sericite and subtract from original analysis:

	$\%SiO_2$	$\%Al_2O_3$	$\%K_2O$	$\%MgO$	$\%Na_2O$	$\%CaO$
Sericite	13.3	11.3	3.47			
Residual	53.9	14.2		0.25	1.7	0.94

Step 2. Calculate proportion of albite in porcelain:
Assuming all Na_2O comes from albite: **% albite** = 1.7/11.8 * 100 = **14.4%**
Calculate how much $SiO_2$and Al_2O_3 are contributed by albite and subtract from this residual analysis:

	$\%SiO_2$	$\%Al_2O_3$	$\%K_2O$	$\%MgO$	$\%Na_2O$	$\%CaO$
Albite	9.9	2.8			1.7	
Residual	44.0	11.4		0.25		0.94

Step 3. Repeat for kaolinite:
Assuming all residual Al_2O_3 comes from kaolinite: **% kaolinite** = 11.4/39.8 * 100 = **28.6%**

	$\%SiO_2$	$\%Al_2O_3$	$\%K_2O$	$\%MgO$	$\%Na_2O$	$\%CaO$
Kaolinite	13.3	11.4				
Residual	30.7			0.25		0.94

Step 4. Quartz: **% quartz** = 30.7/100 * 100 = **30.7%**

	$\%SiO_2$	$\%Al_2O_3$	$\%K_2O$	$\%MgO$	$\%Na_2O$	$\%CaO$
Quartz	30.7					
Residual	–	–	–	0.25		0.94

Predicted mineralogical composition of average Tianqi porcelain body:

	As calculated	*Normalized*	*Normalized without quartz*
Sericite	29.4	28.2	40.6
Albite	14.4	13.8	19.9
Kaolinite	28.6	27.4	39.5
Quartz	30.7	29.4	
Others	1.2	1.2	
Total	104.3	100.0	100.0

The predicted raw material composition is therefore:
Sericite = 28.2%
Albite = 13.8%
Kaolinite = 27.4%
Quartz = 29.4%

Predicted coordinates on triangular diagram are obtained by eliminating quartz and re-normalizing:
Sericite = 40.6%
Albite = 19.9%
Kaolinite = 39.5%

---

those which occur when rock-forming minerals crystallize from magma. The principal factors which determine the distribution of trace elements between different minerals, as first demonstrated by Goldschmidt, are those relating to the size and charge of the ionic species. By studying the tendencies of various pairs of elements to substitute for each other in different silicate minerals, it is possible to observe that similarity of ionic radius is the principal factor. This gives a whole series of pairwise associations, some of which are unsurprising, such as Mg^{2+} and Ni^{2+} (originally thought by Goldschmidt both to have radii of 0.78 Å, but now tabulated as 0.66 and 0.69 Å respectively; *e.g.* Aylward and Findlay, 1974; Table 4). This simply reflects the ability of one ion to physically replace another in the crystal lattice without causing undue distortion. Other pairs are a little less obvious, such as Na^+ and Ca^{2+} (radii now given as 0.97 and 0.99 Å respectively), or K^+ and Ba^{2+} (1.33 and 1.34 Å), where substitution can occur despite a difference in valency (which of course has to be balanced up elsewhere in the structure). Goldschmidt introduced the concept of '*ionic potential*' – the ratio of the ionic charge to

the ionic radius – into the theoretical discussion of the distribution of the elements, and his ideas have since been termed '*Goldschmidt's Rules*' for the relationship between the composition of a liquid and that of the crystals forming from it (*e.g.*, Henderson, 1982; 125). These can be summarized as:

(i) ions of similar radii and the same charge will be incorporated into the crystal in the same ratio as their concentration in the liquid;
(ii) an ion of smaller radius but the same charge will be incorporated preferentially into a growing crystal;
(iii) an ion of similar radius but with a higher charge will be incorporated preferentially into a growing crystal.

These rules form the basis of our understanding of the distribution of the elements in rocks and ores, which, at least in igneous rocks, has long been held to conform to a log-normal law (Ahrens, 1954). When combined with some knowledge of the kinetics of crystal growth, they can also be used for calculations of the temperature (and/or pressure) at which crystallization took place (*geothermometry* and *geobarometry*; Henderson, 1982; 175). This procedure may well have applications in some areas of archaeological chemistry (*e.g.*, in considering the partitioning of trace elements between slag and metal during metallurgical processes), but it has long been realized that these factors are of limited value in considerations of surface phenomena such as weathering, and in aqueous environments in general. Weathering takes place under a very limited range of temperatures, pressures, pH, E_h, and ionic concentrations, and is overwhelmingly conditioned by the behaviour of the various ions in aqueous media. The composition of surface waters is defined by the oxidation state of the ions in solution, the speciation of the ions, and the stability of various solid phases present with respect to the solution – all of which are controlled to a greater or lesser extent by the redox conditions of the environment. Furthermore, one cannot simply use the concept of the ionic radius as an indicator of the behaviour of a particular element, as is possible in higher temperature systems. A classic example of this is the chemical composition of seawater and the sediments which precipitate from it (Millot, 1970; 49). It is well known that the major dissolved species in typical seawater are Na^+ and Cl^- - effectively, seawater can usually be considered as a half molar solution of sodium chloride. In fact, the molar concentrations of sodium and potassium in seawaters are given as approximately 0.486 M and 0.010 M respectively (Henderson, 1982; Table 11.2), giving an approximate ratio of K/Na by weight of 1/28.5. The equivalent figure for fresh water

is approximately 1/10, indicating that sodium is by far the dominant cationic species in surface waters. In contrast, the ratio of K/Na by weight in surface rocks ranges from about 1/2.8 for schists to 1/7.7 for limestones, with an average of 1/3 for sedimentary deposits. It is clear that sediments are significantly enriched in potassium compared to surface waters, which runs counter to Goldschmidt's Rules. The explanation lies in the fact that in an aqueous medium the sodium ion is normally surrounded by a *hydration sphere* of H_2O molecules, whereas potassium ions are not (see Chapter 5). This gives an *effective radius* for these ions in water as nearer 1.83 Å for the sodium ion (compared to 0.97 Å without hydration), whereas potassium is calculated as being around 1.24 Å, compared to its 'normal' radius of 1.33 Å (Millot, 1970; 52). The sodium ion is therefore effectively considerably larger than the potassium ion, and is consequently discriminated against when crystallization and sedimentation occurs (as would be expected from Goldschmidt's Rules, providing the solvation effect is taken into account). Rubidium and caesium behave similarly to potassium, and are therefore similarly concentrated in aqueous sediments, whereas most of the other elements behave as does sodium.

It is clear from this simple example that the environment with which we are most concerned in archaeological chemistry is one in which the aqueous behaviour of the elements is of prime importance. Nevertheless, the geochemical behaviour of the elements is still subject to the thermodynamic and kinetic controls first postulated by Goldschmidt, and the association between elements in clays, sediments, and ultimately in archaeological ceramics is therefore in principle predictable. We should not, therefore, be surprised to find certain inter-element associations in the analysis of archaeological ceramics – in fact, we should be more surprised if we do not! Some common behaviour has been noted by other authors, *e.g.*, Glascock (1992), who observed that Ca is often correlated with elements such as Sr and Ba; Fe, Sc, and other transition elements are well known for exhibiting high correlations, and the rare earths (usually the lanthanides) exhibit strong correlations with each other. This latter point has been commented on geochemically in the archaeological literature (Allen *et al.*, 1975), bringing attention to the well known '*europium anomaly*' in rare earth geochemistry (arising from the fact that most lanthanides exist primarily in the 3^+ oxidation state, but europium can be reduced to 2^+ under natural conditions, giving it a slightly different pattern of behaviour). In general, however, discussion of these correlations is usually restricted to simple observations of their occurrence, and tends to focus on the undoubted influence of correlation on the multivariate data analysis techniques commonly employed (*e.g.*,

Bishop and Neff, 1989). These studies usually conclude that ceramic data cannot be analysed without taking correlation effects into account, and that what often distinguishes different clay sources is not the absolute concentrations of the various trace elements, but the degree of correlation between pairs of elements. The point is illustrated in Figure 4.10 (from Bishop and Neff, 1989), showing the correlation between Fe and Sc in three distinct clay sources. Two sources have approximately the same degree of correlation, but differ in absolute concentrations, but the third shows a different relationship between the two elements. This is simply a reflection of the different geochemical forces acting on a clay deposit as a result of deposition and weathering, and is a very good argument for attempting to understand these processes as part of the scientific approach towards the study of archaeological ceramics.

If the geochemical correlations between elements in clays are little discussed in the archaeological literature, then this paucity is nothing compared to the lack of consideration of post-depositional geochemistry on the composition of archaeological ceramics. Almost every discussion of ceramic provenance studies assumes without question that the composition of the ceramic vessel as recovered from the ground is unchanged from the composition of the newly fired vessel in antiquity. This is

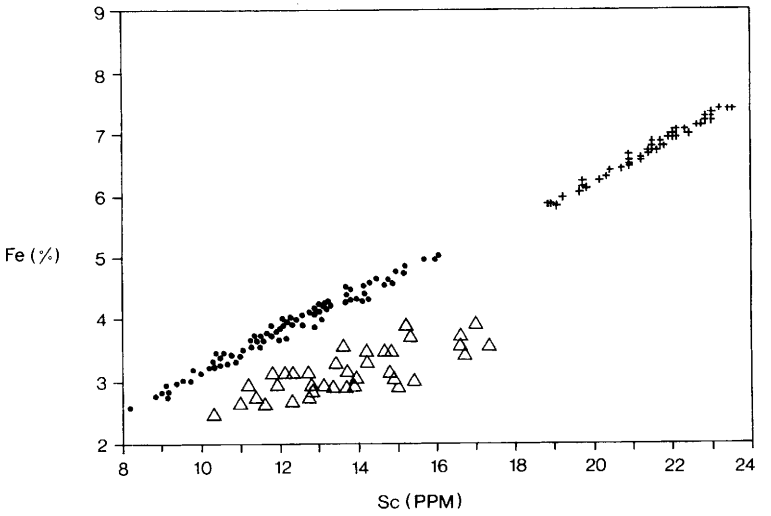


Figure 4.10 *Scatterplot of Fe and Sc values for three different pottery groups, showing the effect of correlation on the data*
(Redrawn with permission from Bishop and Neff, 1989; Figure 2. Copyright 1989, American Chemical Society)

unlikely to be so if there is any water around in the burial environment – a fact widely acknowledged in archaeological bone studies (*e.g.*, Hancock *et al.*, 1989), but largely unappreciated in pottery work. There are a few exceptions. In 1967, Freeth reported a study of some Bronze Age sherds from Lincolnshire, in which he compared the major and minor chemical element analysis (as determined by optical emission spectrometry) of several sherds from the same vessel, some of which were recovered from below the current water table, and some from above. He concluded that two of the oxides measured (CaO and MnO) showed conclusive evidence of post-depositional change, with the CaO being three times greater in the sherds found above the water table. The vessels were originally highly calcareous ('shell tempered'), and it is not clear to what extent these data can be taken as general, but there is a clearly some evidence for a potential problem. More recent work has been summarized by Freestone *et al.* (1985), which shows that the evidence from archaeological material is contradictory, as is the outcome of laboratory simulation experiments. Freestone *et al.* (1985) report the results of electron microprobe analyses of cross-sections from a range of archaeological samples, all of which showed evidence for enhanced P_2O_5 concentrations at the surfaces of the sherds, which is mirrored by the behaviour of CaO. They interpret this as evidence for extensive post-depositional alteration of a wide range of ceramic fabrics. A subsequent (but recurring) suggestion that enhanced levels of phosphate in pottery could be taken as a marker for the presence of residues of blood and fats (*e.g.*, Bollong *et al.*, 1993) has been vigorously (and in our view, correctly) rebuffed by Freestone *et al.* (1994).

Very little consideration has been given to mineralogical change during burial of ceramic material. Maggetti (1982) notes that low-fired ceramics (below 700 °C) in particular are susceptible to mineralogical change, with the secondary products being carbonates (particularly calcite, $CaCO_3$), hydrates (hematite – Fe_2O_3 – hydrating to goethite, $Fe_2O_3.H_2O$), hydrosilicates (zeolites and various clay minerals such montmorillonites), and hydrated sulfates such as gypsum ($CaSO_4.2H_2O$). Some consideration has also been given to the more aggressive chemical environment of seawater (Bearat and Dufournier, 1992), in which long term experiments have suggested that certain ceramics will lose calcium and strontium, and gain magnesium.

These studies clearly have important implications for the interpretation of analytical data from archaeological ceramics, and highlight the importance of petrological and mineralogical analysis to support the chemical data. They clearly show that, under some circumstances at least, archaeological ceramics may undergo considerable alteration, and

suggest that this must be a fertile area for further analytical and theoretical studies, probably using the techniques of geochemical groundwater computer modelling developed for use elsewhere in geochemistry (Pollard, 1993). Another profitable line of enquiry may be the application of the concept of *cation exchange capacity*, familiar for the study of clays, but only sporadically applied to fired ceramics (*e.g.*, Hedges and McLellan, 1976). This is an experimental measure of the proportion of exchangeable cations held by the clay (or ceramic), and may have implications for radiometric dating techniques, as well as for provenance applications using trace elements which might be gained or lost from groundwaters.

THE PROVENANCE OF ARCHAEOLOGICAL CERAMICS: ROMAN FINEWARES

The majority of chemical provenance studies carried out since the 1970s have utilized, for good theoretical reasons, trace element analysis by neutron activation analysis (NAA), and more recently inductively coupled plasma spectrometry (either ICP-AES, using an optical emission detector, or ICP-MS, where the plasma source is coupled to an even more sensitive mass spectrometer; see Chapter 2). A review of the chemical characterization of archaeological ceramics has recently been published by Neff (1992). The example illustrated here, which was carried out at the Research Laboratory for Archaeology and the History of Art, University of Oxford, has been specifically chosen to illustrate the possibility of using major and minor element analysis (via atomic absorption spectrometry (AAS), or even *X*-ray fluorescence, XRF). Although inherently less powerful than the trace element techniques for provenancing, this type of approach does give more insight into the nature of the clays used, and occasionally, as described above in connection with the Japanese porcelain study, allows something to be said about the nature of the raw materials being used. Rarely can this be deduced from trace element data.

'Rhenish' wares are a type of fine Roman pottery made in the Rhineland and Central Gaul (around the Lezoux area), and imported into Britain from the early 2nd to the late 3rd Centuries AD (Symonds, 1992). Their manufacture is closely related to that of *terra sigillata*, or samian – the most famous of the Roman finewares, typically with a dark red glossy surface slip. Both wares are characterized by a fine red fabric. 'Rhenish' wares often show a light or dark grey sandwich-like core in a cross section of the paste, and have a very glossy polished black, or dark red, or dark green colour-coating. They usually occur in the forms of beakers

and cups, and, less frequently, carafes and flagons. Typical decorations are rouletting, indenting, folding, and decorations '*en barbotine*' either under the colour-coating or painted on it in a white paint. Barbotine decoration under the colour-coating is more typical of Central Gaulish vessels, and sometimes includes figure-types commonly found on Central Gaulish *terra sigillata*. The white-painted decoration is more common to vessels from the Rhineland, where they are referred to as *spruchbechers*, or 'motto' beakers, as the decoration often consists of a repeating scroll underneath a 'motto', usually an exhortation to drink.

Chemical research on 'Rhenish' wares was initiated in order to study their distribution in Roman Britain, since it had been found to be very difficult to distinguish visually between vessels from the Rhineland and those from Central Gaul, especially with material from Roman-British occupation sites, which tend to produce small sherds, as opposed to whole vessels from graves. A comparison of published vessels from the Rhineland, France, and Britain suggested the existence of several superficial source indicators, but especially in the limited British literature these seemed speculative and perhaps inadequate. These indicators included: straight, pedestal bases on Central Gaulish vessels, versus curved, beaded bases on East Gaulish vessels; the absence of a grey, sandwich-like core in the fabric of the former, versus its presence with the latter vessels; and the decorations '*en barbotine*', which are under the colour-coating on the former vessels, and above it and in white paint on the latter vessels. Therefore a pilot programme of analysis of 'Rhenish' wares was initiated in Oxford, by AAS and XRF. This pilot programme consisted of twenty samples from several British excavations, divided into two groups thought to correspond to the two production areas according to the visual criteria described above. The results confirmed these two groups as chemically distinct and relatively internally homogeneous, as shown in Figure 4.11, where sections of two typical qualitative XRF spectra are shown. Most striking is the presence of As in the Central Gaulish vessel (estimated average value expressed as As_2O_3 around 0.01%), along with higher values of Zn, Rb, and Sr. The geochemistry of As is complicated by the fact that it exhibits variable valency (existing as 3+ and 5+), but it is usually associated with oxidizing primary mineralization, but can also be concentrated in reduced marine sediments. The general enhancement of trace elements in the Central Gaulish vessels may suggest a reducing marine environment for the deposition of the parent clay beds, but at this stage this is simple speculation.

It had long been suspected that 'Rhenish' wares are closely related to *terra sigillata*. Typologically, the relationship is closer with Central

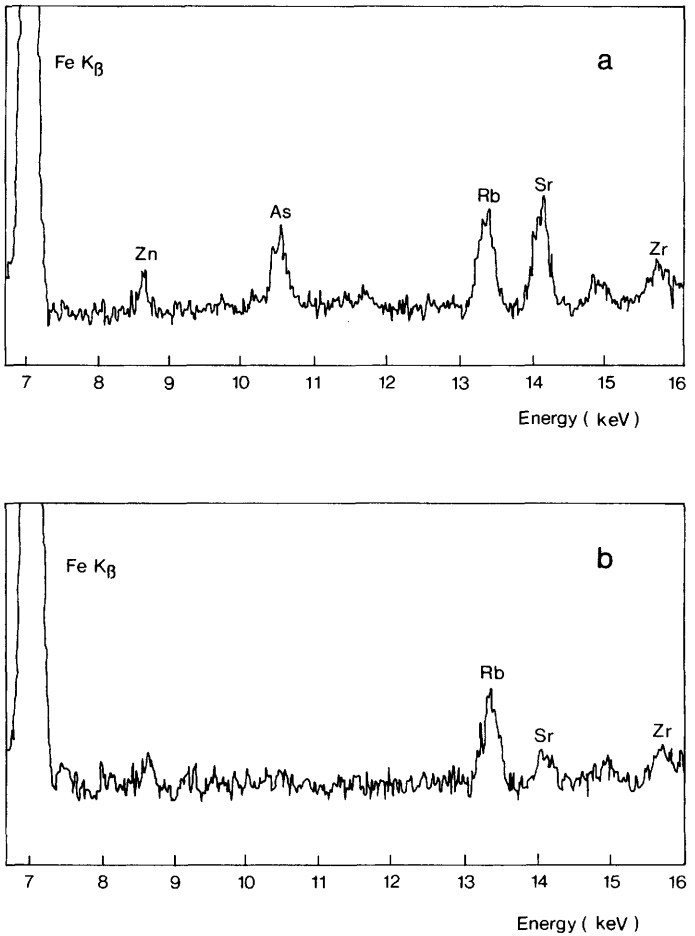


Figure 4.11 *Part of XRF spectra from two 'Rhenish' wares of different suspected origin* (a) *Alcester 76, attributed to Central Gaul;* (b) *Verulamium 58, attributed to Eastern Gaul*
(From Pollard *et al.*, 1981; Figure 1, with permission)

Gaulish wares, where identical applied figure-types are found on Central Gaulish *sigillata* and on Central Gaulish 'Rhenish' vessels. East Gaulish 'Rhenish' wares are less closely related to *sigillata*: the white-painted decoration is found on red vessels only rarely. 'Rhenish' wares are nonetheless the only pottery type made in the Rhineland of a quality similar to that of East Gaulish *sigillata* (or better), during the same period. *Terra sigillata* has been extensively studied, both typologically and

chemically (*e.g.*, Picon *et al.*, 1971, 1975). If it could be shown, therefore, that 'Rhenish' wares from either production area matched chemically with *sigillata* from any particular known workshop in that area, then it could reasonably be surmised that the 'Rhenish' wares and *sigillata* came from the same clay source and were therefore probably made in the same workshop. Because of the absence of stamps or of any obvious typological evolution, 'Rhenish' wares have always been rather broadly dated: if there is a direct link with *sigillata*, which is generally well dated, there ought to be a better chance of dating 'Rhenish' wares more closely. Further chemical analysis was therefore undertaken of fifty more samples of 'Rhenish' wares (from several British excavations) along with fifty samples of *sigillata* identified on typological grounds as being from Trier, Rheinzabern, Blickweiler, Chemery, La Madeleine, and Lavoye (all East Gaulish sites), Lezoux, and Les Martres de Veyre (Central Gaulish sites), and Colchester (the only then known British kiln making *sigillata*: the chemical composition of material from this site has subsequently been considered in more detail by Hart *et al.*, 1987). These results were discussed by Pollard *et al.* (1981). Subsequent work supplemented these sites with *sigillata* samples from the Central Gaulish sites of Clermont Ferrand and Toulon-sur-Allier, the 'Mid-Gaulish' sites of Gueugnon, Jaulges-Villiers-Vineux, and Domecy-sur-Cure (see map, Figure 4.12). These data were discussed in detail by Pollard *et al.* (1982).

The chemical data were analysed using the multivariate techniques developed at the time in Oxford (Pollard, 1986). In the first study, principal components analysis (PCA) and cluster analysis showed that the Lezoux *terra sigillata* and the supposed Central Gaulish 'Rhenish' wares grouped together, well-separated from the East Gaulish material – principally distinguished by higher values of SrO, CaO, As_2O_3, and lower NiO, K_2O, and Cr_2O_3 (Figure 4.13). The East Gaulish vessels occupy a different region of the plot, with the Trier and Rheinzabern groups distinguished from each other, with all of the East Gaulish 'Rhenish' vessels associated with the Trier material. It was concluded, therefore, that not only were Trier and Lezoux the principal production centres for the 'Rhenish' wares imported into England, but also that the same clay was being used for black and red gloss vessels. What was also important archaeologically was that it also supported the separation of the 'Rhenish' vessels into East and Central Gaulish groups on the grounds of the macroscopic criteria described above, meaning that pottery specialists could continue to use these criteria with a greater degree of confidence. Unfortunately, however, a large number of sherds are found which cannot be classified by these visual means. In these cases, chemical analysis appears to be the only reliable guide to

Figure 4.12 *Map of Gaulish* terra sigillata *production centres*
(From Symonds, 1992, by permission of Oxford University Committee for Archaeology)

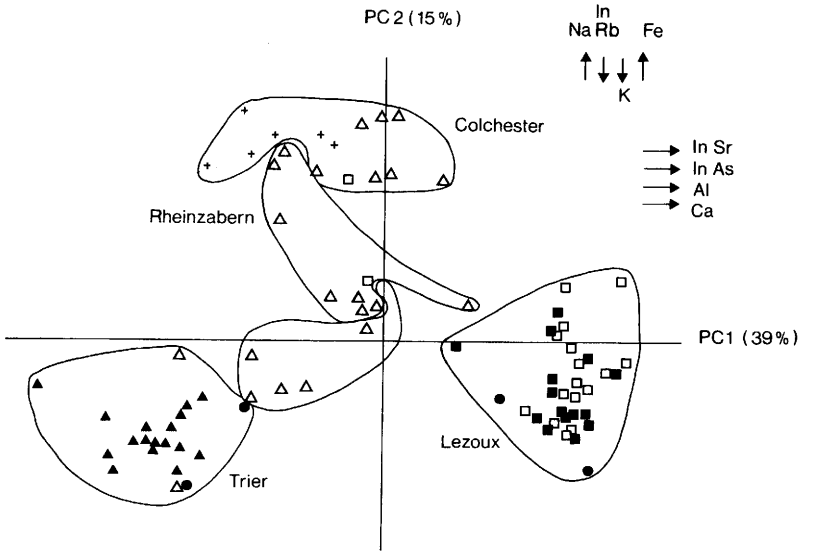


Figure 4.13 *Cluster analysis of Rhenish and* sigillata *analyses, showing association between Lezoux* sigillata *and Central Gaulish 'Rhenish', and Trier* sigillata *and East Gaulish 'Rhenish'*
Key: ■ *Central Gaulish* sigillata; □ *Central Gaulish 'Rhenish';*
△ *East Gaulish* sigillata; ▲ *East Gaulish 'Rhenish';*
\+ *Colchester* sigillata; ● *uncertain 'Rhenish'*
(From Pollard *et al.*, 1981; Figure 4, with permission)

provenance. Only a handful of the 85 'Rhenish' vessels analysed did not appear to originate from these two centres – these may be associated with minor production centres (minor, that is, in terms of supply to Britain), and this encouraged us to extend the study a little further, in order to resolve this problem.

Another notable result from this work was the somewhat surprising success of the XRF technique, which appeared in this case to provide a rapid, non-destructive method for distinguishing sherds from East and Central Gaul, both 'Rhenish' and *terra sigillata*. A 60 second irradiation of the cleaned edge of a sherd was sufficient to produce a spectrum of the type reproduced in Figure 4.11, and on the basis of this an opinion could be given as to the source of the vessel. It is not, of course, possible to distinguish between the various East Gaulish production sites on the basis of such a quick test, but with a small plain body sherd such information has on several occasions proved useful. Ancillary work indicated that the same trace element characteristics were possessed by the surface gloss as

was shown by the body fabric of vessels from both sources (as might be expected if the slip was made from the finest fraction of the clay from the same source as that used for the body fabric), and this offers a useful source indicator when even the minimal damage caused by cleaning the edge is not permissible. A similar suggestion for the discrimination of *sigillata* from Arezzo, La Graufesenque, Lezoux, and Lyon using a plot of the intensity ratios of Ti/Mn versus Fe/K measured on the surface has subsequently been made by Baillié and Stern (1984).

This preliminary work left six of the 85 fragments of 'Rhenish' ware found in Britain unaccounted for. Could any of these have come from any of the smaller Gaulish production centres known to be producing vessels of similar type and quality? It had already been shown that it is possible to distinguish chemically between *sigillata* from closely adjacent kilns, such as Lezoux and Les Martres de Veyre (Picon, unpublished), and so it should be possible to discriminate between the other comparable products from these sources. The problem was approached by analysing separately the analytical data from the three regions under consideration – Central Gaul, 'Mid Gaul' and East Gaul, or the Rhineland. Simplified dendrograms from the cluster analysis of these three data sets are reproduced in Figure 4.14. The first area to be considered in more detail was the Central Gaulish area around Lezoux, including Les Martres de Veyre, Clermont-Ferrand, and Toulon-sur-Allier (see map, Figure 4.12). Unfortunately, no samples were obtainable from the suspected workshops at Neris-les-Bains, St Bonnet/Yzeure, Lubié-Lapalisse, or Vichy, but a selection of vessels were analysed from the museum at Moulins, which might be expected to contain examples of the work of most of the local kilns. The data set contained 96 samples, including the 'Rhenish' wares found in England and attributed to Central Gaul. Cluster analysis of the analytical data identified four distinct groups, each one attributable to one of the four kilns represented, confirming that the products are indeed chemically distinct, and suggesting a secondary attribution for some of the vessels imported into England from this area (Figure 4.14a). The distinction between the groups attributed to Lezoux and Les Martres de Veyre is based on the Cr, Mg, and Ca content, but several of the Les Martres 'Rhenish' wares were to be found in the Lezoux group. It was originally suspected that some of the Les Martres sherds might in reality be products of the Lezoux kilns, and this appears to be the case. The group labelled Clermont-Ferrand consists mainly of vessels from Moulins museum, but the presence of three Clermont-Ferrand sherds suggests this to be the source, despite the fact that the Lezoux group also contains two (possibly three) of the vessels also thought to be manufactured in Clermont-

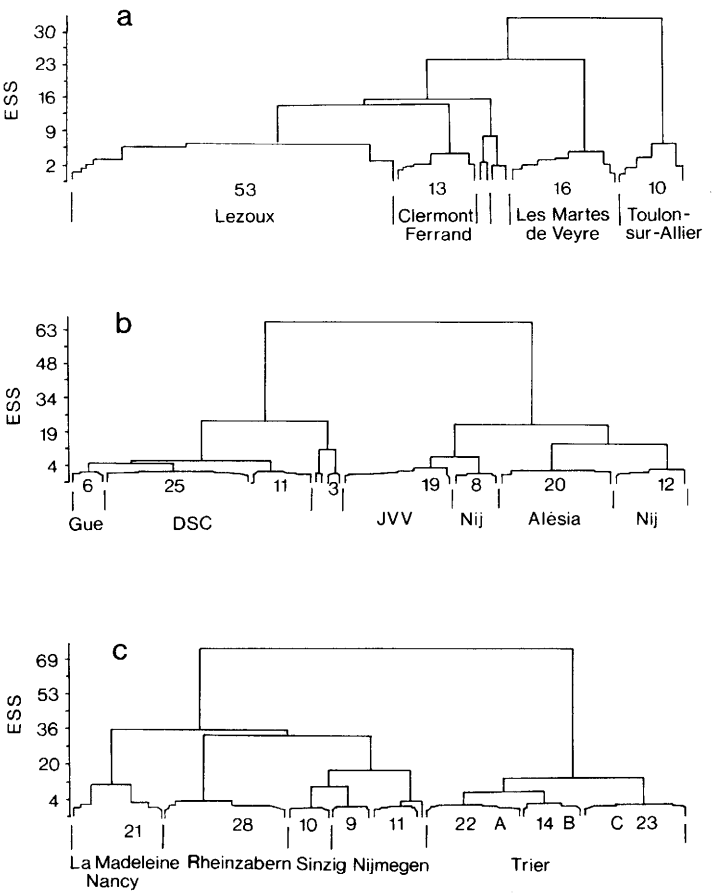


Figure 4.14 *Dendrograms of chemical data from 'Rhenish' wares and* sigillata *from three regions of Gaul: a) Central Gaul, b) Mid Gaul, c) East Gaul*
(From Pollard *et al.*, 1982; Figures 2–4, with permission)

Ferrand. This group is interesting however, because it contains the only British finds not attributed to Lezoux – fragments of three vessels, all found at Milton Keynes.

The second group was attributed to an area loosely described as North Central Gaul or Mid Gaul – the region around Alésia (Figure 4.14b). From a study of the material in French museums, it is obvious that several groups of black colour-coated vessels of similar quality to 'Rhenish' wares were being produced in regions other than Central and East Gaul. A good example of this was found at the Musée Alésia, which

contains a set of high quality vessels characterized by tall, everted rims, and excised decoration. The nearest known workshops producing *sigillata* are Gueugnon and Jaulges-Villiers-Vineux, and these, together with Domecy-sur-Cure, are also thought to have been making black colour-coated vessels. Samples were obtained from these three workshops and from the Museum to determine whether (a) the museum samples were made at Lezoux and (b) if not, whether they were from one of the neighbouring *sigillata* production centres.

Subsequently, results from the analyses of fourth Century colour-coated vessels from Nijmegen, originally obtained for comparison with Trier, were added to the data set, because preliminary clustering of all the data suggested that they were more closely related to the Mid Gaulish material. The dendrogram resulting from Ward's method of cluster analysis is more difficult to interpret than that for Central Gaul, but nine clusters appear to be significant, seven of which are identifiable as associated with Gueugnon (Gue), Domecy-sur-Cure (DSC), Jaulges-Villiers-Vineux (JVV), Alésia, and two groups containing material only found at Nijmegen (Nij). There is some exchange of samples between the Jaulges-Villiers-Vineux group and that attributed to Gueugnon, but the chemical differences are quite clear, and we may suggest on the basis of chemical analysis that the original attributions are incorrect.

The final area to be studied was East Gaul (Figure 4.14c). This had two purposes, one of which related to the Nijmegen vessels, with their implications for fourth Century production at Trier after the decline around 270 AD. This has been discussed above, and a Trier origin for these vessels has been discounted on chemical grounds. The second purpose, by analogy with the work on Central Gaul, was to confirm that all the vessels labelled as East Gaulish 'Rhenish' in Britain did in fact originate in Trier, and not from any other East Gaulish *sigillata* centre. Unfortunately, most of the minor *sigillata* kilns are only represented by two or three sherds, and this is not enough to define a chemical signature for these kilns – they appear to have been classified with either Trier or Rheinzabern, and this is purely an artefact of the analysis, resulting from the much larger numbers of Trier and Rheinzabern samples included. The classification suggested nine significant clusters, identified as La Madeleine A, La Madeleine B, Rheinzabern, Sinzig, Nijmegen A, Nijmegen B, and three groups from Trier, labelled A, B, and C. The archaeological interpretation of these results gives rise to several difficulties. There is nothing associated with the group labelled 'Trier A' to confirm this attribution, although none of the *sigillata* associated with this group give convincing proof of another source. The close chemical proximity of this group to the other two clusters identified with Trier on

the dendrogram is the sole evidence for a Trier origin. Two scenarios are possible – (a) the group labelled Trier A does not originate in Trier, but from some other chemically very similar kiln site, or (b) the production of 'Rhenish' and *sigillata* at Trier was not as closely related at Trier as at Lezoux, resulting in slight differences in the fabric. There could also be a temporal difference in the production of the two types of vessel, with some associated slight change of clay source. A third explanation is also possible, relating to the clustering procedures used. Ward's method can sometimes yield unacceptable divisions in dense data, arbitrarily dividing up large groups into a number of clusters which are approximately spherical in multivariate space (Pollard, 1982). Other clustering algorithms, such as average linkage (ALCA), when applied to this data set do not divide the Trier material in this way, but neither do they separate clearly the groups attributed to Rheinzabern, Sinzig, and Nijmegen, indicating the sensitivity of such interpretations to the exact data analysis procedures used. A time variation in the composition of the ceramics produced at such long-lived production centres, as in the *sigillata* production at Lezoux (Picon, 1971), has been postulated, but there is no firm evidence for this at Trier, since none of the *sigillata* analysed from Trier show any tendency to split into different compositional groups. (Subsequent work on Lezoux *sigillata* (Argyropoulos, 1992) has suggested that once a satisfactory raw material recipe had been obtained for *sigillata,* the clay sources used by at least six workshops in Lezoux remained constant throughout the second Century AD.) The resolution of this problem is important to our understanding of imported Roman finewares in Britain, since half of the East Gaulish 'Rhenish' wares from England fall into this 'Trier A' group (the other half being securely associated with Trier *sigillata* in the 'Trier C' group). It is clearly a matter of some concern to decide whether we are dealing with an artefact of the data analysis employed, the products of a minor *sigillata* kiln in the Trier region, or a time distinction in the composition of the products from Trier itself. This is a matter which can only be resolved by further analysis of well-provenanced material, and serves to illustrate the usual outcome of chemical provenance work – some questions can be confidently answered, but others are inevitably left undecided.

SUMMARY

The example of chemical provenancing briefly described here is atypical of the wider field of analytical studies of archaeological ceramics, but does illustrate both the power of the technique and some of the potential

problems, particularly in the interpretation of the data. In terms of the problem of identifying the Continental sources of 'Rhenish' wares found on British sites, the work has been highly successful. With very few exceptions, all the samples analysed came from either Lezoux or, in all probability, Trier, and the clear chemical discrimination allowed the typological criteria to be examined from a firm base. The clear compatibility of 'Rhenish' wares with *sigillata* (at least at Lezoux) has allowed the archaeological study of 'Rhenish' vessels to proceed at a much faster pace by comparison with the much better known *sigillata*, and led us to believe that many of the smaller 'Rhenish' and *sigillata* production sites in France could be chemically distinguished. The chemical characteristics of some of the samples found at Milton Keynes compared very closely with those of samples from Clermont-Ferrand, and these are, so far, the only vessels of this type found in Britain not to be attributed to either Lezoux or Trier. This is a remarkable tribute to the value of chemical studies of archaeological ceramics, when combined with an archaeologically coherent research design.

Methodologically, there are inevitably some questions which must always be considered. Do the 'control sherds' adequately represent the output of all of the relevant vessels from all of the relevant kiln sites? There is always potentially the problem of 'kiln X' – an unknown (or at least, uncharacterized) kiln which is producing wares similar to those under consideration, but not included either because it has not been identified archaeologically, or because it was not known to produce the type of ware under question. This problem should be at least considered since a geochemical knowledge of the type of clay deposits being exploited may restrict the geographical extent of the possible clay sources. The question of the representativeness of the samples is more difficult. Most kiln sources are characterized by at best a hundred analyses, probably covering a range of 'qualities' of vessels produced, a range of vessel types, and, possibly, a time span of tens or even hundreds of years. Essentially the archaeological chemist is relying on the quality control procedures in force in antiquity to ensure that the sample is representative of the range of compositions produced! It is undoubtedly essential to match as closely as possible the fabric, shape, quality, and date of the control sample with the unknown material. Again, this emphasizes the need for extremely good archaeological definition of a project to give any chance of success. Finally, of course, the data analysis needs to be handled with extreme care, from a knowledge of the propensities of the different types of approach which might be applied. It is all too easy to rely on one technique, just because it gives an easily-interpreted picture.

REFERENCES

Ahrens, L.H. (1954). The log-normal distribution of the elements 1 (A fundamental law of geochemistry and its subsidiary). *Geochimica et Cosmochimica Acta* **5** 49–75.

Allen, R.O., Luckenbach, A.H. and Holland, C.G. (1975). The application of instrumental neutron activation analysis to a study of prehistoric steatite artifacts and source material. *Archaeometry* **17** 69–83.

Argyropoulos, V. (1992). *Chemical Studies of the Roman Samian Pottery Industry of Central Gaul.* Unpublished Ph.D. Thesis, Department of Archaeological Sciences, University of Bradford.

Aylward, G.H. and Findlay, T.J.V. (1974). *S.I. Chemical Data*, John Wiley, Chichester (2nd edn.).

Baillié, P.J. and Stern, W.B. (1984). Non-destructive surface analysis of Roman terra sigillata: a possible tool in provenance studies? *Archaeometry* **26** 62–68.

Bearat, H. and Dufournier, D. (1992). Ceramics alteration due to contact with seawater. Paper presented at *Archaeometry 92*, 28th International Symposium on Archaeometry, 23–27th March 1992, Los Angeles (abstract only).

Bishop, R.L. and Lange, F.W. (eds.) (1991). *The Ceramic Legacy of Anna O. Shepard.* University Press of Colorado, Boulder, Colorado.

Bishop, R.L. and Neff, H. (1989). Compositional data analysis in archaeology. In *Archaeological Chemistry IV*, ed. Allen, R.O., American Chemical Society Advances in Chemistry Series 220, Washington DC, pp. 57–86.

Bollong, C.A., Vogel, J.C., Jacobson, L., van der Westhuizen, W.A. and Sampson, C.G. (1993). Direct dating and identity of fibre temper in pre-contact Bushman (Basarwa) pottery. *Journal of Archaeological Science* **20** 41–55.

Brady, N.C. (1974). *The Nature and Properties of Soils.* Macmillan, New York (8th edn.).

Bragg, W.L. (1937). *The Atomic Structure of Minerals.* Cornell University Press, New York.

Brindley, G.W. and Lemaitre, J. (1987). Thermal, oxidation and reduction reactions of clay minerals. In *Chemistry of Clay and Clay Minerals*, ed. Newman, A.C.D., Mineralogical Society Monograph No. 6, Longmans, London, pp. 319–370.

Brown, G. (1984). Crystal structure of clay minerals and related phyllosilicates. In *Clay Minerals: Their Structure, Behaviour and Uses*, eds. Fowden, L., Barrer, R.M. and Tinker, P.B., Royal Society, London, pp. 1–20.

Deer, W.A., Howie, R.A. and Zussman, J. (1962, 1963). *Rock-Forming Minerals*. Longmans, London (5 vols.).

Evans, R.C. (1966). *An Introduction to Crystal Chemistry*. Cambridge University Press, Cambridge (2nd edn.).

Freestone, I.C., Meeks, N.D. and Middleton, A.P. (1985). Retention of phosphate in buried ceramics: an electron microbeam approach. *Archaeometry* **27** 161–177.

Freestone, I.C., Middleton, A.P. and Meeks, N.D. (1994). Significance of phosphate in ceramic bodies: discussion of paper by Bollong *et al.* *Journal of Archaeological Science* **21** 425–426.

Freeth, S.J. (1967). A chemical study of some Bronze Age sherds. *Archaeometry* **10** 104–119.

Glascock, M.D. (1992). Characterization of archaeological ceramics at MURR by neutron activation analysis and multivariate statistics. In *Chemical Characterization of Ceramic Pastes in Archaeology*, ed. Neff, H., Prehistory Press, Madison, Wisconsin, pp. 11–26.

Goldschmidt, V.M. (1954). *Geochemistry*. Clarendon, Oxford.

Grim, R.E. (1968). *Clay Mineralogy*. McGraw-Hill, New York (2nd edn.).

Hancock, R.G.V., Grynpas, M.D. and Pritzker, K.P.H. (1989). The abuse of bone analyses for archaeological dietary studies. *Archaeometry* **31** 169–179.

Hart, F.A., Storey, J.M.V., Adams, S.J., Symonds, R.P. and Walsh, J.N. (1987). An analytical study, using inductively-coupled plasma (ICP) spectrometry, of samian and color-coated wares from the Roman town at Colchester together with related Continental samian wares. *Journal of Archaeological Science* **14** 577–598.

Hedges, R.E.M. and McLellan, M. (1976). On the cation exchange capacity of fired clays and its effect on the chemical and radiometric analysis of pottery. *Archaeometry* **18** 203–207.

Heimann, R.B. (1989). Assessing the technology of ancient pottery: the use of ceramic phase diagrams. *Archeomaterials* **3** 123–148.

Henderson, P. (1982). *Inorganic Geochemistry*. Pergamon Press, Oxford.

Impey, O.R., Pollard, A.M., Wood, N. and Tregear, M. (1983). An investigation into the provenance and technical properties of Tianqi porcelain. *Trade Ceramic Studies* **3** 102–158.

Jones, R.E. (ed.) (1986). *Greek and Cypriot Pottery: A Review of Scientific Studies*. British School at Athens Fitch Laboratory Occasional Paper 1, Athens.

Kingery, W.D., Bowen, H.K. and Uhlmann, D.R. (1976). *Introduction to Ceramics*. John Wiley, New York (2nd edn.).

Maggetti, M. (1982). Phase analysis and its significance for technology and origin. In *Archaeological Ceramics*, eds. Olin, J.S. and Franklin, J.D., Smithsonian Institution Press, Washington DC, pp. 121–133.

Millot, G. (1970). *Geology of Clays: Weathering, Sedimentology, Geochemistry.* Chapman and Hall, London.

Murad, E. and Wagner, U. (1991). Mössbauer spectra of kaolinite, halloysite and the firing products of kaolinite – new results and a re-appraisal of published work. *Neues Jahrbuch für Mineralogie-Abhandlungen* **162** 281–309.

Neff, H. (ed.) (1992). *Chemical Characterization of Ceramic Pastes in Archaeology.* Prehistory Press, Madison, Wisconsin.

Norton, F.H. (1952). *Elements of Ceramics.* Addison-Wesley, Cambridge, Massachussets.

Orton, C., Tyers, P. and Vince, A. (1993). *Pottery in Archaeology.* Cambridge University Press, Cambridge.

Picon, M., Carre, C., Cordoliani, M.L., Vichy, M., Hernandez, J.A. and Mignard, J.L. (1975). Composition of the La Graufesenque, Banassac and Montans terra sigillata. *Archaeometry* **17** 191–199.

Picon, M., Vichy, M. and Meille, E. (1971). Composition of the Lezoux, Lyon and Arezzo samian ware. *Archaeometry* **13** 191–208.

Pollard, A.M. (1982). A critical study of multivariate methods as applied to provenance data. In *Proceedings of the 22nd Symposium on Archaeometry, University of Bradford, 30th March–3rd April 1982*, eds. Aspinall, A. and Warren, S.E., University of Bradford Press, Bradford, pp. 56–66.

Pollard, A.M. (1986). Multivariate methods of data analysis. In *Greek and Cypriot Pottery: a Review of Scientific Studies*, ed. Jones, R.E., British School at Athens Fitch Laboratory Occasional Paper 1, Athens, pp. 56–83.

Pollard, A.M. (1993). Groundwater modelling in archaeology – the need and the potential. Paper presented at *Archaeological Science 93*, University of Bournemouth, September 1993 (Proceedings in press).

Pollard, A.M., Hatcher, H. and Symonds, R.P. (1981). Provenance studies of 'Rhenish' pottery by comparison with *terra sigillata*, *Revue d'Archéométrie, Actes due XX Symposium International d'Archéométrie, Paris 26–29 Mars 1980,* Vol II, 175–185.

Pollard, A.M., Hatcher, H. and Symonds, R.P. (1982). Provenance studies of 'Rhenish' wares – a concluding report. In *Proceedings of the 22nd Symposium on Archaeometry, University of Bradford, 30th March-3rd April 1982*, eds. Aspinall, A. and Warren, S.E., University of Bradford Press, Bradford, pp. 343–354.

Pollard, A.M. and Wood, N. (1984). Development of Chinese porcelain technology at Jingdezhen. In *Proceedings of the 24th International Archaeometry Symposium*, eds. Olin, J.S. and Blackman, M.J., Smithsonian Institution, Washington, pp. 105–114.

Putnis, A. (1992). *Introduction to Mineral Sciences.* Cambridge University Press, Cambridge.

Renfrew, C. (1977). Introduction: Production and exchange in early state societies, the evidence of pottery. In *Pottery and Early Commerce: Characterization and trade in Roman and later ceramics*, ed. Peacock, D.P.S., Academic Press, London, pp. 1–20.

Rice, P.M. (1987). *Pottery Analysis: A Sourcebook.* University of Chicago, Chicago.

Ries, H. (1927). *Clays. Their Occurrence, Properties and Uses.* Chapman and Hall, London.

Singer, F. and Singer, S.S. (1963). *Industrial Ceramics.* Chapman and Hall, London.

Symonds, R.P. (1992). *Rhenish Wares. Fine Dark Coloured Pottery from Gaul and Germany.* Oxford University Committee for Archaeology Monograph No. 23, Oxford.

Vandiver, P.B., Soffer, O., Klima, B. and Svoboda, J. (1989). The origins of ceramic technology at Dolni Vestoniče, Czechoslovakia. *Science* **246** 1002–1008.

von Meyer, E. (1891). *A History of Chemistry.* Macmillan, London.

Chapter 5

The Chemistry and Corrosion of Archaeological Glass

INTRODUCTION

In the modern world, we are accustomed to taking the chemical stability of glass very much for granted – we rely on the durability of glass for so many things, such as windows and (until the widespread availability of plastics) bottles, as well as its use in the chemical laboratory as an extremely inert container. In addition to its apparent inertness, glass has a number of other beneficial properties, such as its transparency, or the ability to take on virtually any colour as the result of the addition of small amounts of transition metals.

Even with modern glass, this apparent chemical stability is largely an illusion: for example, Bacon (1968) observed that an ordinary soda-lime-silica glass bottle full of water will lose approximately 30 mg of material per year into the solution. The situation with archaeological glass is considerably more complex. During the Roman period and through the early Medieval period (to about 1100 AD) in Europe, glass was made from either soda-rich plant ash and pure sand, or *natron* (mineral hydrated sodium carbonate – $Na_2CO_3.10H_2O$) mixed with a calcareous sand (Newton, 1982), both resulting in a glass with a remarkably modern composition. Like modern glass, it is relatively durable, although the majority of excavated examples show evidence of surface iridescence due to deterioration (see later section). During the European Medieval period (*ca.* 1100–1600 AD), however, the 'building boom' in great cathedrals required the supply of large quantities of coloured glass to provide the so-called 'stained glass windows' (technically, the majority from the earlier period are painted glass, not stained glass – see below), which was satisfied by glass manufactured using beech ash as the source of the alkali. This '*forest glass*' has proved to be a major problem for modern-day

conservators, since its composition has rendered it almost water soluble in extreme cases, and the maintenance of the surviving glass in a large cathedral has been likened to the mythological painting of the Forth Rail Bridge in Scotland – as soon as the job is completed, it is time to start again! The cause of this poor durability in Medieval glass is to be found largely in the composition of the glass itself, although storage in damp conditions during the Second World War as protection against bomb damage has exacerbated the problem in some cases. Contrary to popular opinion, the increased atmospheric pollution during the 20th Century, whilst clearly not beneficial, has probably not been the primary cause of this deterioration.

This chapter reviews the chemistry and structure of silicate glasses, and looks at the chemistry of the colouring process. It then summarizes the work on the relationship between the composition and durability of Medieval European window glass. The final section looks briefly at the much more complex decay mechanisms of buried glass.

THE STRUCTURE AND CHEMISTRY OF GLASS

Before describing the structure of glass, it is important to first define the term 'glass', and then to specify the particular branch of this large class of materials with which we are concerned. The definition given by the American Society for Testing and Materials is: '*Glass is an inorganic product of fusion which has cooled to a rigid condition without crystallizing*'. This by no means covers all substances which are found in the vitreous state (*e.g.*, toffee made from heated sugar), nor does it describe all the current possible methods of manufacturing glasses, although it is sufficient for archaeological purposes. An alternative structural definition quoted by Holland (1964) is '*the glass-like or vitreous state is believed to be that of a solid with the molecular disorder of a liquid frozen into its structure*'. A more formal statement of this is given by Paul (1990; 4): '*glass is a state of matter which maintains the energy, volume, and atomic arrangement of a liquid, but for which the changes in energy and volume with temperature and pressure are similar in magnitude to those of a crystalline solid*'. This then may also include organic polymers and resins as well as non-oxide glasses, but it is generally better to consider the vitreous state as an additional state of matter (to be added to solid, liquid, and gas) rather than to define a glass by the process used in its manufacture. In archaeological studies we are exclusively concerned with oxide glasses based on silica, and with few exceptions belonging to the *alkali–alkaline earth–silica* family of glasses (much of what is said here also applies to ceramic glazes, although this also includes systems such as lead silicates not described here). The production of such

Table 5.1 *Characteristic reference temperatures (°C) in glass working* (Adapted with permission from Holloway, 1973; 19)

Reference point	*Definition*	*Viscosity (poise)*	*96% SiO_2 (°C)*	*Soda-lime silica (°C)*
Working point	Sufficiently soft for shaping	10^4	–	1000
Softening point	Glass tubes can be bent in a flame	$10^{7.6}$	1500	700
Annealing point	Internal stresses removed in a few minutes	$10^{13.4}$	900	510
Strain point	Highest temperature from which the glass can be rapidly cooled without serious internal stress	$10^{14.5}$	820	470

glasses is still the most important aspect of the glass industry, although many other oxide glasses of technical importance have been developed during the last century. New manufacturing techniques now allow a much wider range of materials to be made in the vitreous state, including products such as metallic glasses, which are now of great technological importance (*e.g.*, Greer, 1995; see below).

An important distinguishing feature of a glass is that it shows no discontinuous change of any measurable property on cooling from the liquid to the vitreous state. Hence it is necessary to define various reference points, usually in terms of *viscosity*, in order to discuss the glass-forming process. Table 5.1 gives the commonly accepted reference points defined by the viscosity of the melt, a brief definition of each point, and the corresponding temperatures for a typical soda-lime-silica glass and 96% vitreous silica. The viscosities are quoted in the CGS units of poise, with dimensions (g cm^{-1} s^{-1}), as is conventional in glass technology (the SI equivalent is kg m^{-1} s^{-1}). For comparison, glacier ice has a viscosity of 1.2 x 10^{14} poise, glycerol $10^{1.2}$ poise, and water at 25 °C approximately 10^{-2} poise. This Table incidentally also shows clearly the benefit in terms of workability which is achieved by modifying the composition of pure silica glass with the addition of alkali and alkaline earths – roughly a factor of two reduction in the reference temperatures.

The specific volume (volume per unit mass) of a glass melt is shown in Figure 5.1 as a function of temperature. The solid line shows the result of allowing the melt to cool slowly and crystallize, with a characteristic

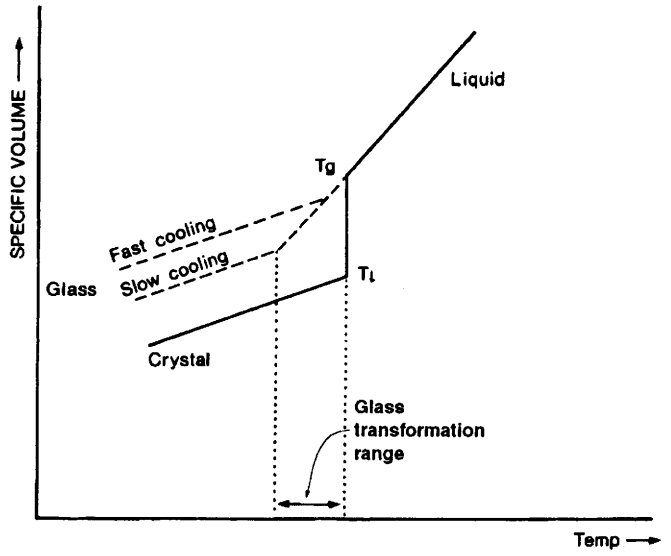


Figure 5.1 *Transformation from a liquid to a crystalline or glassy state* (Adapted from Paul, 1990; Figure 1.2, by permission of Chapman and Hall)

sharp drop in specific volume at the melting point, corresponding to the phase transition from a liquid to a crystalline solid. This temperature depends on the composition of the melt and is known as the *liquidus temperature* (T_l), defined as the temperature above which no crystalline solid exists. In the glassy state (the broken line) achieved by a slower cooling rate, the curve bends at some temperature below T_l which depends on the rate of cooling, and attains a gradient similar to that of the crystalline state. The bend defines the *glass formation temperature* (T_g) for that particular melt and cooling rate. Below T_g the melt has become too viscous for molecular movement, and so the system is metastable with respect to the crystalline state, and has the structure of the liquid at T_g 'frozen' into the solid. Glasses of the same composition exhibit a range of temperatures over which vitrification occurs (called the *glass transformation range*) and the exact position of T_g depends on the rate of cooling of the melt. Properties of the glass which depend on the structure (*e.g.*, ionic conductivity, electrical resistance, chemical resistance) depend on T_g and hence on the thermal history of the glass. This temperature is also known as the *fictive temperature.*

The first theory of the structure of glass to become widely accepted was that of Zachariasen (1932), called the *random network theory* [now

commonly referred to as the *continuous random network* (CRN) theory]. This arose from the following comparison of the behaviour of crystalline and vitreous silica:

(i) the mechanical properties of the glass are equivalent to the crystalline form over a large range of temperatures. Thus the bonding forces must be similar;
(ii) *X*-ray diffraction studies of the time showed no periodic structure in the vitreous state;
(iii) the vitreous structure cannot be entirely random because the intermolecular distances do not fall below certain minimum values.

From these observations Zachariasen concluded that glass can be thought of as an infinitely large unit cell containing an infinite number of atoms. The difference between the crystalline and the vitreous state is, therefore, found in the presence or absence of periodicity – what we would now describe as '*long range order*', which is a fundamental property of a crystalline structure. This theory explains many of the differences between glasses and crystals, such as:

(i) the optical isotropy of glasses results from the random atomic arrangement;
(ii) in a random network no two atoms occupy exactly identical sites in terms of their chemical environment, and so the lack of abrupt changes of state can be understood. The bond energies all vary slightly (contrasting with the fixed bond energies in an ideal crystal) and hence the thermal breakdown of the network occurs over a large range of temperatures in a glass – it gradually softens into a liquid rather than melts at a characteristic temperature, as is the case with a crystalline solid. (A glass is conventionally said to have 'melted' when its viscosity is reduced to between 10^8 and 10^2 poise);
(iii) the composition of glass is not stoichiometric – it is a mixture, not a compound;
(iv) there are no crystal cleavage planes in glass – hence the characteristic conchoidal fracture patterns.

It also allowed him to predict on the basis of molecular geometry those oxides which ought to be capable of forming glasses. Goldschmidt had previously observed that the radius ratio of the metals and oxygen (R_A/R_O) in glass forming oxides of the formula A_mO_n was between 0.2 and 0.4, and that *tetrahedral co-ordination* was also necessary (*i.e.*, four oxygen atoms surround the metal atom in a tetrahedral shape – see

Figure 5.2). Zachariasen generalized this theory by giving the following rules for the formation of an oxide glass:

(i) Oxygen atoms are linked to not more than two other atoms.
(ii) The number of oxygen atoms surrounding the cations must be small.
(iii) The oxygen polyhedra share corners with each other, not edges or faces. To give a three dimensional structure at least three corners in each polyhedra must be shared.

On the basis of these rules, he predicted that B_2O_3, SiO_2, GeO_2, P_2O_5, As_2O_5, P_2O_3, As_2O_3, Sb_2O_3, V_2O_5, Sb_2O_5, Nb_2O_5, and Ta_2O_5 may all occur as a single substance in the vitreous state. Of these, B_2O_3, SiO_2, GeO_2, P_2O_5, As_2O_5, and As_2O_3 are known to do so, thus giving good support to Zachariasen's model.

Experimental support for this theory came in a series of papers from Warren and co-workers (Warren, 1937; Warren and Biscoe, 1938). They obtained experimental *X*-ray scattering curves from vitreous silica and applied the Fourier analysis technique developed for the study of liquid structure (Warren, 1937) to obtain a radial distribution curve. This is a function which gives the number of atoms about any atom and their average distance from it. From peaks in the radial distribution curves,

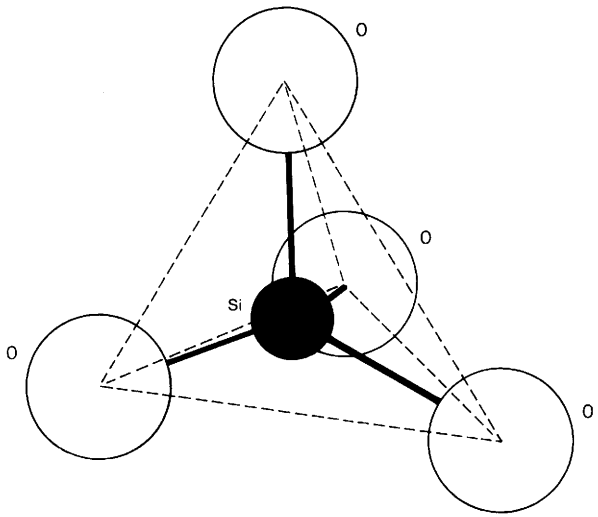


Figure 5.2 *Tetrahedral co-ordination of four oxygens around a silicon atom* (After Volf, 1961; 20 by permission of Pitman Publishing)

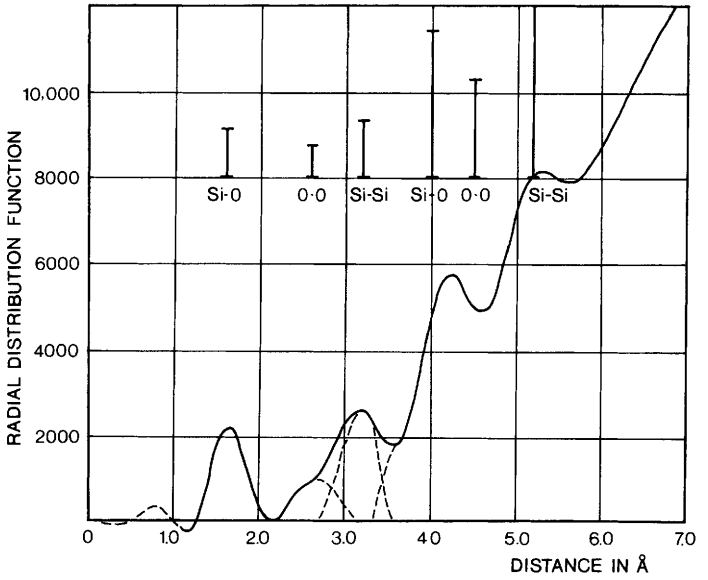


Figure 5.3 *Radial distribution curve for vitreous SiO_2 from* X*-ray studies. Peaks correspond to average atomic separations of Si–O, O–O, etc.* (Warren, 1937; Figure 2, by permission of American Institute of Physics)

the average atomic separations, up to a distance of about 7–8 Å (10 Å = 1 nm = 10^{-9} m), can be measured, and from the area beneath each peak an estimate of the co-ordination numbers (number of surrounding atoms) can be made. For vitreous silica (Figure 5.3) Warren found four oxygens tetrahedrally distributed around the silicon, with an average Si–O bond distance of 1.62 Å (compared with 1.60 Å in crystalline silica). The oxygen–oxygen separation in the tetrahedron was found to be 2.65 Å, and further regular interatomic distances were identified out to about 4.5 Å. From this, he concluded that Zachariasen's model was a valid interpretation of the structure of vitreous silica. Warren and Biscoe (1938) studied binary soda–silica glasses in a similar manner and found an extra peak due to the average Na–O separation at 2.35 Å. They concluded that the average structure of this binary SiO_2–Na_2O glass is a silicon atom tetrahedrally surrounded by four oxygen atoms at a distance of 1.62 Å. The sodium ions are situated interstitially in the network, surrounded on average by six oxygens at 2.35 Å. Arguments based on co-ordination numbers rule out the existence of discrete molecules with the formula Na_2O, $Na_2Si_2O_5$, Na_2SiO_3, *etc.*, but the structure deter-

mined by this method is an average, and the presence of small regions showing crystalline form cannot be ruled out on this evidence (see below). Warren calculated the maximum size allowable for crystalline structures in the glass on the basis of the measured diffraction patterns, and obtained a figure of 7 Å. This is approximately the size of the crystalline silica unit cell, so the term 'crystal' cannot meaningfully be applied to these minuscule aggregations. Subsequent studies using more sophisticated techniques such as nuclear magnetic resonance (NMR- see Chapter 2) to study the distributions of the oxygens, and extended *X*-ray absorption fine structure (EXAFS) to look at the oxygen distributions around the sodium ions in binary soda-silica glasses (Greaves *et al.*, 1981) have emphasized the importance of this short-range order in glass, but have substantially supported the views formed in the 1930s.

Figure 5.4 shows this classical representation of glass structure in two dimensions, following the work of Zachariasen and that of Warren and Biscoe. The compound R_2O_3 is shown in its crystalline and vitreous forms (R in this model is a trivalent cation), and also the vitreous structure when modified by the addition of a substance of the form M_2O. Zachariasen chose to illustrate the structure of glass using a trivalent oxide (such as boron oxide, B_2O_3) simply for ease of representation in two dimensions. In order to picture the structure of a silicate glass, Figure 5.4 should be considered as a 'slice' through the full three dimensional network, remembering, however, that the tetrahedral coordination of the oxygen around the silicon means that in reality the silicon and oxygen atoms are not in the same plane, and therefore the 'slice' should not be regarded as flat – more as a crinkly sheet, with the oxygens slightly lower than the silicon.

As a result of the random network theory, glass-forming oxides have been classified as *network formers, network modifiers*, and *intermediates*. Network formers are those which may be found in the vitreous state as pure substances, and include SiO_2, B_2O_3, GeO_2, P_2O_5, and As_2O_5. The network modifiers disrupt the continuity of the network, changing the physical and chemical properties, as illustrated in Figure 5.4c. This category includes the alkali metal oxides (*e.g.*, Na_2O, K_2O) and the alkaline earth oxides (*e.g.*, CaO, MgO). Intermediates are oxides which can either enter the network as a network former or occupy interstitial holes (*i.e.*, act as a network modifier), but are unable to form glasses themselves. Included in this class are Al_2O_3, TiO_2, and ZrO_2. In archaeological glasses, the principal network former is silica (SiO_2), with the alkalis (soda, Na_2O and potash or potassia, K_2O) and alkaline earths (calcia or lime, CaO and magnesia, MgO) as the network modifiers. Alumina (Al_2O_3) plays an important role in stabilizing the network and

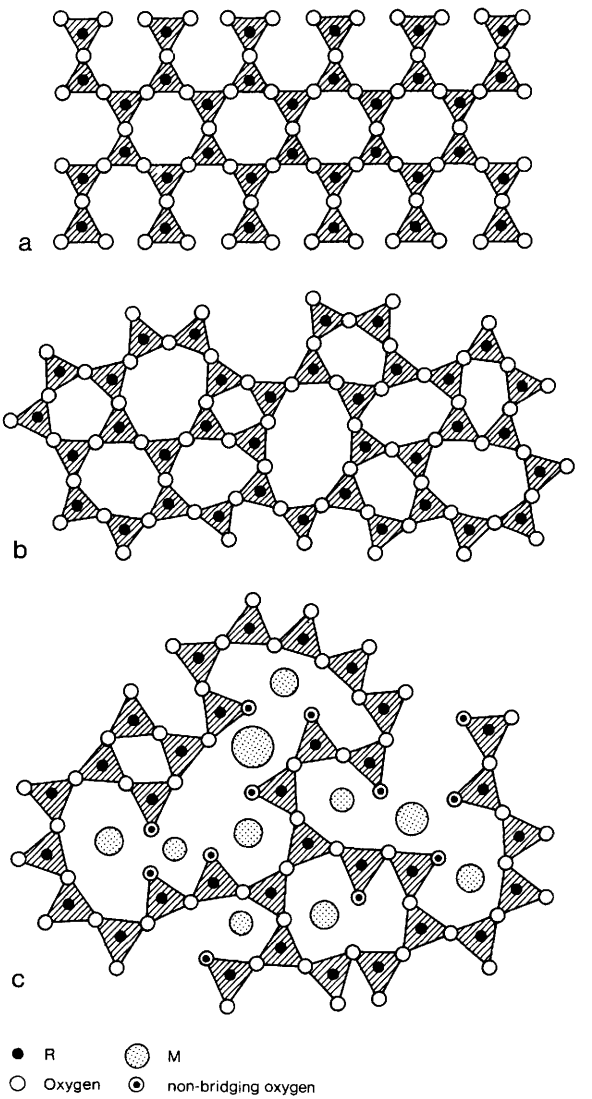


Figure 5.4 *Schematic representation of the vitreous state of a compound of stoichiometry R_2O_3, and as modified by a compound M_2O, according to the random network theory:* (a) *R_2O_3 in the crystalline state;* (b) *R_2O_3 in the vitreous state;* (c) *vitreous R_2O_3 modified by M_2O, showing bridging and non-bridging oxygen sites*

(Reprinted with permission from Zachariasen, 1932; Figures 1a and 1b (copyright 1932, American Chemical Society), and Warren and Biscoe, 1938; Figure 4, by permission of the American Ceramic Society)

is the most important of the glassmaking intermediates. As can be seen from Table 5.1, although silica in the vitreous state is a perfectly well-behaved glass, its value is limited as a usable substance by its very high working temperatures – not softening until 1500 °C, for example. These temperatures are far too high to have been achieved routinely by the ancient glassworkers, and it was only by the discovery of the fluxing effect of the alkalis on these thermal properties that glassmaking became possible. The traditional story of the discovery of the beneficial properties of alkali on glassmaking as first told by Pliny in the 1st Century AD is recounted by Vose (1980). Phoenician merchants camping on the banks of the River Belus used the cakes of natron (natural sodium carbonate) to support their cooking pots on the fire – in the morning they found the natron had fused the sand to form a glass. As all subsequent commentators have pointed out, the story is unlikely to be true, since the temperature of a cooking fire is unlikely to have reached the 1000 °C or more required to produce glass, but it clearly indicates a knowledge of the use of alkalis to modify the properties of silicate glasses, and that the River Belus was an important source of glass-making sand in antiquity.

From Figure 5.4c, it is apparent that some of the oxygen ions in the modified structure are bonded to two silicon atoms, and some to only one, with the resulting excess negative charge being balanced by an interstitial ion of the type M^+. Oxygen ions which join two polyhedra are known as *bridging oxygens*, and those linked to only one *non-bridging*. Stevels (1960–61) characterized the structure of the network with four parameters X, Y, Z, and R, defined as:

X = average number of non-bridging oxygen ions per polyhedron,

Y = average number of bridging oxygen ions per polyhedron,

Z = average total number of oxygen ions per polyhedron,

R = ratio of total number of oxygen ions to total number of network formers.

Z is known from the identity of the principal network former; for silicates, Z is 4. R can be deduced from the molar composition of the glass, although in 'real' glasses R is not predictable precisely because of the presence of intermediate oxides. Knowing Z and R allows the values of X and Y to be calculated for a particular glass, using the simple relationships $X + Y = Z$ and $X + Y/2 = R$. Stevels demonstrated that many properties of the glass depend on the value of Y, showing that glasses of different chemical composition, but with the same Y value, have the same physical properties, such as viscosity and thermal expansion coefficient. From studies of the variations in properties with changing composition, he was able to show that a marked change in behaviour occurs when the value of Y falls below 3. This figure marks the

point at which polyhedra exist which are connected by only two points of contact with neighbouring polyhedra. Thus the rigidity of the network decreases significantly and properties which depend on ionic transport (such as electrical conductivity, *etc.*) show a significant change. In binary soda-silicate glasses, for example, this point corresponds to the composition $Na_2O.2SiO_2$ (33.3 mole % Na_2O, 66.67 mole % SiO_2) and the gradient of a plot of electrical conductivity against composition shows a marked change at this value (Stevels, 1960–61; Figure 3). [Mole % is the measure of the abundance of a particular component by number of (theoretical) molecules, rather than by weight].

A similar result can be obtained by studying the molar ratio of silicon to oxygen atoms in a silicate glass. In vitreous SiO_2 the ratio is 0.5, simply calculated from the stoichiometric ratio of 1 silicon to 2 oxygens, and this figure falls as the number of non-bridging oxygen atoms increases (*i.e.*, as a result of adding extra oxygen via the network modifying oxides). Most chemical analyses of glasses are reported as weight percent oxide, but it can be seen that many of the important properties depend, not on this, but on the molar percentage (*i.e.*, the relative number) of atoms present. The molar ratio Si/O can be calculated for any glass from its chemical (weight % oxide) analysis using the formula due to Huggins and Sun (1943; eqn. 16):

$$Si/O = \frac{f_{Si}}{60.06 \sum \frac{n_M \times f_M}{W_M}}$$

where f_{Si} = wt fraction of SiO_2 (= wt % SiO_2/100),
60.06 = molecular weight of SiO_2,
n_M = no. of oxygens in formula for oxide (M_MO_N),
W_M = molecular weight of oxide M_MO_N,
and f_M = wt fraction of M_MO_N (= wt % M_MO_N/100),

and the summation is over all the oxides in the glass, including SiO_2. Table 5.2 shows the relationships between the calculated value of Si/O, the number of bridging and non-bridging oxygen ions, and the probable bonding of the silicon in the glass, where the marked bond signifies a bridging (Si–O–Si) bond. As the value of the Si/O ratio increases, the number of bridging oxygens increase and the degree of 'polymerization' of the network increases. These figures have important consequences for the interpretation of the chemical durability of glass, as discussed below.

An alternative theory of glass structure, attributed to Lebedev in the Soviet Union in 1921 (Vogel, 1977; Porai-Koshits, 1977), is the *crystallite theory*. This model maintained that glass is made up of small crystalline regions varying in size from 10 Å (three to four silica units) up

Table 5.2 *Relationship between the Si/O ratio in a glass, the number of non-bridging oxygens per polyhedra, and the structure* (Adapted from Volf, 1961; 20 by permission of Pitman Publishing)

Si/O Molar ratio	*Number of bridging oxygens per tetrahedron*	*Number of non-bridging oxygens per tetrahedron*	*Pattern of structure*
<0.286	1	3	— Si
0.286–0.333	1–2	2–3	— Si to — Si
0.334–0.400	2–3	1–2	— Si to — Si —
0.401–0.444	3–4	0–1	— Si— to —Si—
0.500	4	0	— Si —

to perhaps 300 Å. A critical review of these two competing views on the nature of glass has been published by Porai-Koshits (1990). As mentioned above, the experimental work of Warren specifically ruled out the existence of crystalline regions larger than about 10 Å, but it is appreciated that the random network theory presents only an average picture of the structure, and says nothing about the local distribution of ions. Porai-Koshits states that regions of maximum order in the continuous network may be conveniently termed 'crystallites', and, using this definition of a crystallite, the two theories are seen to blend into one another. In 1971, Soviet academicians concluded that modern methods of analysis made it impossible to hold on to the original concept of crystallites in glass (Porai-Koshits, 1985), and yet direct evidence for structural inhomogeneity on the sub-micron to micron scale has subsequently been obtained using electron microscopy and small angle *X*-ray scattering (SAXS). Application of this technique to glasses, together with direct observation by electron microscopy, has conclusively demonstrated the phenomenon of phase separation in many vitreous systems (Vogel, 1977). The composition of the various phases in multicomponent glasses has been confirmed using electron microprobe analysis. The current understanding of the structure of glass may be summarized by combining the comments of Vogel (1977) and Porai-Koshits (1977):

(i) In one component glasses (such as pure vitreous silica), no inhomogeneous structure occurs, except for frozen thermal density fluctuations corresponding to the situation in the melt at the glass transformation temperature. These glasses are adequately described by the random network theory of Zachariasen and Warren.
(ii) All multicomponent glasses whose composition is intermediate between two stable compounds exhibit phase separation, the compositions of each phase approximating to stable compositions. Maximum phase separation occurs in compositions which are furthest removed from the stable compositions.
(iii) The addition of further components such as colourants does not lead to a homogeneous distribution of additives. Highly charged cations will concentrate in anion-rich phases.

There is now ample evidence to support the view that glasses cannot simply be regarded as a homogeneous continuous network, as predicted by the random network theory, but nor is it a series of domains of highly ordered crystallites, as originally proposed by Lebedev. For example, recent detailed studies of sodium silicate glasses using molecular dynamic simulations to model the properties and structure of glasses have shown that the sodium ions are not randomly distributed through the structure, but associate with non-bridging oxygens to produce silica-rich and alkali-rich regions on the atomic scale within the glass, thus supporting the above consensus view (Huang and Cormack, 1990).

The phase structure of glasses has a significant effect on their physical properties, which is discussed below with reference to chemical durability. The magnitude of the phase separation can be altered by heat treatment, and enhanced or reduced by the addition of various oxides to the melt. In particular, the addition of alumina to commercial soda–lime–silica glasses reduces the tendency to phase separation, improving chemical resistance (Doremus, 1973). A detailed study of the microstructure of soda–lime–silica glasses has been published by Burnett and Douglas (1970). The control of phase separation in the melt is now commercially important for processes such as the production of 96% silica glass for laboratory use from phase-separated sodium borosilicate glasses by the Vycor process (Paul, 1990). In this process, a glass composed of approximately 8% Na_2O, 20% B_2O_3, and 72% SiO_2 is heat treated at between 500 and 800 °C, inducing the separation of two phases, one silica rich and the other rich in boric acid. When treated with dilute hydrochloric acid, the boron-rich phase dissolves, leaving a highly porous but almost pure silica glass which, when heated to

1200 °C, becomes more compact and dense. The advantage is that the vitreous silica has been produced at a temperature much below that which would have been necessary to melt a glass of pure silica. Phase separation phenomena are also vitally important in understanding that other branch of silicate chemistry important in archaeology – that of glaze manufacture. One spectacular example is the production of Jun ware glazes by Chinese potters during the Song dynasty (960–1263 AD) and later, in which the characteristic sky blue colour is produced by an emulsion of two liquids forming in the glaze at high temperature, and subsequently being retained by being cooled at exactly the correct rate (Kingery and Vandiver, 1986) – a phenomenal achievement for potters working 1000 years ago!

The vitreous state should therefore be seen as an addition to the traditional three states of matter – one that is described by an essentially random network characteristic of a liquid, but which may still retain inhomogeneities. The crucial factor in deciding whether a melt will crystallize or become vitreous is kinetically controlled – the rate of cooling relative to the bond energies involved. The stronger the chemical bonds, the more sluggish will be the rearrangement processes close to the melting point, and the more likely the melt will be to undergo the glass transition process rather than crystallization, which requires a great deal of molecular reorganization. Silica, with a high Si–O bond energy (368 kJ mol^{-1}), is particularly prone to vitrify, but almost any melt can be vitrified if the rate of cooling is fast enough. This is the basis of the '*splat cooling*' process which can produce metallic glasses, in which a jet of molten metal is sprayed onto a rotating cooled drum, achieving a temperature drop of 1000 °C in a millisecond, causing the metal to vitrify (Greer, 1995).

It should be noted that the discussion in this section has used the terms 'atom', 'ion', and 'molecule' virtually indiscriminately, reflecting the usage of these terms in the glass chemistry literature. The Zachariasen model is phrased in terms of atoms and assumes total covalency of bonding, which gives the required directionality of the bonds, but does not admit much flexibility in the bond lengths, as is required in the model. Ionic or metallic bonds do not have the required directionality. Subsequent workers have modified the theory to account for a 'mixed' (covalent – ionic) bond which is a more realistic picture of the bonding in real glasses, but these modifications have little impact on the structural models as outlined, and are generally ignored in the literature (Paul, 1990). One final point to be emphasized is that although glass analyses are conventionally reported as 'weight percent oxides' or 'molar percent oxides', the presence of compounds such as Na_2O or CaO in a glass is

purely fictitious, in that compounds of this stoichiometry do not physically occur in the glass. Remembering that this oxide notation is purely for convenience is particularly important when dealing with transition metal oxides such as those of Fe^{II} or Fe^{III}, since the assumption of a valence state for a particular transition metal is also largely conventional (see below). The assumed valency of the metal could, however, affect the total analysis of the glass when the iron content is reported as one or other 'fictitious oxide' such as FeO or Fe_2O_3. This is discussed in more detail in the next section.

THE COLOUR OF GLASS

The colour of a glass is usually attributable either to the presence of small amounts of transition metal ions or to metallic atoms within the structure, and the colours resulting from various additives have been studied for well over 100 years (Weyl, 1976). Exceptions to this rule are *opaque glasses*, in which the opacity is often due to the presence of a second finely distributed immiscible phase within the glass, giving rise to the scattering of light which results in a milky or opaque appearance. Another case is the presence of thin layers on the surface of the glass, which can give rise to an iridescent appearance as the result of diffraction phenomena in the surface layers. The latter often arises as a result of weathering (or corrosion in the case of buried glasses), and is a characteristic feature of many otherwise well-preserved archaeological glasses. The former is usually more important in the production of ceramic glazes or enamelled metalwork, apart from certain classes of artefact such as glass beads, which are deliberately made to be opaque. This section concentrates on the chemical explanation of the colours found in transparent glasses.

The classic treatise on the subject of the colour of silicate glasses is that of Weyl (1976). He describes the colours produced by metallic ions in glasses (including the various oxidation states of Fe, Mn, Cr, V, Cu, Co, Ni, U, Ti, W, Mo, and the rare earths), and also those produced by metal atoms in glass, principally Au, Ag, and Cu, as well as describing the effect of replacing the oxygen in the glass with other anions. The results of his synthesis are well-known in archaeology, although the detailed chemistry of the causes of the different colours is less well-appreciated. Table 5.3 lists the colours produced by the traditional colouring ions, together with the modern raw materials used to impart such colours. As can be seen, the colour produced by a metal ion depends not only on its oxidation state, but also on the position it occupies in the glass structure. The reasons for this are explained below.

Table 5.3 *Some of the chromophores responsible for the colours in ionically coloured glasses* (Adapted with permission from Weyl (1976; 60) and Vogel (1994; 232)

Transition metal	*Modern raw material*	*Colouring ion*	*Colour in tetrahedral coordination (network former)*	*Colour in octahedral coordination (network modifier)*
Chromium	chromium(III) oxide Cr_2O_3	Cr^{III}	–	green
	potassium dichromate(VI) $K_2Cr_2O_7$	Cr^{VI}	yellow	–
Copper	copper(II) oxide CuO copper(II) sulfate $CuSO_4.5H_2O$	Cu^{I}	–	colourless to red, brown fluorescence
		Cu^{II}	yellowish-brown	blue
Cobalt	cobalt(III) oxide Co_2O_3 cobalt(II) carbonate $CoCO_3$	Co^{II}	blue	pink
Nickel	nickel(II) oxide NiO nickel(III) oxide Ni_2O_3 nickel(II) carbonate $NiCO_3$	Ni^{II}	purple	yellow
Manganese	manganese(IV) oxide MnO_2 potassium manganate(VII) $KMnO_4$	Mn^{II}	colourless to faint yellow, green fluorescence	weak orange, red fluorescence
		Mn^{III}	purple	–
Iron	iron(II) oxalate $FeC_2O_4.2H_2O$	Fe^{II}	–	absorbs in infrared
	iron(III) oxide Fe_2O_3	Fe^{III}	deep brown	weak yellow to pink
Uranium	uranium oxide U_3O_8 sodium uranate $Na_2U_2O_7.3H_2O$	U^{VI}	yellowish-orange	weak yellow, green fluorescence
Vanadium	vanadium(V) oxide V_2O_5	V^{III}	–	green
		V^{IV}	–	blue
		V^{V}	colourless to yellow	–

Even this table of colours is by no means exhaustive – for example, Fe can give a range of colours from pale yellow through to blues and greens, depending on the oxidation state of the ion, the presence of other transition metal ions, and the composition of the base glass. Indeed, an understanding of the principles underlying the colour of glasses is far more complex than one might believe, and requires a discussion of factors such as the redox conditions in the furnace and the nature of the co-ordination sphere around the colouring ion – all concepts familiar to inorganic chemists.

Before attempting to explain the phenomenon of colour in glasses, it is instructive to look at aqueous solutions coloured by transition metal ions. A well-known undergraduate chemistry experiment is to observe the difference in colour between an aqueous solution of copper sulfate and a solution of copper sulfate crystals dissolved in a mixture of water and concentrated ammonia solution. The aqueous solution is the familiar pale blue-green of copper solutions, but as ammonia is added the colour deepens to a much more intense blue. The explanation of this apparently simple change requires a detailed investigation of *co-ordination chemistry*. Even the original blue crystals of copper sulfate are not what they seem – in fact they are the pentaquo compound $CuSO_4.5H_2O$ – in other words, the central copper ion is surrounded by six groups (called *ligands*) – one sulfate, and five water molecules. These six ligands are arranged in a distorted octahedron – four of the water groups in a square plane surrounding the copper, plus one water below, and the sulfate above the plane (Figure 5.5), and the co-ordination is termed *octahedral*. The term 'distorted' implies that the Cu–O bond lengths to the square planar water groups are equal, but are not the same as the Cu–O bond to the

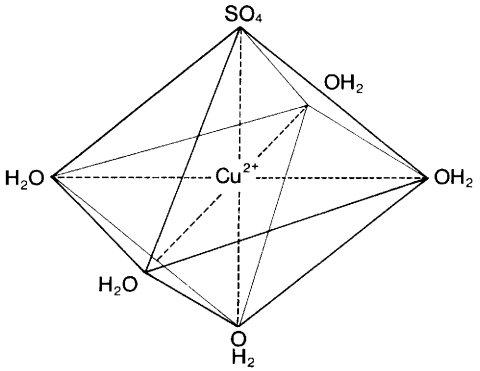


Figure 5.5 *Octahedral co-ordination of the Cu^{2+} ion in copper sulfate, $CuSO_4.5H_2O$*

other water group, or the Cu–O bond length to the oxygen in the sulfate group. Anhydrous copper sulfate – without the five waters – would be colourless, indicating that the colour of the crystal is strongly related to presence of this *co-ordination sphere* around the copper ion, which is in fact the case.

The colour of the transition metal ions, which is one of their major characteristics, arises directly from the interaction between the outer orbital electrons of the transition metal and the electric field created by the presence of the co-ordinating ligands. The theory of this is called *ligand field theory*, and is well-covered in most basic textbooks on inorganic chemistry (*e.g.*, Cotton and Wilkinson, 1976). Ligand field theory is an extremely powerful tool, which explains all manner of spectroscopic, electronic, and magnetic properties of the transition elements. A simplified version, based purely on electrostatic considerations, is called *crystal field theory*, which allows a relatively simple explanation to be given for the causes of colour in transition metal compounds. Transition metals are defined as those elements which have partially filled *d*-orbitals (or *f*, but we will restrict this discussion to the *d*-block elements: see the Periodic Table, Appendix 5). As explained in Appendix 1, the *d*-orbitals are made up of five possible energy levels, and because each energy level can accommodate two electrons (one 'spin up' and one 'spin down'), they can therefore hold a maximum of ten electrons before they are complete. In free space (*i.e.*, far removed from any electric or magnetic field), the five energy levels are exactly equivalent in energy, and therefore a transition metal ion with, say, three *d*-electrons (*e.g.*, Cr^{III}, or Mn^{IV}) could put one electron in any three energy levels, leaving the other two empty. (Putting two electrons in the same orbital whilst another orbital of the same energy level is empty is energetically unfavourable). However, when the transition metal ion is co-ordinated by other ions or molecules, as it invariably is in a crystal or in solution, then there is an interaction between the charges on the neighbouring ions, atoms, or molecules and the *d*-orbitals of the transition metal ion. Because the five *d*-orbitals have different shapes in space around the nucleus, some are affected more than others by the surrounding charges, and the precise nature of the interaction depends on the relationship between the fixed geometry of the *d*-orbitals and the distribution of nearest neighbours around the ion. In octahedral co-ordination (as shown in Figure 5.5), for example, three of the *d*-orbitals are depressed in energy compared to the *d*-orbital energy in the free ion, and two are increased. In tetrahedral co-ordination (*i.e.*, with four co-ordinating ligands making a tetrahedral shape), the pattern is different, with three orbitals being raised and two depressed. The energy difference between

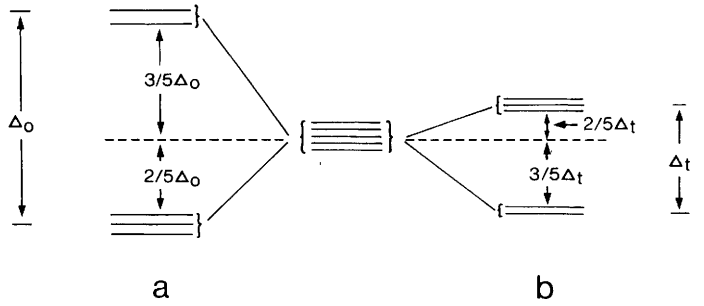


Figure 5.6 *Energy level diagram of the splitting of the d-orbitals of a transition metal ion as a result of* (a) *octahedral co-ordination and* (b) *tetrahedral co-ordination, according to the crystal field theory*
(Adapted from Cotton and Wilkinson, 1976; Figure 23-4. Copyright 1976 John Wiley & Sons, Inc. Reprinted by permission of the publisher)

the upper and the lower groups is not the same in these two cases (Δ_o is the energy difference in octahedral splitting, Δ_t the tetrahedral: see Figure 5.6). Now when the outer *d*-electrons are being allocated to orbitals, there is not a free choice for the ion. For our ion with three *d*-electrons to distribute, in the octahedral case all three can occupy one each of three lower orbitals, but in the tetrahedral case there is a choice between putting two in one each of the lower and the third in one of the upper orbitals, or pairing two electrons (one 'spin up' and one 'spin down') into one of the lower orbitals. These two choices are known as 'high spin' and 'low spin' respectively, since the first has three unpaired electrons, and the second only has one. The number of paired electrons in the *d*-orbital is of paramount importance in dictating the magnetic properties of the compound, and is the basis for an atomic explanation of the magnetic properties of materials.

The splitting of the *d*-orbital energies by the presence of other atoms has one further point of central importance here. It now becomes possible for electrons in the lower energy state to be promoted to the upper energy state, and to drop down again, emitting a quantum of electromagnetic energy equal to the energy difference between the split orbitals, as described in Chapter 2. (In fact, this so-called *d–d* band transition is forbidden by the laws of quantum mechanics, but it happens anyway because of quantum mechanical tunnelling.) The energy difference between the split orbitals (Δ_o or Δ_t) is small, but when converted into a wavelength using the usual formula [$E = h\nu$, or $E = hc/\lambda$, where E is the energy difference between the orbitals, h is Planck's constant (6.626×10^{-34} J s), c is the speed of light (3×10^8 m s^{-1}) and λ and ν are

the wavelength and frequency of the emitted radiation, respectively] it transpires that the electromagnetic radiation produced is in the visible region of the spectrum, unlike much of the other radiation arising from transitions between atomic orbitals. This then is the source of the colour in most inorganic compounds. One corollary is that, because the energy difference between the split orbitals depends on the geometry of the co-ordinating ligands, the colour of a compound can change if there is a change in co-ordination, *i.e.*, if an ion is surrounded by six water molecules in octahedral co-ordination, the colour will be different to that produced by the same ion surrounded tetrahedrally by only four water molecules. Likewise, if one or more of the water molecules is replaced by another ligand, then the strength of the electrostatic interaction will change, the energy splitting of the *d*-orbitals will change, and so the colour will again change.

In view of this, let us return to our simple experiment. When the crystals of hydrated copper sulfate dissolve in water, the simple formula:

$$CuSO_4 \rightleftharpoons Cu^{2+} + SO_4{}^{2-}$$

is an acceptable version of the truth in most cases, but it implies that there is a 'bare' copper ion free to move through the solvent water molecules, which is not in fact the case. Because water molecules develop a dipole moment, with the oxygen becoming slightly negatively charged and the two hydrogens slightly positive, the slight negative charge on the oxygen is attracted to the positive charge on the Cu^{2+}. Geometrical considerations mean that only six waters can cluster around the copper ion. This means that the copper(II) ion exists in water as the hexaquo ion $[Cu(H_2O)_6]^{2+}$, with the six waters octahedrally distributed around the copper ion. (The notation copper(II), Cu^{II}, and Cu^{2+} all imply the divalent copper ion). The familiar pale blue-green colour is characteristic of *d-d* band transitions within this octahedral copper(II) hexaquo complex. When the same crystals are dissolved in water containing various additions of ammonia, there is a gradual replacement of the solvating water molecules by ammonia molecules, resulting in a series of complexes ranging from $[Cu(NH_3)(H_2O)_5]^{2+}$ up to $[Cu(NH_3)_4(H_2O)_2]^{2+}$ resulting in a deeper blue colour as the number of ammonia groups increase. (It is usually very difficult to entirely replace all the waters in copper(II) complexes). This is an example of the chemical nature of the co-ordination sphere altering the colour of a transition metal ion.

It is not difficult to translate this knowledge to an explanation of the colour of transition metal ions in glass. It is conventional to talk of the glass as a solid 'solvent' for the colouring ions, which either exist as

interstitial network modifiers (just like Na^+) or as substitutes for the silicon in the network, and to discuss the co-ordination of the non-bridging oxygens around the metal ion in just the same way as the water was treated above. Interstitial ions usually have at least six or more oxygens surrounding them, and are approximately octahedrally co-ordinated. Certain ions (such as Fe^{3+}) can also substitute for the silicon in the network itself, and thus will have fourfold (approximately tetrahedral) oxygen co-ordination. From the above arguments, it becomes obvious to assume that the colour of a glass will change if the oxygen co-ordination changes from four to six – the equivalent in a crystal of changing from tetrahedral to octahedral co-ordination, although in a glass the geometry will not be quite so regular. According to Weyl (1976: see also Table 5.3), for example, if cobalt ions replace the sodium ions interstitially in a silicate glass, the resulting colour will be pink. If the same ions substitute for the silicon in the network, then the colour will be deep blue. Nickel ions give a yellow colour if they act as network modifiers, but purple if acting as network formers. Normally, the substituting ion will distribute itself between both possible sites, depending on the details of the manufacturing process and the furnace conditions, and the result for nickel would be a mixture of yellow and purple colours, which would normally balance to give a greyish glass. A secondary influence might also be the nature of the other network modifiers present, since these may also have an influence on the crystal field experienced by the colouring ion. It is conceivable, for example, that the colour developed by an ion in a glass with soda as the only alkali might be different from that given by the same ion in the same glass, but with potash substituted as the alkali. This factor has received relatively little attention, in view of the lack of commercial interest in potash glasses, although it may be historically relevant because of the use of 'forest glass' in the great cathedrals. A full treatment of the influence of the co-ordination of transition metal ions on the colour of glasses is given in Bamford (1977) and Paul (1990). An interesting footnote to this debate is the observation, made by Newton (1978), that the use of beech wood ash to make coloured glass as advocated by the Medieval monk Theophilus (Hawthorne and Smith, 1979; Book II, Chapter 4) might have been influenced by the fact that all the colours required (bar red, which used additional copper) could be obtained from glass made in such a way simply by controlling the furnace conditions, *i.e.*, without any explicit addition of colourants. It might also explain the observation that deep blue glasses (coloured by cobalt) were made principally from soda glass, either because potash glass does not give such a deep colour (an idea which could be tested experimentally), or because of conservatism

in the use of recipes. Clearly, whatever the answer to this, Medieval glassmakers were adept co-ordination chemists!

Colouration by metal atoms as opposed to ions in glasses is less complicated chemically, but the physics of the situation tends to be more involved. The more noble of the transition metals (Cu, Ag, and Au) can exist in glasses as metallic clusters; Cu and Au give rise to the highly-prized red colours in glasses and glazes, and silver is responsible for the yellow colour in the true Medieval *stained glass*. In red glass, the colour is imparted by a fine dispersion of precipitated metal throughout the glass, and is a result of scattering from the particles rather than absorption. The colour depends on the size of the precipitated particles, which in turn depends on the cooling programme adopted in the manufacturing process (Weyl, 1976). One of the technical marvels of the ancient glassmakers art, the 4th Century AD Lycurgus Cup, which is dichroic – red in transmitted light but green in reflected light – is coloured by a very fine distribution of small (<10 nm) metallic particles, made up on average of 66% Ag and 31 % Au, with 3 % Cu (Barber and Freestone, 1990). Silver in glass is particularly important in a discussion of Medieval glass, because it is the cause of the yellow colour in later Medieval *stained glass* – earlier glass is more correctly termed *painted*, because an opaque paint is simply applied to the surface of a transparent piece of glass. The manufacturing process for stained glass is that the silver is diffused into the surface after the base glass has been made, a process said to have been invented in Ulm in 1460 AD (Weyl, 1976). The process is known as *cementation*, and involves the application of a paste of a silver compound (usually the oxide mixed with a clay binder) to the surface of the glass, and heating to about 600 °C, at which temperature the silver ions diffuse into the glass by exchange with the sodium ions. Once the diffusion has taken place, the silver ions are reduced to the metal by a *redox reaction* with the other transition metal ions present (Fe^{II} or similar), as follows:

$$Ag^{+} + Fe^{2+} \rightleftharpoons Ag + Fe^{3+}$$

The final stage is to control the temperature programme in order to promote the growth of the silver droplets to the size required to produce the yellow colour, in the same way as the production of ruby red glasses.

Clearly then, in glasses coloured by metal ions, the co-ordination chemistry of the transition metal ion has a major influence on the colour. The other major influence is the oxidation state of the metal ion, since variable valency is another characteristic of the transition metals. All other things being equal, for example, iron in the Fe^{II} form will give a pale blue colour, whereas Fe^{III} gives rise to a brown colour in tetrahedral

co-ordination. The two principal factors which control the oxidation state of the transition metals are the redox conditions in the furnace (*i.e.*, the partial pressure of oxygen in the furnace atmosphere, or more simply the degree of oxidation or reduction achieved) and redox reactions between the various transition metal ions present in the melt – this is returned to below.

Control of the furnace conditions, together with the rate of cooling, is the traditional kernel of the glassmaker's art, and the primary means of control over the colour and quality of the product – more specifically, the glassmaker's art consists of matching the production process to the recipe to produce the desired result. Up until recent times, all glasses contained impurities from the raw materials, and the production of truly colourless glass required careful control over the recipe and the furnace conditions. The most important trace elements in this respect are iron and manganese. Most ancient and historical glasses contain significant levels of iron (0.3–1.5 wt % oxide) as an impurity from the sand used, which imparts the 'natural' green tinge in archaeological glass. Under oxidizing furnace conditions, iron forms the Fe^{III} species, which usually imparts darker green or brown colours, whereas under reduction Fe^{II} predominates, giving paler blues and greens. The reason for this actually relates to the previous discussion – Fe^{III} has an ionic radius of 0.64 Å compared to 0.74 Å for Fe^{II}, and is more likely to occupy the tetrahedral site of silicon (ionic radii of Si^{IV} is 0.42 Å) in the network, whereas Fe^{II} is more likely to occupy the larger octahedral site vacated by sodium (Na^{I} ionic radius 0.97 Å).

The exact redox state of the iron in the melt depends on the redox potential of the furnace (strictly, the partial pressure of oxygen in the atmosphere: Paul, 1990; 148), but also on the total amount of iron present, and on the redox reactions between the various possible redox couples in the melt – the most important one being that between manganese and iron. The equilibrium between Fe^{II} and Fe^{III} is affected by the total amount of iron present in the glass – as the total iron increases, the Fe^{III} state predominates over the Fe^{II} (Jackson, 1992), thus darkening the colour as the iron increases. Manganese in a glass can act as both a colourant and as a decolorizer – the latter partly by oxidizing the iron present, and partly by compensating for the resulting green colour by its own purple colour in the oxidized Mn^{III} state – Mn^{II} is a faint yellow or brown, and not a strong colourant. The Mn–Fe redox couple is as follows:

$$Mn^{3+} + Fe^{2+} \rightleftharpoons Mn^{2+} + Fe^{3+}$$

or using an oxide formulation:

$$Mn_2O_3 + 2FeO \rightleftharpoons 2MnO + Fe_2O_3$$

Under the conditions normally found in a glassmaking furnace, the equilibrium lies well over to the right – so much so, in fact, that Mn_2O_3 does not usually occur until all the iron present is in the Fe^{III} form. Hence it is conventional to report the manganese present in glass in terms of MnO, and the iron as Fe_2O_3, although it can be seen that this may not be entirely true. The resulting colour is therefore due largely to the Fe^{III}, which, although it gives a dark colour, is a weak chromophore (*i.e.*, a lot of Fe^{III} is needed to give a measurable colour). The well-known decolorizing effect of manganese requires excess manganese to be present in the Mn^{II} state, when the purple colour counteracts the green. More generally, Bamford (1977; 82) gives the following table to help work out which oxidation states will predominate when any two transition metals are present in a melt:

$$\Rightarrow$$
$$Cr^{6+}/Cr^{3+}$$
$$Mn^{3+}/Mn^{2+}$$
$$Ce^{4+}/Ce^{3+}$$
$$As^{5+}/As^{3+}$$
$$Sb^{5+}/Sb^{3+}$$
$$Fe^{3+}/Fe^{2+}$$
$$\Leftarrow$$

If a particular pair of transition metals are present (such as Cr and Fe), then the lower of the pair in this table (in this case, iron) will tend to exist as the left hand (oxidized) form of the ion (Fe^{3+}) and the upper of the pair will be in the right hand (reduced) state (Cr^{3+}). This tendency is shown by the direction of the arrows. In most ancient glasses iron is present, so the iron will always tend to be in the Fe^{III} state, and any other transition metal in its lower oxidation state. This is important not just from the point of the colour, but also because in glasses with a high transition metal content (such as vitrified nuclear waste or natural glasses, *i.e.*, obsidian) the redox state of these metals can influence the durability (Jantzen and Plodinec, 1984). For example, Fe^{III} improves the durability, whereas Fe^{II} reduces it (see below).

It has been repeatedly demonstrated that a straightforward chemical analysis using a technique which simply measures the total iron or manganese present (as is usually the case, even if the data *appears* to be in the form of oxide measurements) does not give an adequate guide to the colour, and that a knowledge of the exact oxidation states of the transition metal ions is required (Sellner *et al.*, 1979). This implies that

techniques such as Mössbauer spectroscopy or electron spin resonance are useful in this context, or direct measurement of the visible absorption spectrum of the glass compared to the absorption spectra of the various transition metal ions (Green and Hart, 1987).

It is clear that the colour of a glass is the result of a complex interplay between the co-ordination of the transition metal ions, and by redox reactions between the various ions present, and the redox potential in the furnace. The traditional archaeological view that colour can be simply related to the presence of various 'colouring agents' can only be regarded as a very crude guide.

THE DECAY OF MEDIEVAL WINDOW GLASS

The *durability* of glass may be defined as its resistance to attack by water, aqueous acid or base solutions, steam, or atmospheric agents. The particular case of attack of glass by water combined with atmospheric gases (*e.g.*, SO_2, CO_2) is termed *weathering*. The resistance of glass to attack by these agents is assessed commercially by several standard tests described in detail by Volf (1961). In general these involve grinding and sieving the glass to a specified size, subjecting a known weight to extraction by a fixed volume of liquid at a known pH and temperature for a specified time, then determining the weight of each component extracted chemically. The resistance of a particular glass to any one of the above agents is then classified according to the amount of leached material found.

Over the last fifty years, much literature has appeared on the reactions of glass with water, and a comprehensive review of the subject can be found in Clark *et al.* (1979) and Newton (1985). Most of this work has been on simple systems (usually binary soda–silica or ternary soda–lime–silica glasses), but it has provided a good framework for the understanding of the processes involved. In the following sections, we review the evidence from studies on binary glasses, which has given an understanding of the processes involved. We then consider ternary systems (Na_2O–CaO–SiO_2 and K_2O–CaO–SiO_2) which allows a comparison to be made of the different behaviour of Na^+ and K^+ in such systems. We briefly consider the rôle of phase separation in the context of ionic diffusion through glasses, and then review the evidence for the durability of synthetic Medieval glasses, which leads to a short discussion of the behaviour of mixed alkali glasses. The section concludes with a summary of some work done on relating durability to bulk composition of 'real' Medieval glass from York Minster.

Douglas and Isard (1949) investigated the reaction of water with soda–

lime–silica glass and concluded that a double-diffusion process occurs, with H^+ ions from the water entering the network to establish electrical neutrality as Na^+ ions are removed. Strictly, this statement should be phrased in terms of the *hydronium ion*, H_3O^+, since this is the correct ionic form of the positive ion resulting from the dissociation of water, but it is conventional to refer to it as the H^+ ion, and it makes little difference except when dealing with ionic transport phenomena. If electroneutrality were not maintained by the above mechanism, an electrical double layer would form, rapidly bringing the process to a halt. The mathematics of this double diffusion process shows that if one diffusion coefficient is much less than the other, then the process may be regarded as a single diffusion problem with a diffusion coefficient similar to that of the slower ion. The electrical conductivity of silicate glasses is known to be due to the diffusion of alkali ions (Morey, 1954), and Douglas and Isard deduced that, since the activation energies for electrical conduction and for water leaching were found to be very nearly equal (0.77 and 0.79 eV respectively), the diffusion of sodium ions was the rate controlling factor.

The basis of the currently-accepted mechanism for the attack of glass by aqueous solutions was proposed by Charles (1958), after observing that vitreous silica is not attacked by water vapour even over long periods, whereas multi-component glasses show signs of leaching into the solution relatively quickly. He concluded that dissolution occurs via the 'terminal structure' of alkali ions (*i.e.*, those associated with the non-bridging oxygen sites) (Equation 5.1):

$$\begin{array}{c} | \\ -\mathrm{Si}-\mathrm{O}-\mathrm{Na} \\ | \end{array} + \mathrm{H_2O} \rightarrow \begin{array}{c} | \\ -\mathrm{Si}-\mathrm{OH} \\ | \end{array} + \mathrm{Na^+} + \mathrm{OH^-} \qquad (5.1)$$

This equation represents the hydrolysis of the salt of a weak acid. The sodium ion migrates into the solution, and the hydroxyl ion released in the glass can then attack the otherwise stable silica network, converting a bridging oxygen into a non-bridging site, thus disrupting the network (Equation 5.2):

$$\begin{array}{ccc} | & | & \\ -\mathrm{Si}-\mathrm{O}-\mathrm{Si}- & + \mathrm{OH^-} \rightarrow & \begin{array}{c} | \\ -\mathrm{Si}-\mathrm{OH} \\ | \end{array} + \begin{array}{c} | \\ -\mathrm{Si}-\mathrm{O^-} \\ | \end{array} \\ | & | & \end{array} \qquad (5.2)$$

The non-bridging oxygen ion so produced is now capable of dissociating another water molecule (Equation. 5.3):

$$\begin{array}{c} | \\ -\mathrm{Si}-\mathrm{O}^- \\ | \end{array} + \mathrm{H_2O} \rightarrow \begin{array}{c} | \\ -\mathrm{Si}-\mathrm{OH} \\ | \end{array} + \mathrm{OH}^- \qquad (5.3)$$

These equations may be summarized as follows, but the intermediate stages involving the non-bridging oxygen sites are of crucial importance, as demonstrated by the observed resistance of vitreous silica to aqueous attack (Equation 5.4):

$$\begin{array}{c} | \\ -\mathrm{Si}-\mathrm{O}- \\ | \end{array} \begin{array}{c} | \\ \mathrm{Si}- \\ | \end{array} + \mathrm{H_2O} = 2 \left[\begin{array}{c} | \\ -\mathrm{Si}-\mathrm{OH} \\ | \end{array} \right] \qquad (5.4)$$

Excess hydroxyl ions are produced during the reaction, and, depending on the environment of the glass, they may accumulate in the corrosion layer, increasing the pH of the attacking solution and accelerating the dissolution of the network. Once the pH has risen to greater than 9, the silica network begins to break up and silicon is removed into solution as $Si(OH)_4$. Ion exchange of H^+ for Na^+ is referred to as Stage I of the process, and the break-up of the network Stage II (Clark *et al.*, 1979). These equations explain the following observations:

(i) Corrosion of glass proceeds faster in basic solutions.
(ii) Pure silica is resistant to aqueous attack at around neutral pH.
(iii) Glasses are normally resistant to acid attack (except HF, where the fluoride ion behaves like the hydroxyl group), since the hydroxyls produced tend to be neutralized before the network is attacked.
(iv) Faster corrosion rates are observed with steam than with water at the same temperature because dilution effects on the hydroxyl concentration are reduced.

Charles also made another important observation, namely that corrosion of an expanded glass structure (*i.e.* resulting from a high fictive temperature) proceeds faster than corrosion of a compacted glass structure (annealed) even if conditions and glass composition are the same. Thus the weathering characteristics of a piece of glass also depend on its thermal history.

Douglas and El-Shamy (1967) studied the rate of leaching of alkali and silica from binary glasses as a function of pH, and El-Shamy *et al.* (1972) extended this work to include soda-lime-silica glasses. Results for a binary potash-silica glass are shown in Figure 5.7. Below pH 9 the rate of alkali extraction is approximately constant and that of silica negligible.

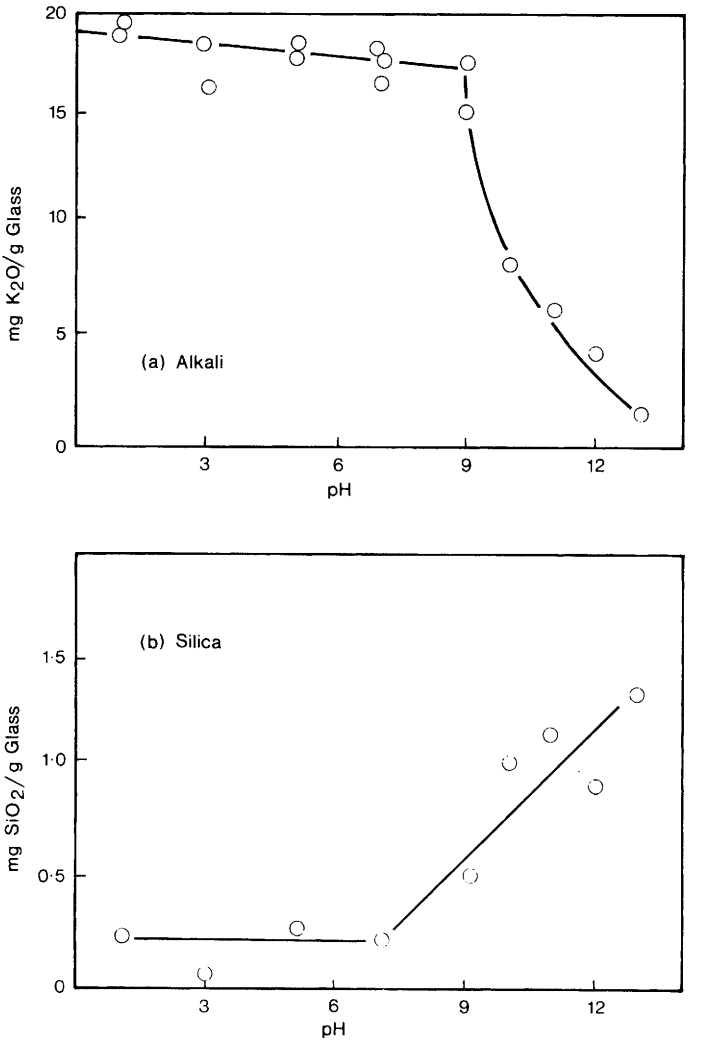


Figure 5.7 *pH dependence of alkali and silica extraction from binary $15K_2O/85SiO_2$ glass at 35 °C. The y-axis shows the amount of alkali and silica extracted from a given weight of glass over a fixed time period*
(Douglas and El-Shamy, 1967; Figures 5 and 6, reprinted by permission of the American Ceramic Society)

This corresponds to attack via Equation 5.1, with the concentration of OH^- ions being insufficient to break up the silica network as described by Equation 5.2. Above pH 9 the rate of alkali extraction falls and the quantity of silica extracted rises. The rate of extraction of silica depends

strongly on the alkali content of the glass and the pH of the attacking solution. Douglas and El-Shamy confirmed that the alkali in solution produced by ion exchange in the glass affected the rate of extraction of silica, as proposed by Charles, by leaching identical glasses in water in two vessels, one containing a cation exchange resin to remove the alkali ions. The extraction rate of silica in the alkali-free solution was significantly less than in the alkaline solution. Thus the rate of extraction of silica depends on the rate of removal of alkali. Addition of calcium oxide to the glass reduced the quantity of alkali extracted provided the CaO content of the glass did not exceed 10 mole %. No calcium was detected in the solution. Above 10 mole % calcium was extracted at pHs below four, with a corresponding increase in the alkali extraction. This reaction is described by the Equation 5.5:

$$\begin{array}{c} | \\ -\mathrm{Si}-\mathrm{O}-\mathrm{Ca}-\mathrm{O}- \\ | \end{array} \begin{array}{c} | \\ \mathrm{Si}- \\ | \end{array} + 2\mathrm{H}^{+} = 2 \left[\begin{array}{c} | \\ -\mathrm{Si}-\mathrm{OH} \\ | \end{array} \right] + \mathrm{Ca}^{2+} \qquad (5.5)$$

The terminal network sites so produced then provide a path for the easy movement of alkali ions, explaining the accompanying increase in alkali extraction.

In a comprehensive series of experiments, Douglas, El-Shamy and co-workers studied the leaching rates of powdered glasses with differing compositions in aqueous solutions. The apparatus and experimental method used is described by Rana and Douglas (1961a). These authors (Rana and Douglas, 1961a, 1961b) studied the rates of extraction of all components from soda-lime-silica and potash-lime-silica glasses in water, and concluded that two mechanisms were operating:

(i) Over short time periods and at low temperatures the rate of alkali extraction varied with the square root of time.
(ii) Over long time periods the amount of alkali and silica extracted varied linearly with time.

It was found that the rate of extraction of potash from potash glasses was always greater than the extraction rate of soda from the equivalent soda glass, even though the electrical conductivity of the potash glass is lower. Douglas and El-Shamy (1967) also observed that the ratio of alkali to silica in the solution is greater than the ratio in the glass, implying a preferential removal of alkali. This must lead to a leached layer being formed on the glass which is depleted in alkali, the thickness of which depends on time and the composition of the glass. For glasses of

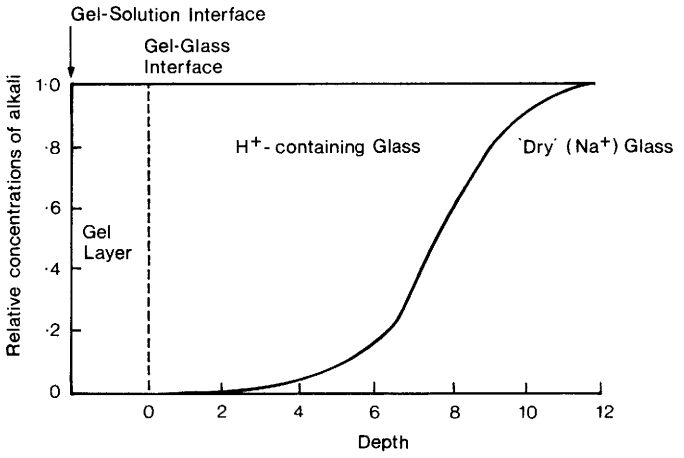


Figure 5.8 *Model of hydrated glass surface showing the development of the hydrated layer* (Doremus, 1975; Figure 3, by permission of Elsevier Science)

low durability the leached layer may extend through the whole bulk of the glass. The transition from a square-root dependent time coefficient to a linear dependence can be understood in terms of the time taken for a leached layer to reach equilibrium. Initially, the process will be diffusion-controlled as the alkali and hydrogen ion exchange occurs. When equilibrium is reached, there are two reaction planes, the gel/solution interface (*gel* is the term applied to the porous silica network) and the glass/gel boundary. A simple diagram of the situation is given in Figure 5.8, from Doremus (1975). At the gel/solution interface, the porous silica network is being broken down by the action of hydroxyl groups, at a rate which depends on the OH^- concentration at the surface. This concentration depends on the pH of the attacking solution, the composition of the glass and the experimental conditions. The gel layer itself is an open network allowing relatively free ionic movement. At the glass/gel interface, the ion exchange reaction is proceeding, with H^+ ions entering the network and alkali ions leaving it. Thus at equilibrium there is a leached layer (the gel) of constant thickness, which progresses steadily into the glass as the silica network at the gel/solution interface is destroyed. The reaction between glasses and water can be treated mathematically as a double diffusion problem with a moving boundary, as shown by Frischat in his monograph on ionic diffusion in oxide glasses (1975). Subsequent work by Hench and co-workers has extended this simple picture by defining six types of glass surface

resulting from the attack of glass by water (summarized by Newton, 1985), which model more closely the behaviour of real glasses in contact with water of varying pH.

An important advance in the understanding of the chemical behaviour of glasses in aqueous solution was made in 1977, when Paul (1977) published a theoretical model for the various processes based on the calculation of the standard free energy (ΔG°) and equilibrium constants for the reactions of the components with water. This model successfully predicted many of the empirically-derived phenomena described above, such as the increased durability resulting from the addition of small amounts of CaO to the glass, and forms the basis for our current understanding of the kinetic and thermodynamic behaviour of glass in aqueous media.

In addition to the composition of the glass, the corrosion resistance is determined by the degree of phase separation present and the nature of the surface before the glass is subject to corrosion. Other factors which may affect the behaviour are the presence of any crystalline particles, either due to heat treatment or from incomplete melting of the raw materials, and the presence of any stresses, either internal or external. The surface composition of glass can be modified by reactions which occur during the manufacturing process, and on contact with the atmosphere. These include the migration of alkali ions to the surface during manufacture to satisfy the 'dangling' bonds caused by termination of the silica network, partially compensated by the high temperature volatility of the alkalis. Fresh glass surfaces are also de-alkalized very rapidly by contact with atmospheric water vapour. Several techniques are available for determining the composition of the surface layer – refractive index measurements (Pfund, 1946), infrared reflection spectroscopy (IRRS) (Sanders *et al.*, 1972), infrared microspectroscopy (Cooper *et al.*, 1993), Auger electron spectroscopy (Dawson *et al.*, 1978), and electron loss spectroscopy (ELS; Pollard *et al.*, 1980). All these methods confirm that the chemical composition of the surface is a complicated function of age and thermal history.

The influence of phase separation on the physical properties of glass has been summarized by Porai-Koshits (1977): '*the viscosity of a two phase glass depends on the character of distribution of the high viscous phase, the electrical conductivity – on the distribution of the conducting phase, the chemical durability – on the distribution of the chemically more unstable phase.*' The effect of microstructure in the sodium silicate system has been studied by Redwine and Field (1968), who observed two types of phase structure after differing heat treatments of glasses with the same composition – two interconnected phases, or a dispersion of particles in a continuous phase. In the second

case, the properties of the glass were very similar to those of the continuous phase, but in the first case the properties depended on the amount and composition of the low-soda phase. Frischat (1975) investigated the effect of phase separation on the diffusion of ^{22}Na through a soda-silica glass, and concluded that two diffusion paths are possible, one through the soda-rich phase, and one with a lower diffusion coefficient through the silica-rich phase. Reviews of the phase separation tendencies of multi-component glasses are given by Doremus (1973) and Vogel (1977). The magnitude of phase separation is controlled by the addition of certain substances to the melt, as well as by heat treatment. One of the most important additives, as already noted, is alumina, which, when added in small quantities (< 2.5%), considerably reduces the tendency to phase separation. This is the basis of the manufacture of chemically resistant silicate glasses (Volf, 1961). Of importance to the study of ancient glasses, particularly 'forest glass' is the fact that phosphorus pentoxide (P_2O_5) is known to increase the tendency to phase separation, as is the addition of small quantities of halogens to the melt (Vogel, 1977).

Modern chemically stable glasses have been developed empirically over the years, well in advance of the theoretical understanding of their chemical properties. Many of the glasses manufactured today for technological purposes are not in the soda-lime-silica system (*e.g.* Pyrex), but the poor durability of Medieval glass has been known for many years. In 1907, Heaton gave a paper to the Society of Arts (Heaton, 1907) in which he discussed the chemistry of Medieval glasses. He identified the main alkali as potash, derived from wood ash instead of seaweed, which produces a soda glass. The wood ash also supplied lime, magnesia, and phosphorus pentoxide. The diversity of composition of Medieval glasses is attributable to the variability of composition of burnt wood, not just from species to species but even from different parts of the same tree. El-Shamy (1973) studied the reactions with aqueous solutions of simulated Medieval glasses (quaternary K_2O–CaO–MgO–SiO_2 glasses) with systematically varying composition. Replacement of SiO_2 by CaO (above 10 mole % CaO) resulted in an increased extraction of all oxides except SiO_2 in acid solution, even though the MgO and K_2O contents were unchanged. In acid solution the leaching of K_2O, CaO, and MgO was greater, in accordance with Equation 5.5. Replacement of SiO_2 by MgO had the same effect. Replacing SiO_2 by K_2O caused a considerable increase in the extraction of all oxides in water. In glasses with 15 mole % K_2O, complete de-alkalization was observed within 72 hours. Replacement of K_2O by CaO or MgO has the effect of reducing the rate of alkali extraction by one quarter for 5% replacement. It seems

that CaO and MgO are interchangeable, as replacement of CaO by MgO had no effect.

El-Shamy noticed a sudden decrease in the chemical resistance of these glasses when the SiO_2 content fell below 66 mole % SiO_2, and related this to the geometry of the structure. Below this point, every silicon atom becomes associated with a basic atom as second neighbour, so there is always an interconnecting path of non-bridging oxygen sites available for the diffusion process. This is the same argument as was used by Stevels (1960–61) to explain the importance of the point at which the number of bridging oxygen ions (Y) falls below 3.

The fact that potash glasses are considerably less durable than equivalent soda glasses is attributable largely to the greater ionic radius of the K^+ ion (1.33 Å compared with 0.95 Å for Na^+). Thus as the melt passes through the glass formation temperature, the network is formed around the much larger K^+ ions, resulting in a more open structure which is more easily attacked by water. The difference in corrosion resistance of sodium and potassium binary silicate glasses with 30 mole % alkali has been studied by Hench (1975). Initially (at t = 1 min), a silica-rich film was developed by de-alkalization, penetrating deeper into the bulk in the potash glass. After 100 min, the silica network of the potash glass was being dissolved, whilst the soda glass suffered de-alkalization but not dissolution. After a long time (10^3 min), both glasses were being destroyed because of the high alkali content. Addition of calcium and aluminium tended to cause a protective film on the glass, preventing dissolution.

When two alkalis are present together in a glass, the glass exhibits non-linearity in many of its physical properties as the ratio of one alkali to the other is changed, the total alkali being kept constant. This is called the *mixed alkali effect* (MAE), and is most pronounced in the properties which depend on ionic transport – electrical conductivity, diffusion, chemical durability. Day (1976) has reviewed the effect, and discussed the many theories which attempt an explanation. Most properties show a pronounced maximum or minimum at a particular alkali ratio, and Day states that mixed alkali glasses have a higher durability than single alkali glasses. He reports a minimum in the durability for a molar potash/soda ratio of 0.7–0.8. Hench (1975), however, claims that there is no minimum in the durability curve, and that previous observations of minima are due to presenting the results in terms of weight rather than molar percentages. He also observed that addition of calcium to the mixed alkali glasses did not produce a minimum. A more comprehensive experimental study of the MAE in mixed-alkali silicate and mixed-alkali

lime silicate glasses by Dilmore (1977) led to the following conclusions about the MAE:

(i) Some glasses do exhibit a minimum in the concentration of leached alkalis – in general these are glasses of high durability.
(ii) The presence of MAE is a function of the pH of the attacking solution. Below pH 9, selective leaching is the dominant mode of decay, and the MAE may be present. Above pH 9, total dissolution takes place and the MAE is unimportant.

Hence it is concluded that in multicomponent silicate glasses, the presence or absence of a MAE is a complicated function of composition and environment, and not necessarily entirely dependent on the ratio of one alkali to the other.

In a large analytical study of the relationship between the chemical composition and corrosion behaviour of Medieval window glass from York Minster (Pollard, 1979), a total of more than 200 samples were analysed, dating mainly from the twelfth, fourteenth, and fifteenth Centuries. A summary of the findings has been published (Cox *et al.*, 1979). The glasses were visually classified according to weathering characteristics into four broad categories: unweathered (u), pitted (p), pitted-and-crusted (p/c), and crusted (c). This classification is a gross simplification of the true variation in weathering characteristics, since no two corroded samples show exactly the same features. In particular, the pitted-and-crusted group contains a wide variety of specimens: some exhibit pits which contain a white deposit, whilst others show a severe loss of volume caused by total dissolution of the surface. Obviously, these four groups are extremely crude and can only be considered as a rough guide. It is also possible that some of the glasses which now exhibit pure pitting once had a crust in the pits and have subsequently lost them in the course of restoration, since the history of each piece is not precisely known.

The samples were chosen to give as broad a spread as possible of dates and weathering characteristics, but the group is not representative of the true range available in York Minster. No thirteenth Century glass was available at the time although several large windows of thirteenth Century work exist in the Minster. Only three glasses from the sixteenth Century were analysed, as this period marks the end of the Medieval period. One piece of eleventh Century glass was added to the group – this was an excavated deep blue glass which was completely unweathered. Amongst the twelfth and fourteenth Century glasses the most common decay mechanism is pitting, with relatively few cases of

crusting, and few unweathered pieces surviving. By the fifteenth Century the majority of pieces are unweathered, with crusting in particular becoming very rare.

The complete results of the quantitative analysis of the samples by *X*-ray fluorescence are given in Pollard (1979). The results of the XRF analysis were converted from weight to molar percentages, which were then used to calculate triangular co-ordinates according to the formula based on that given by Iliffe and Newton (1976). The three axes are network formers 'SiO_2', alkali network modifiers ('R_2O') and alkaline earth network modifiers ('RO'). The formulae for calculating these co-ordinates from the molar percentages are:

$$\text{'}SiO_2\text{'} = SiO_2 + P_2O_5 + 2(Al_2O_3 + Fe_2O_3)$$
$$\text{'}R_2O\text{'} = Na_2O + K_2O - (Al_2O_3 + Fe_2O_3)$$
$$\text{'RO'} = CaO + MgO + MnO + CuO + ZnO + PbO$$

The justification for adding twice the molar concentration of the trivalent oxides to the 'SiO_2' figure is that they can enter the network and also immobilize the alkali cations, which is why they are also deducted from the 'R_2O' value.

The data are most simply summarized in the triangular diagram shown in Figure 5.9. It is immediately apparent that there is a correlation between weathering behaviour and composition: the 60% 'SiO_2' value is the lower limit for durable glasses. This diagram also shows another striking feature, that is the existence of a small group of high 'SiO_2', low 'RO' glasses, toward the top left hand side, well removed from any other group in the sample. This group contains five samples, which form a distinctive set in that they are all soda glasses, all deep blue and all date from the twelfth Century. The existence of such a group is of great historical interest, and poses several problems for the art historian. The high sodium, low magnesium, phosphorus, and potassium values of these glasses indicate a different manufacturing process, and possibly a different place of origin from the main body of the glasses analysed, perhaps a coastal site. The composition is very similar to that of Saxon glass as reported by Cramp (1970), but the pieces are larger than the common Saxon samples found. Possibly the glass was remade from Saxon glass, or possibly it was imported from a different source because of its rich colour, which the blue potash glass of the period does not seem able to imitate (see discussion above about the possible differences in colour between equivalent soda and potash glass). Some further work on this subject is discussed below.

It can be seen from Figure 5.9 that the chemical distinction between

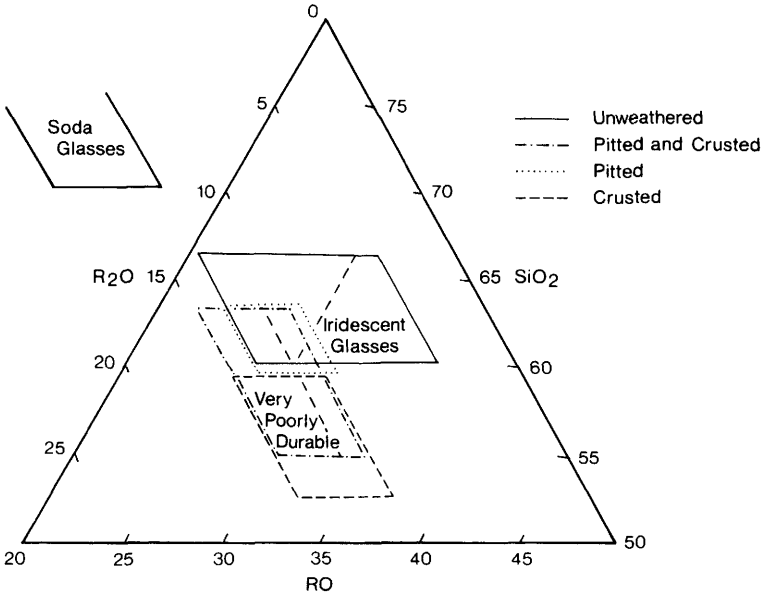


Figure 5.9 *Triangular diagram showing the distribution of the various weathering classifications of York Minster glass*
(Pollard, 1979; Figure 46)

pitted, pitted-and-crusted, and crusted glasses is a relatively arbitrary one. All these glasses are characterized by 'R_2O' values between 10 and 15%, and it seems again that 60% 'SiO_2' is an important figure. Above this, decay is predominantly by pitting, below it by crusting, but a large number of pitted-and-crusted glasses straddle this line indicating that there is no real distinction between the groups. Above 60% 'SiO_2' it is likely that the network formers are linked by enough bridging oxygen atoms to resist the general dissolution of the network, giving rise to localized decay (pits). Below this point the network is more disrupted by modifiers and a uniform removal of the surface is possible. In his study of the durability of synthetic Medieval glasses discussed above, El-Shamy (1973) observed a sudden drop in the chemical resistance of his glasses as the SiO_2 content dropped below $^2/_3$ of the molar composition. This silica content, according to El-Shamy, marks the point at which, on average, each silicon atom has a basic ion as second neighbour, leading to an interconnected path of cation sites, with a corresponding increase in ionic mobility. In the glasses analysed here, aluminium and phosphorus will act as network formers, and this may lower the critical value to the observed 60% 'SiO_2'. In addition to the main regions in Figure 5.9, two

further groupings are marked; a very poorly durable group characterized by high alkali content, and a group of otherwise unweathered glasses which show an iridescent surface, characterized by high alkaline earth content. Many glasses in this study were found to have lost a considerable amount of material from the surface, either uniformly over the whole surface or unevenly, as if by severe pitting. These glasses are examples of the poorest durability, although it is quite possible that glasses of even worse durability were manufactured, but have not survived. Despite the fact that great care was taken to produce a representative surface (Cox and Pollard, 1977), it is also possible that glasses of such low resistance to corrosion have moved on the triangular diagram as the alkali has been leached from the whole bulk, resulting in a glass which has an artificially lower 'R_2O' and a higher 'SiO_2' co-ordinate.

The types of weathering behaviour observed in the glasses from York Minster are not necessarily typical of the decay of Medieval glass in general. Several weathering phenomena commonly observed in other cathedrals are rare in York. In particular, glasses forming a thick heavy crust are not at all common in the Minster. As part of an attempt to extend the study to these other types, further analyses were carried out on generally unweathered Medieval glass. When plotted on the triangular diagram, four (two each from Winchester and Chartres) were found to occupy positions typical of the blue twelfth Century soda glasses found in York. Again, all were dark blue twelfth Century, and all are soda glasses (9.5–17.4 mole % Na_2O), with low MgO (0.6–0.8 mole %), low K_2O (0.4–1.2 mole %) and low CaO (7.4–8.9 mole %). These characteristics agree very well with the York group, and we must conclude that dark blue soda glasses were fairly widespread during the twelfth Century. Subsequently, Cox and Gillies (1986) studied more 12th Century blue soda glass from York, and concluded that it was not rare in the Minster, and postulated that the glass was Roman (or slightly later), re-melted in France, possibly with the addition of some potash glass, and imported into England. This ties in with the brief description given around 1100 AD by Theophilus (Hawthorne and Smith, 1979; Book 2 Chapter 12), where he describes '*different kinds of glass ... found in mosaic work in ancient pagan buildings*' and also '*various small vessels ... which are collected by the French ... They even melt the blue in their furnaces, adding a little of the clear white to it, and they make from it blue glass sheets which are costly and very useful in windows.*' There appears to be good evidence from modern analytical data to substantiate this description, and the widespread distribution of such glass suggests that the re-use of ancient glass may have been more common than was previously thought.

The structural chemistry of the weathering crusts themselves have

been studied using a combination of *X*-ray diffraction, infrared spectroscopy, atomic absorption analysis, and electron microscopy (Gillies and Cox, 1988a, 1988b; Schreiner, 1988). A number of phases were identified, most commonly gypsum ($CaSO_4.2H_2O$), syngenite [$K_2Ca(SO_4)_2.H_2O$], hydrated silica ($SiO_2.xH_2O$), and calcite ($CaCO_3$), but no pattern was found between the mineralogy of the corrosion products and the climatic conditions or orientation of the window. The best relationship appeared to be that between the bulk composition of the glass and the identity of the weathering products.

The main conclusion of this study is that the weathering behaviour of Medieval glass is dictated primarily by chemical composition, and does not depend significantly on the aspect of the window or the level of pollution in the atmosphere, as has often been thought. The distinction between pitted and crusted glasses is shown, in general, to be an arbitrary one, with the two groups coalescing into a band on the triangular diagram, and exhibiting a slow gradation of weathering characteristics. It is here that the environment of the glass may have some influence on the corrosion, governing whether a crust forms or not as the surface is attacked by water. In general, the conclusions here agree reasonably well with the data from accelerated corrosion work on simulated Medieval glasses, although the figure from this of 66.67 mole % SiO_2 as the lower limit of good durability is higher than the experimental figure. In this study, a figure of 60% 'SiO_2' seems more appropriate, where the 'SiO_2' co-ordinate includes the intermediate oxides of aluminium and phosphorus, which may explain the difference.

THE CORROSION OF BURIED GLASS

Although buried glass is of more general interest in archaeology than window glass, the detail of its corrosion behaviour has received considerably less attention in the archaeological literature. This is in part due to the general assumption by archaeologists that the glass which has survived is relatively inert, although this is not necessarily the case (Newton and Davison, 1989), but also because of the complexity of the soil environment (Pollard, 1993). The interaction of groundwaters and vitreous material is, of course, crucial to the long-term stability of buried vitrified nuclear waste, and this aspect has been extensively studied using accelerated corrosion techniques (Lutze, 1988). The value of the study of natural (geological) glasses has long been understood in the context of validating these models, but more recently it has come to be appreciated that a study of archaeological material also has a role to play (Cox and Ford, 1993).

The observation of the characteristic corrosion products on buried glass goes back at least as far as Brewster (1863), who reported the characteristic flakiness and surface iridescence. Under the microscope he observed a fine laminar structure with a range of thicknesses between 0.3 and 15 μm which he deduced was responsible for the iridescence via interference phenomena. Subsequent chemical and structural analyses of these multiple layers using electron microscopy, *X*-ray diffraction, and Raman microspectroscopy (Cox and Ford, 1993) has concluded that these layers are depleted in virtually all oxides except SiO_2, and consist of poorly crystalline hydrated silicates and aluminosilicates, with deposits of calcite, calcium phosphate, and manganiferous minerals on the surface, thought to be of external origin. Contrary to the early suggestion that these layers may reflect annual variations in ground conditions, and could be counted to give an age for the glass in the same way as dendrochronology (Brill and Hood, 1961), subsequent studies have shown a number of different morphologies of weathering crusts, including parallel layers, hemispherical layers, and zigzag banding (Cox and Ford, 1993). The likely explanation of this is that the dissolution products from the glass (hydrated silica, plus ions of the alkalis and alkaline earths, depending on the bulk composition of the glass) diffuse out into the attacking aqueous solution, exceed the solubility product as a result of pH changes, and precipitate as some form of hydrated, poorly crystalline silicate (or aluminosilicate) mineral, possibly incorporating other species which may be present from the attacking solution. The cyclic nature of the events which gives rise to the laminations may be explained by the precipitation event influencing the local pH and ionic concentrations in the solution, and thus temporarily inhibiting the precipitation process. It is clear from this that the process depends very strongly on local conditions, and not on any annual cycles, as was postulated by Brill and Hood (1961).

The precise nature of the chemical interaction between the buried glass and the local groundwaters (and, indeed, between any buried archaeological object and its environment) is exceedingly complicated to explain precisely, but is fundamental to our understanding of the long-term alteration and survival of archaeological evidence. Some attention has been paid to the survival of buried bone, since it is agreed that the long-term chemical and physical alteration of bone is largely a result of interaction with groundwater. This has been considered in some detail because of phenomena such as the *post mortem* uptake of uranium, an understanding of which is important for uranium series dating of bone and teeth and for estimating the dose rate in electron spin resonance dating of teeth (Aitken, 1990). Analytical studies of the diagenesis of the

inorganic bone fraction have concentrated on experimental observations of diffusion profiles measured in either archaeological bone or from bone immersed in known concentration synthetic solutions. These studies have been restricted to a consideration of the interaction of bone with water – what might be termed a 'local' study. Little attention has been paid to the wider view of bone (or any other archaeological artefact) in its geochemical environment. It is clear that the relationship between sedimentary context, groundwater composition, and artefact mineralogy is necessarily complex, but also that it is this holistic approach which must be developed if anything meaningful is to be achieved in terms of understanding and predicting the nature of the interaction.

A commonly expressed view is that the 'burial environment' of an object can in some way be reconstructed in the laboratory after excavation providing a sample is taken of the sediment containing the object. This is a gross oversimplification of the relationship between the soil chemistry, the groundwater chemistry, and the chemistry of the object itself. A chemical analysis of the sediment for an element such as strontium will give little indication of the level of strontium in the groundwater moving through the sediment and the object, even if the analysis is relevant in terms of assessing the 'available' strontium rather than the total strontium in the deposit – a point made in the context of bone by Williams (1988). Control of the uptake of elements from sediment into groundwater is far more subtle than this, and depends not only on the prevailing aqueous environment (temperature, pH, redox potential) but also on the mineralogical form of the element, and the concentration (more accurately, activity) of other dissolved species. In most natural dilute groundwaters, for example, the pH is entirely controlled by the $CaCO_3$–H_2O–CO_2 equilibria (Garrels and Christ, 1965; 75). In soil solutions, phosphate solubility is often controlled by the aluminium content (activity) of the soil (Lindsay, 1979; 169). The concentration of trace elements such as strontium in groundwaters will therefore almost certainly be controlled by the activity of other species – in this case, probably calcium and magnesium.

There is, of course, a very good reason for this lack of knowledge about the detailed interaction between soil, groundwater, and archaeological object. Field measurements of the important parameters such as pH, E_h (redox potential), dissolved CO_2 levels, and the various trace element concentrations are notoriously difficult to perform without unbalancing the equilibrium of the system being studied. Not only that, it is well established that an important factor affecting the chemical properties of groundwater is not simply the dissolved trace element or

cationic concentrations, but the form which the element is in – the *speciation*. For example, the actual chemical form of dissolved phosphate changes from being predominantly the PO_4^{3-} ion at pHs greater than 12 to the $H_2PO_4^-$ species between pH 2 and 7.5. This has an automatic knock-on effect on the stability of various soil minerals, which will in turn effect the composition of the groundwater, and so on. Although much of these basic data on soil chemistry have been published (*e.g.*, Lindsay, 1979), the situation rapidly becomes so complex that it becomes difficult to deal effectively with real systems. For over 20 years, geochemists have used computer modelling in order to provide further understanding of how such systems might behave. Computer models of the chemistry of dilute aqueous solutions such as surface groundwaters (as opposed to deeper brines) have been in existence since the mid-1960s, and are now commercially available and documented (*e.g.*, Bassett and Melchior, 1989). This type of program can be used to predict the stability fields for a large number of minerals over the range of naturally encountered conditions of pH and E_h. Although widely used in geochemistry, little use has been made of these programs in archaeology, with the exception of the work done by Thomas and colleagues in order to understand the corrosion behaviour of copper in the burial environment (Thomas, 1990). There is a clear need to apply these tools (together with a knowledge of chemical kinetics), not only to aid our understanding of the interaction between buried glass and its environment, but also to increase our understanding of the nature of the surviving archaeological record in general.

SUMMARY

A study of the chemistry and corrosion of glass provides ample opportunity for the application of a wide range of inorganic and physical chemical principles. The structure of glass was for many years a challenging area of research for physical chemists, with two competing theories – the continuous random network model of Zachariasen and coworkers and the 'crystallite' model of Lebedev. As with many scientific debates, the truth lies somewhere in between, and, in fact, when examined closely the two competing models merge into one another. With improved instrumentation for studying short range atomic order, the subtleties of the vitreous state, including the influence of phase separation on various properties such as durability, became much clearer, and the debate has now largely resolved itself. The explanation of the colours observed as a result of the presence of small amounts of transition metals in glasses is a particularly interesting application of

inorganic chemistry, since it involves a knowledge of co-ordination chemistry, crystal field theory, and redox reactions. Some of the most spectacular examples of the glassmakers art, however, also involve physical principles such as the dichroic effect produced by the scattering of light from a colloidal dispersion of gold-silver alloy precipitates in glass.

The corrosion and decay of Medieval window glass has received a great deal of attention for many years, and has provided a number of surprises. It has been shown, for example, that the decay of the potash-based coloured glass in the great cathedrals of northern Europe is not primarily caused by increased atmospheric pollution, but by aqueous attack on glasses of inherently low durability. One of the principal factors in this low durability is the composition of the glass itself – partly because of the lower durability of potash glasses as opposed to soda glasses, but probably more importantly because the total proportion of 'network formers' in the glass is lower than the critical value required to produce a continuous stable network. This conclusion has been validated by chemical analysis of the glasses themselves, by accelerated corrosion tests on synthetic glasses of Medieval composition, and by thermodynamic calculations of theoretical stabilities.

The corrosion of buried glass is considerably less well-understood by archaeologists, despite a considerable body of data collected as a result of the importance of a knowledge of the interaction between vitreous materials and groundwaters to the nuclear waste disposal industry. The basic process of the leaching out of glass components and the cyclic precipitation of poorly-crystalline silicate minerals can be explained, but the detail of the interaction between buried glass and its groundwater environment requires further modelling in order to become predictive or reconstructive. The value of chemical studies of the corrosion of archaeological glass is now appreciated as being of great value in validating the accelerated testing results of nuclear waste disposal materials, but it suggests that the natural long-term laboratory that is the archaeological record is not yet fully exploited in terms of its relevance to modern-day chemistry.

REFERENCES

Aitken, M.J. (1990). *Science-based Dating in Archaeology*. Longman, London.

Bacon, F.R. (1968). The chemical durability of silicate glass Part One. *Glass Industry* **49** 438–446.

Bamford, C.R. (1977). *Colour Generation and Control in Glass*. Glass Science and Technology 2, Elsevier, Amsterdam.

Barber, D.J. and Freestone, I.C. (1990). An investigation of the origin of the colour of the Lycurgus Cup by analytical transmission electron microscopy. *Archaeometry* **32** 33–45.

Bassett, R.L. and Melchior, D.C. (1989). Chemical modeling of aqueous systems: an overview. In *Chemical Modeling of Aqueous Systems II*, eds. Melchior, D.C. and Bassett, R.L., American Chemical Society, Washington, pp. 1–14.

Brewster, D. (1863). On the structure and optical phenomena of ancient decomposed glass. *Philosophical Transactions of the Royal Society of Edinburgh* **23** 193–204.

Brill, R.H. and Hood, H.P. (1961). A new method for dating ancient glass. *Nature* **189** 12–14.

Burnett, D.G. and Douglas, R.W. (1970). Liquid-liquid phase separation in the soda-lime-silica system. *Physics and Chemistry of Glasses* **11** 125–135.

Charles, R.J. (1958). Static fatigue of glass I. *Journal of Applied Physics* **29** 1549–1553.

Cooper, G.I., Cox, G.A. and Perutz, R.N. (1993). Infra-red micro-spectroscopy as a complimentary technique to electron-probe micro-analysis for the investigation of natural corrosion on potash glasses. *Journal of Microscopy* **170** 111–118.

Clark, D.E., Pantano, C.G. Jr. and Hench, L.L. (1979). *Corrosion of Glass.* Books for Industry and the Glass Industry, New York.

Cotton, F.A. and Wilkinson, G. (1976). *Basic Inorganic Chemistry.* Wiley, New York.

Cox, G.A. and Ford, B.A. (1993). The long-term corrosion of glass by ground-water. *Journal of Materials Science* **28** 5637–5647.

Cox, G.A. and Gillies, K.J.S. (1986). The *X*-ray fluorescence analysis of Medieval blue soda glass from York Minster. *Archaeometry* **28** 57–68.

Cox, G.A. and Pollard, A.M. (1977). *X*-ray fluorescence analysis of ancient glass – the importance of sample preparation. *Archaeometry* **19** 45–54.

Cox, G.A., Heavens, O.S., Newton, R.G. and Pollard, A.M. (1979). A study of the weathering behaviour of Mediaeval glass from York Minster. *Journal of Glass Studies* **21** 54–75.

Cramp, R. (1970). Decorated window glass and millefiori from Monkwearmouth. *Antiquaries Journal* **50** 327–335.

Dawson, P.T., Heavens, O.S. and Pollard, A.M. (1978). Glass surface analysis by Auger Electron Spectroscopy. *Journal of Physics C* **11** 2183–2193.

Day, D.E. (1976). Mixed alkali glasses – their properties and uses. *Journal of Non-Crystalline Solids* **21** 343–372.

Dilmore, M.F. (1977). *Chemical Durability of Multicomponent Silicate Glasses.* Unpublished PhD Thesis, University of Florida.

Doremus, R.H. (1973). *Glass Science.* Wiley, New York.

Doremus, R.H. (1975). Interdiffusion of hydrogen and alkali ions in a glass surface. *Journal of Non-Crystalline Solids* **19** 137–141.

Douglas, R.W. and El-Shamy, T.M. (1967). Reactions of glasses with aqueous solutions. *Journal of the American Ceramic Society* **50** 1–8.

Douglas, R.W. and Isard, J.O. (1949). The action of water and of sulphur dioxide on glass surfaces. *Journal of the Society of Glass Technologists* **33** 289–335.

El-Shamy, T.M. (1973). The chemical durability of K_2O–CaO–MgO–SiO_2 glasses. *Physics and Chemistry of Glasses* **14** 1–5.

El-Shamy, T.M., Lewins, J. and Douglas, R.W. (1972). The dependence on the pH of the decomposition of glasses by aqueous solutions. *Glass Technology* **13** 81–87.

Frank, S. (1982). *Glass and Archaeology.* Academic Press, London.

Frischat, G.H. (1975). *Ionic Diffusion in Oxide Glasses.* Aedermannsdorf.

Garrels, R.M. and Christ, C.L. (1965). *Solutions, Minerals and Equilibria.* Freeman Cooper, San Francisco.

Gillies, K.J.S. and Cox, A. (1988a). Decay of medieval stained glass at York, Canterbury and Carlisle. Part 1. Composition of the glass and its weathering crusts. *Glastechnische Berichte* **61** 75–84.

Gillies, K.J.S. and Cox, A. (1988b). Decay of medieval stained glass at York, Canterbury and Carlisle. Part 2. Relationship between the composition of the glass, its durability and the weathering products. *Glastechnische Berichte* **61** 101–107.

Greaves, G.N., Fontaine, A., Lagarde, P., Raoux, D. and Gurman, S.J. (1981). Local structure of silicate glasses. *Nature* **293** 611–616.

Green L.R. and Hart F.A. (1987). Colour and chemical composition of ancient glass: an examination of some Roman and Wealden glass by means of Ultraviolet-Visible-Infra-red Spectrometry and Electron Microprobe analysis. *Journal of Archaeological Science* **14** 271–282.

Greer, A.L. (1995). Metallic glasses. *Science* **267** 1947–1953.

Hawthorne, J.G. and Smith, C.S. (trans.) (1979). *Theophilus, De Diversis Artibus.* Dover Publications, New York.

Heaton, N. (1907). Mediaeval stained glass: its production and decay. *Journal of the Society of Arts* **55** 468–484.

Hench, L.L. (1975). Characterization of glass corrosion and durability. *Journal of Non-Crystalline Solids* **19** 27–39.

Holland, L. (1964). *The Properties of Glass Surfaces.* Chapman and Hall, London.

Holloway, D.G. (1973). *The Physical Properties of Glass.* Wykeham, London.

Huang, C. and Cormack, A.N. (1990). The structure of sodium silicate glass. *Journal of Chemical Physics* **93** 8180–8186.

Huggins, M.L. and Sun, K.-H. (1943). Calculations of density and optical constants of a glass from its composition in weight percent. *Journal of the American Ceramic Society* **26** 4–11.

Iliffe, C.J. and Newton, R.G. (1976). Using triangular diagrams to understand the behaviour of Medieval glass. *Verres et Réfractaires* **30** 30–34.

Jackson, C.M. (1992). *A Compositional Analysis of Roman and Early Post-Roman Glass and Glassworking Waste from Selected British Sites.* Unpublished PhD Thesis, Department of Archaeological Sciences, University of Bradford.

Jantzen, C.M. and Plodinec, M.J. (1984). Thermodynamic model of natural, Medieval and nuclear waste glass durability. *Journal of Non-Crystalline Solids* **67** 207–223.

Kingery, W.D. and Vandiver, P.B. (1986). *Ceramic Masterpieces.* Free Press, New York.

Lindsay, W.L. (1979). *Chemical Equilibria in Soils.* Wiley-Interscience, New York.

Lutze, W. (1988). Silicate glasses. In *Radioactive Waste Forms for the Future,* eds. Lutze, W. and Ewing, R.C., North-Holland, Amsterdam, pp. 3–159.

Morey, G.W. (1954). *Properties of Glass.* Reinhold, New York.

Newton, R.G. (1978). Colouring agents used by Medieval glassmakers. *Glass Technology* **19** 59–60.

Newton, R.G. (1982). *The Deterioration and Conservation of Painted Glass: A Critical Bibliography.* Corpus Vitrearum Medii Aevi Great Britain Occasional Paper II, British Academy, Oxford University Press, Oxford.

Newton, R.G. (1985). The durability of glass – a review. *Glass Technology* **26** 21–38.

Newton, R.G. and Davison, S. (1989). *Conservation of Glass.* Butterworth, London.

Paul, A. (1977). Chemical durability of glasses; a thermodynamic approach. *Journal of Materials Science* **12** 2246–2268.

Paul, A. (1990). *Chemistry of Glasses.* Chapman and Hall, London (2nd edn.).

Pfund, A.H. (1946). The aging of glass surfaces. *Journal of the Optical Society of America* **36** 95–99.

Pollard, A.M. (1979). *X-Ray Fluorescence and Surface Studies of Glass, with Application to the Durability of Mediaeval Window Glass.* Unpublished D.Phil. Thesis, Department of Physics, University of York.

Pollard, A.M. (1993). Groundwater modelling in archaeology – the need and the potential. Paper presented at *Science and Site: Evaluation and Conservation,* 8–10th September 1993, Bournemouth University (Proceedings to be published).

Pollard, A.M., Matthew, J.A.D. and Heavens, O.S. (1980). Electron loss spectra of silicate glasses. *Physics and Chemistry of Glasses* **21** 167–170.

Porai-Koshits, E.A. (1977). The structure of glass. *Journal of Non-Crystalline Solids* **25** 87–128.

Porai-Koshits, E.A. (1985). Structure of glass: the struggle of ideas and prospects. *Journal of Non-Crystalline Solids* **73** 79–89.

Porai-Koshits, E.A. (1990). Genesis of concepts on structure of inorganic glasses. *Journal of Non-Crystalline Solids* **123** 1–13.

Rana, M.A. and Douglas, R.W. (1961a). The reaction between glass and water. Part I. Experimental methods and observations. *Physics and Chemistry of Glasses* **2** 179–195.

Rana, M.A. and Douglas, R.W. (1961b). The reaction between glass and water. Part 2. Discussion of the results. *Physics and Chemistry of Glasses* **2** 196–205.

Redwine, R.H. and Field, M.B. (1968). The effect of microstructure on the physical properties of glasses in the sodium silicate system. *Journal of Materials Science* **3** 380–388.

Sanders, D.M., Person, W.B. and Hench, L.L. (1972). New methods for studying glass corrosion kinetics. *Applied Spectroscopy* **26** 530–536.

Schreiner, M. (1988). Deterioration of stained medieval glass by atmospheric attack. Part 1. Scanning microscopic investigations of the weathering phenomena. *Glastechnische Berichte* **61** 197–204.

Sellner, C., Oel, H.J. and Camera, B. (1979). Unterscuchung alter gläser (waldglas) auf zusammenhang von zusammenensetzung, farbe und scmelzatmosphäre mit der elektronenspektroskopie und der elektronenspinresonanz (ESR). *Glastechnische Berichte* **52** 255–264.

Stevels, J.M. (1960–61). New light on the structure of glass. *Philips Technical Review* **22** 300–311.

Thomas, R.G. (1990). *Studies of Archaeological Copper Corrosion Phenomena.* Unpublished PhD thesis, School of Chemistry and Applied Chemistry, University of Wales College of Cardiff.

Vogel, W. (1977). Phase separation in glass. *Journal of Non-Crystalline Solids* **25** 172–214.

Vogel, W. (1994). *Glass Chemistry.* Springer-Verlag, Berlin (2nd edn.).

Volf, M.B. (1961). *Technical Glasses.* Pitman, London.

Vose, R.H. (1980). *Glass.* Collins, London.

Warren, B.E. (1937). *X*-ray determination of the structure of liquids and glass. *Journal of Applied Physics* **8** 645–654.

Warren, B.E. and Biscoe, J. (1938). Fourier analysis of *X*-ray patterns of soda-silica glass. *Journal of the American Ceramic Society* **21** 259–265.

Weyl, W.A. (1976). *Coloured Glasses.* Society of Glass Technology, Sheffield.

Williams, C.T. (1988). Alteration of chemical composition of fossil bones by soil processes and groundwater. In *Trace Elements in Environmental History,* eds. G. Grupe and B. Herrmann, Springer-Verlag, Berlin, pp. 27–40.

Zachariasen, W.H. (1932). The atomic arrangement in glass. *Journal of the American Chemical Society* **54** 3841–3851.

Chapter 6

The Chemical Study of Metals – the European Medieval and Later Brass Industry

INTRODUCTION

With the exception of gold and platinum, the great majority of the metal artefacts used in antiquity were obtained by smelting metalliferous ores to give the various metals and alloys required (Craddock, 1995). One of the earliest goals of archaeological chemistry has been to track these metal objects back to their ore source using trace element analysis, since this may be a direct method of reconstructing ancient trading patterns. In this context, we are using the term 'trace element' to imply a component of the alloy which was not deliberately added, and which may be considered to be characteristic of the ore from which the metal (or one of the metals, in an alloy) was extracted. Such 'trace elements' may well be present in amounts up to a few percent, and the issue of deliberate addition or accidental incorporation is contentious, to say the least (*e.g.*, Pollard *et al.*, 1990). Although the principle of trace element provenancing is very similar to that employed successfully with other archaeological materials [*e.g.*, obsidian (Chapter Three) and pottery (Chapter Four)], when applied to metals it has always been a process viewed with considerable scepticism in some quarters, and now it would be fair to say that the majority opinion of archaeometallurgists is that precise chemical provenancing of metal objects is not in general possible. This is because the high temperatures and extreme redox conditions involved in processing the ores and finished metal mean that some, at least, of the control over the trace element composition of the finished product is exerted by thermodynamic and kinetic considerations during the processing. In fact, it could be argued that the trace element composition of the finished artefact is governed largely by the

chemical equilibria established during these processes, and that this information might therefore be useful in reconstructing the processes used.

In general, of course, it is probably safest to say that the trace element composition of a finished metal artefact is controlled by a range of factors, including the composition (mineralogical and chemical) of the ore source(s) involved, the thermodynamics and kinetics of the processes used, and deliberate human factors such as the mixing of metals to produce the desired alloy. The influence of these factors is discussed in more detail in Chapter 9. It is no great surprise that the huge but relatively simplistic metal analysis programmes of ancient European copper alloy artefacts carried out for the purpose of provenance during the 1950s and 60s simply created confusion (Budd *et al.*, in press). It would, however, be unwise to take the extreme view that chemical analysis of archaeological metal objects is never worthwhile. The study of the metallographic structure of metal artefacts (using various microscopic techniques on prepared surfaces or sections) is extremely well-established, and is absolutely vital for the understanding of manufacturing processes (*e.g.*, Scott, 1991). Significant additional information can nevertheless be obtained about manufacturing processes from chemical analysis, as is demonstrated in the study of European Medieval brasses presented below.

Brass is an alloy of copper with a minor proportion of zinc – typically in archaeological samples up to about 30%, for reasons explained below. It is an extremely interesting alloy, largely because of the volatility of metallic zinc (boiling point 906 °C), which means that zinc cannot easily be reduced to a metal from its common ore sources without distillation. Consequently, metallic zinc is rare (some would say absent) in the archaeological record before the later Medieval period, although brass first appears in Europe in Roman times. In the absence of metallic zinc, the alloy could not have been made by melting together ingots of copper and zinc, and some other process is required. The Romans perfected a technique known as the *calamine process*, described below, which allows brass to be made without recourse to metallic zinc, but which incidentally puts kinetic or thermodynamic controls on the maximum levels of zinc achievable in brass made by this process. Once the distillation process for the manufacture of zinc had been mastered, brass of any composition could be made by a simple mixing of copper and zinc, and thus the upper limit of zinc content imposed by the calamine process became largely obsolete. Thus a measurement of the level of zinc in a brass alloy can be extremely informative about the manufacturing process involved.

One result of the peculiar history of brassmaking is that since the alloy composition has changed in an explicable way over time as a result of changing manufacturing processes, authenticity studies of supposedly ancient brass artefacts becomes a real possibility. The so-called 'Drake Plate' (Hedges, 1979) discussed below is an excellent example of this. Given the high market value of Medieval European brass artefacts, particularly scientific instruments, the ability to give an independent (scientific) opinion on the authenticity of disputed artefacts is extremely important, and one which is now routinely taken advantage of by most large museums.

In this chapter, following a description of what is believed to be the ancient process for the manufacture of brass, we review the early history of brassmaking, and the Medieval and later European industry, largely from literary sources. Following a brief discussion of the problems of the analysis of museum specimens of brass objects, we then summarize some analytical data on European brass tokens and coinage which throws an interesting perspective on the literary evidence. We conclude with an example of the application of this knowledge to the study of European scientific instruments.

THE PRODUCTION METHODS OF BRASS IN ANTIQUITY

The question of how ancient brass was produced has been discussed extensively. As noted above, the two simplest possibilities are the *direct process* – the mixing of metallic copper with metallic zinc, presumably by adding the zinc to molten copper – and the *calamine* (also known as *cementation*) process, which was in use in Europe by the Roman period. Widespread use of the direct process is thought to be unlikely in Europe before the 18th Century AD because of the difficulty in producing metallic zinc. The calamine process involves heating together solid broken metallic copper and zinc oxide or zinc carbonate ($ZnCO_3$) mixed with charcoal in a closed crucible. Zinc carbonate was traditionally known in Britain as *calamine*, although mineralogically it is now correctly referred to as *smithsonite*, since the term *calamine* used to be used for the silicate mineral *hemimorphite* [$Zn_4Si_2O_7(OH)_2.H_2O$] in the USA. At a temperature of between 906 and around 1000 °C, (*i.e.*, hot enough to vaporize the reduced zinc metal, but not hot enough to melt unalloyed copper, which has a melting point of 1083 °C) the zinc vapour is absorbed by the solid copper, and at the end of the process the temperature is raised, the alloy melts and can be poured into an inclined stone mould to form sheet brass. Alternatively, the metal may melt during the process,

when the solidus temperature (the temperature below which no liquid phase exists) of the alloy produced falls to the temperature of the furnace – the binary Cu–Zn phase diagram can be used to predict this 'melting point' of the alloy for a given uptake of zinc (Figure 6.1). It is generally stated (but not necessarily true!) that the nature of this solid copper–gaseous zinc reaction is such that the amount of zinc entering the alloy is limited by the reaction equilibrium set up at the temperature of the furnace, and that a maximum zinc uptake of about 28% is observed because the alloy would melt if more zinc was absorbed. This was apparently more or less confirmed in a series of experiments carried out by Werner (1970) and, more comprehensively, by Haedecke (1973). In fact, Werner put an upper figure of 30% on the zinc uptake, and

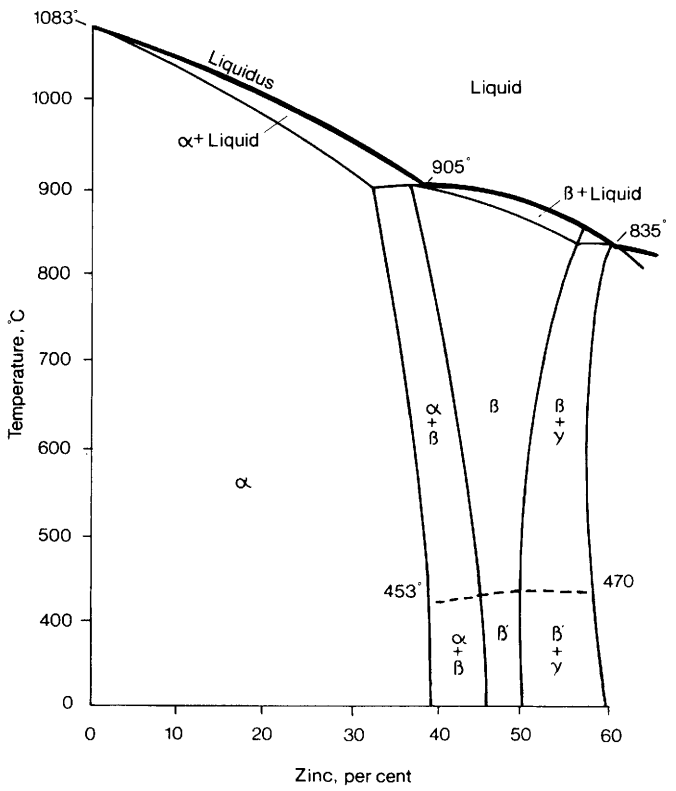


Figure 6.1 *Binary Cu-Zn phase diagram*
(Adapted from Rollason, 1973; Figure 201)

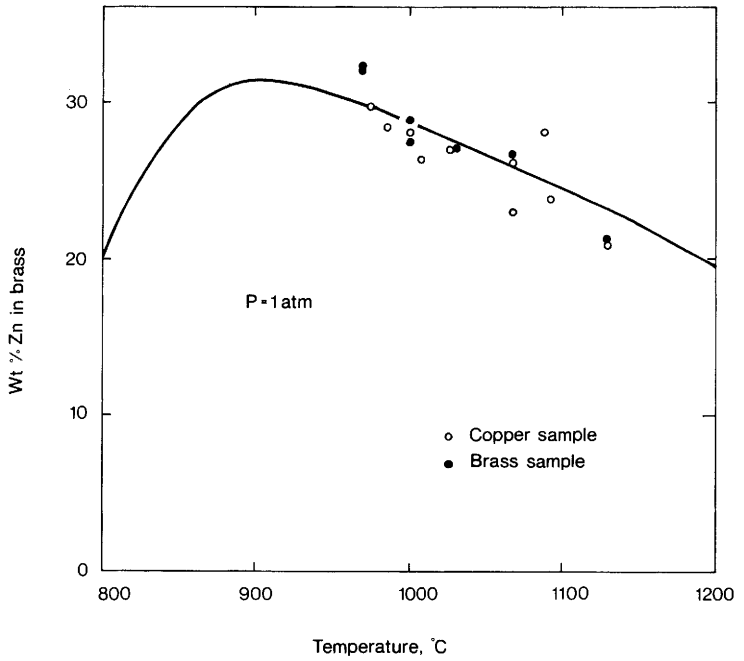


Figure 6.2 *Uptake of zinc during the calamine process as a function of temperature* (Adapted from Haedecke, 1973; Figure 5, with permission)

Haedecke gives a graph of zinc uptake as a function of furnace temperature (his Figure 5, reproduced as Figure 6.2) which shows a theoretical curve together with some supporting experimental data which peaks at a zinc uptake of 31% ($\pm$ 1%) at just over 900 °C. His work included starting with solid samples of both copper and brass, and he concluded that the final zinc level in the brass depended only on the temperature of the furnace, providing the temperature was above 970 °C – below this temperature, kinetic factors limit zinc absorption to a thin ($<$ 1 mm) surface layer. From these data there can be little doubt that the uptake of zinc is limited by thermodynamic or kinetic considerations during the calamine process. However, in the light of the importance subsequently placed on these figures of uptake limitations in the interpretation of the ancient manufacturing techniques (and consequently in authenticity judgements), there appears to be scope for further study of the details and, in particular, the mechanism of the limitations on zinc uptake. There can be no doubt that the existing interpretations are inadequate.

THE EARLY HISTORY OF BRASS AND ZINC

The early history of brass is discussed only briefly here, drawing mainly on the work of Tylecote (1976), Craddock *et al.* (1980), and several chapters in Craddock (1990), all of which summarize the analytical and documentary evidence. The very earliest appearance of zinc in copper alloys cannot necessarily be interpreted as the deliberate production of brass, since the principal primary ore of zinc (ZnS – *sphalerite, blende*, or *black jack)* commonly occurs in association with sulfides of copper and lead, and many sulfidic copper ores contain traces of zinc. It is not unlikely that the smelting of sulfidic copper ore under suitable conditions might give rise to impure copper containing traces of zinc, although the roasting step normally required for the smelting of sulfidic ores might be expected to limit the uptake because of the volatility of zinc. Despite this, both Craddock and Tylecote accept that arsenical copper or tin bronzes with up to 8% Zn can be explained as the unconscious use of a mixed ore, retaining a variable concentration of zinc. Tylecote asserts that the 23.4% Zn contained in a pin from Gezer, an Early Bronze Age city in Palestine, is the earliest example of a deliberate brass, dating to 1400–1200 BC, but Craddock (1978) is more sceptical, noting that no other Palestinian Bronze Age metalwork has yet been published with a similar composition, that the pin in question was without a reference number and described by the excavator as 'a small corroded worthless pin', and that the site was also occupied until Hellenistic times, throwing some possible doubt on the exact date of the pin. Craddock suggests that the earliest occurrence of brass should be placed in the 8th–7th Century BC in Asia Minor, represented by finds of fibulae ('safety pin'-type brooches) from the Gordion Tomb in Phrygia, southwestern Anatolia (alloys with 10% Zn, and little tin or lead). Sporadic examples have been published of later 1st Millennium BC brasses with up to 10% Zn, culminating in an Etruscan (Italian) statuette of a naked youth dated to the 3rd or 2nd Centuries BC with around 11% zinc (Craddock, 1978; Figure 2). Brass coinage containing between 13 and 26% Zn has been reported from ancient Anatolia (Bithynia and Phrygia, in modern day Turkey), securely dated to the 1st Century BC. One example, dating to between 90 and 75 BC from Amisus in Bithynia and Pontus (Northeastern Turkey) was found to contain 19% zinc (with 0.83% Pb, 0.31% Sn, and 0.5% Fe, the remainder Cu) which is thought to represent the earliest known brass coin definitely produced by the cementation process (Preston, 1980). There can be little doubt that brass was being produced, almost certainly by the cementation process, by the end of the first Millennium BC in what is now Turkey, and probably other areas of the Mediterranean.

The early history of metallic zinc has been reviewed in some detail by Craddock (1990). Although a number of artefacts from the classical world have been reported as being made from zinc, the one which has excited most attention and has survived close scrutiny is the small sheet (6.6 × 4 cm) of zinc found in 1939 in the Athenian Agora, stratified with pottery and coins of the 4th–2nd Centuries BC. It was subsequently analysed by emission spectrography, and found to be of zinc, with 1.3% Pb, 0.06% Cd, and traces of Cu, Fe, Mn, Mg, with very faint traces of Sn, Ag, Si, and Cr (thought to be a residue from the cleaning process applied after excavation). Although the levels of lead, iron, and cadmium are similar to modern zinc, the presence of the other traces were not as would be expected in a modern sample, although it was recognized that a relatively short period of burial could have been responsible for their presence. The clinching argument was felt to be the metallographic structure, which showed moderate hammering, rather then the casting or hot rolling which has been common in European zinc manufacture since about 1805 AD.

Craddock (1990) accepts this piece as evidence that metallic zinc, although apparently extremely rare, was known in antiquity, and turns to the classical authors for support. The Greek author Strabo in his Geography (drawing on the now lost 'Philippica' of the 4th Century BC writer Theopompus) writes of a stone from Andeira (in Phrygia, modern day Turkey) which when 'treated in a furnace with certain earth it yields drops of false silver. This, added to copper forms the mixture which some call *oreichalkos*' (Craddock, 1990; 4, from a translation of Strabo by Caley, 1964). 'Oreichalkos' is generally taken as referring to brass, and therefore the 'false silver' is identified as zinc. This reference by Strabo is thought to refer to the occasional condensation of reduced zinc on the walls of furnace flues, which has been observed to happen in later periods, and was referred to in the German mining regions as '*conterfei*'. It would appear that the direct process of zinc manufacture was known to the ancient Greeks, albeit possibly as a great rarity.

Another reference, this time probably to the calamine process, appears in the works referred to as a pseudo-Aristotelian compilation called 'On Marvellous Things Heard', probably written in the 3rd Century BC. As translated by Caley (1967; 67), the relevant passage runs:

> '*the bronze of Mossynoeci is very shiny and light in colour, though tin is not mixed with the copper, but a kind of earth which occurs there is smelted with it. But they say that the discoverer of the mixing process did not instruct anyone else, so that the bronze objects formerly produced there are superior, whereas those made subsequently are not.*'

The Mossynoeci were a people who lived in Asia Minor, on the shores of the Black Sea, again in modern day Turkey.

It seems clear enough from both analytical and documentary sources that the manufacture of brass was discovered some time during the First Millennium BC, probably in modern day Turkey. Under the influence of the Greeks, this knowledge spread throughout the Mediterranean, finally culminating in its use for the reformed coinage issues of Augustus in around 23 BC, which meant that the alloy was widely available throughout the Roman Empire (Craddock *et al.*, 1980). Its popularity is attested by its widespread use as a coinage metal from this time, and ultimately as the alloy of choice for a wide range of items of jewellery (Bayley, 1990) – presumably because of its golden appearance, and the relative ease of manufacture using the calamine process.

Following the decline of classical influence in Europe, knowledge of the manufacture of brass by the calamine process became enmeshed in the Alchemist's search for *aurifaction* – the production of gold from base metals (Pollard, 1988). Some of this lore from Ptolemaic and Roman Alexandria has been preserved in the writings of the likes of Zosimus (3rd Century AD) and Mary the Jewess (2nd Century AD), who, significantly, is traditionally credited with the invention of the retort for distillation (Holmyard, 1957). This information was translated into Arabic after the conquest of Egypt (640 AD) for the use of the Islamic proto-chemists, and was eventually returned to Europe by way of the Moors in Spain. It is very easy to see the relationship between brassmaking – the addition of a special 'earth' to a base metal producing a psuedo-gold – and the 'great work' of alchemy – the isolation of the 'philosopher's stone', which transmutes base metal into gold.

Needham (1974) gives a very full and well-documented account of the early history of brass and zinc in the Far East. Briefly, he concludes that the production of brass from calamine and copper was practised in China by the 3rd or 4th Century BC. The relative agreement of the dates of the first production by this method in both East and West caused Needham to speculate that knowledge of the process diffused from some intermediate place – he suggests Persia, on the grounds that brass was being imported into China from there before 590 AD, and Chinese sources often say that Persia is the source of 'real' brass. Brass was also known in India by 646 AD, but the work of Craddock outlined above so far points to Asia Minor as the most likely source of such knowledge. In their analytical study of copper-based Chinese coinage, Bowman *et al.* (1989) showed the essential conservatism of the Chinese mint, in that, although copper alloys formed the basis of

the official issue, brass was not introduced until sometime between 1503 and 1505 AD, but became the sole alloy in use after 1527 AD. Chinese literary sources, particularly the *T'ien Kung K'ai Wu* (written around 1637 AD) refer to an alloying procedure involving six or seven parts of copper to three or four parts of zinc, allowing for 25% loss of the zinc through vaporization, giving an approximate composition of 30% zinc in the final alloy. Zinc metal was clearly available at this time, and was referred to as *wo chienn* ('poor lead').

The actual isolation of zinc metal on an appreciable scale seems to have occurred first in China in the 10th Century AD (Xu, 1990), using an upwards distillation procedure from secondary (oxidized) zinc minerals. Earlier finds of metallic zinc (such as that at the Agora, noted above) are possibly explained by the chance condensation of small quantities of zinc in the furnace during the production of lead and silver from mixed ores. Much attention has been focused in recent years on northern India, particularly the Zawar region, following the spectacular finds of *in situ* downward distillation zinc furnaces using sulfide ores (Craddock *et al.*, 1990; Hegde, 1989). The dating evidence from this site is somewhat tantalizing: radiocarbon dates from the associated deep mines show that these were in use during the first Millennium BC, but the evidence for the actual use of the downward distillation process dates to some time after the 10th Century AD, with the principal industrial phases at Zawar dating to the 14th Century AD (Craddock *et al.*, 1990). Metallic zinc, or '*spelter*', was exported to Europe from the Far East after 1605 AD, with England importing about 40 tons per year between 1760 and 1780. In 1745, a ship called the *Gotheberg* sank with her cargo of porcelain, tea, silk, and zinc. On recovery and analysis of the metal, around 1870, it was found to be better than 98.99% Zn, with traces of iron and antimony.

If we accept a date of around 1000 AD for the commencement of the distillation of zinc on a large scale, then, following the work of Craddock (1978), all earlier brasses should contain less than 28% Zn, as this is the approximate upper limit for the calamine process at around 1000 °C. Above this temperature, the process is more efficient, but it is said that the brass produced melts and the active surface area for the process is thus reduced. By granulating the copper and therefore increasing the surface area, the maximum can be pushed to around 33% Zn, but it is unlikely that this was done in Europe until the 18th Century (see below). This model is supported by the analytical data: Craddock's work on Roman brass indeed shows an upper limit of about 28% zinc.

THE MEDIEVAL AND LATER EUROPEAN BRASS INDUSTRY

In Medieval Europe, brass continued to be made by the cementation process, particularly in Germany, which was famous for its skill in metalwork. Many of the early alchemical writers mention brass production – Albertus Magnus (1193–1280 AD), Thomas Aquinas (1225–1274), Roger Bacon (1214–1292) [See Grant (1974) and Partington (1961)]. More importantly, cementation is mentioned in the famous technical treatises of Theophilus (*On Divers Arts*; Hawthorne and Smith, 1979), Agricola (*De Re Metallica*; Hoover and Hoover, 1950), and Biringuccio (*Pirotechnia*; Smith and Gnudi, 1959). Theophilus, probably a German monk and metalworker, in his *De Diversis Artibus* written around 1100 AD (dating according to Hawthorne and Smith), gives a very clear and well-known account of a calamine-type process (Book III, chapter 66; 143–144):

'*... when the crucibles are red-hot take some calamine, ... that has been [calcined and] ground up very fine with charcoal, and put it into each of the crucibles until they are about one-sixth full, then fill them up completely with the above mentioned [crude] copper, and cover them with charcoal.... Now, when the copper is completely melted, take a slender, long, bent iron rod with a wooden handle and stir carefully so that the calamine is alloyed with the copper.... Put calamine in them all again as before and fill them with copper and cover them with charcoal.... When it is once more completely melted, stir again very carefully and remove one crucible with tongs and pour out everything into [little] furrows cut in the ground. Then put the crucible back in its place. Immediately take calamine as before and put it in, and on top as much of the copper that you have [just] cast as it can hold. When this is melted as before, stir it and add calamine again and fill it again with the copper you have [just] cast and allow it to melt. Do the same with each crucible. When it is all thoroughly melted and has been stirred for a very long time pour it out... This alloy is called coarse brass and out of it are cast cauldrons, kettles and basins*'

It is noteworthy that the process described here does not involve closed crucibles – it simply appears to rely on the depth of the crucible (see Hawthorne and Smith, 1979, Figure 16 for a reconstruction of the furnace and crucibles used) and the charcoal cover to produce the necessary conditions for the solid–vapour reaction to occur. The above description is clearly written by a man with first hand experience of the process. Elsewhere, however (Book III, chapter 48), he states that there are many sorts of gold, and proceeds to describe 'Spanish Gold', which is an alchemical transmutation of copper treated with 'powder of Basilisk and human blood'. This ambivalent approach to the nature of brass

seems to have vanished by the 16th Century AD, which saw the publication of Biringuccio's *Pirotechnia* (1540) and Agricola's *De Re Metallica* (1556) which give (particularly the latter) a wealth of practical detail which dispels much of the mythology associated with metal production.

Paracelsus, the Swiss physician (1493–1541 AD) appears to be the first European to describe metallic zinc: '*zinc, a bastard of copper, a peculiar metal, but often adulterated by foreign materials. It is of itself fusible, but does not admit to hammering.*' Libavius (1540–1616) obtained some zinc from the East Indies via Holland in 1597. Löhneyss, in 1617, described zinc ('contrefey') as a metal that serves to imitate gold, and reports the distillation of zinc in a lead smelting furnace in Rammelsburg. At an earlier date, Erasmus Ebener showed that zinc could be used to make brass instead of calamine (ca.1550–1560). The documentary evidence from this fascinating period in the history of chemistry and metallurgy is discussed by Partington (1961).

By 1529 in England, Henry VIII was so worried about the supplies of 'strategic metals' for ordnance that he forbade the export of brass, copper, bell metal, and latten. During the reign of Elizabeth I (1558–1603) brass was made in England, but the mining of the rich copper deposits was almost totally neglected (Hamilton, 1967). In 1566 she had persuaded German miners and metallurgists to come to Britain, and a copper mine was opened at Keswick in the Lake District under the auspices of the *Mines Royal Company* (1568), which was, like all other mining concerns of the time, a Crown monopoly. Following the discovery of suitable calamine in the Mendips, the *Mineral and Battery Works* was established, and produced the first English brass at Tintern Abbey wireworks (1568; Day and Tylecote, 1991). The Mineral and Battery Works were conferred with the sole right to mine calamine, and also to produce wire and battery products (largely vessels made from hammered and shaped sheets) using the new method of water-powered machinery, also imported from Germany. However, Tintern brass could not be made malleable enough for the satisfactory production of a battery – possibly due to sulfur being added from the coal used, or the poor quality of the copper supplied by the Mines Royal Company (Day, 1973) – and the works soon ceased production of brass and turned to manufacturing iron wire, for use in the local wool combing industry. The zinc content of this first English brass was recorded as 20% (28 lb in weight added to 1 cwt of copper), but this was soon improved to 24% (35 lb weight added; Day, 1973).

The rights of brass production were subsequently leased by mills at Isleworth and Rotherhithe, although illegal brassworks were also in

operation, including one in the Bristol area, which was soon to become an important centre. Under the Crown monopoly, however, the industry declined, and by 1660 good quality calamine was being exported abroad and finished brass being imported. During the Civil War, brass was scarce, and copper was imported from Sweden because the Keswick mines had closed. By about 1670 foreign competition (including copper from Japan!) had virtually closed the English brassworks, and import controls were imposed to try to protect the flagging industry. The monopoly of the Mines Royal Company was abolished in the Mines Royal Act of 1689 by William and Mary, and many new mines and brassworks came into being – *e.g.*, the English Copper Company in Cornwall (1691) and the Dockwra(y) Copper Company, Esher, Surrey (1692). The export of copper was legalized in 1694 (but not to France!). In 1702 the influential Bristol Brass and Wire Company was established, again using Dutch and German labour. The calamine process was still poorly understood, and the quality of the brass was said to be low, resulting in the continued importation of better quality Dutch brass. The Bristol and Esher companies combined in 1709, and in 1711 they petitioned the House of Commons to give protection from imported brass. A counter-petition was lodged by the metal workers, who feared that the standard of the available brass would fall as a result. The petition was refused, but was resubmitted in 1722.

It appears that the principal obstacle to the production of good quality brass was the purity of the copper available – Day (1973) gives a figure of 92% for the purity of copper around 1700. In the early 18th Century several technical improvements were made to the process of copper refining, such as the replacement of coal by coke in the reverberatory furnaces in use at Bristol in 1710. The resulting improvement to the quality of English brass meant that by 1740 the amount of imported material was drastically reduced without legislation. Hamilton (1967) states that 'high purity copper was achieved by 1778'. English copper was used for coinage from 1714 (copper coinage commenced in England in 1613) – before that, copper for coinage had been imported from Sweden.

It was the 18th Century which saw the first changes in European brass manufacturing technology since Roman times. In 1723 Nehemiah Champion of Bristol obtained a patent (no. 454) for an improved version of the calamine process using granulated copper (prepared in the same way as lead shot) instead of broken pieces of copper. This was said to improve the uptake of zinc from 28.6% to 33.3% (40 lb of copper yielding 60 lb of brass instead of 56 lb). As noted above, zinc metal had been known in the East for at least 700 years by this time, and some had found its way to Europe, where chemists had realized its role in the pro-

duction of brass. A small amount was used in the production of gilding metals (called *pinchbeck, etc.* – see below), which only contain 10–15% Zn. It must have been too expensive, however, to use in the production of true brass, particularly so since calamine was very cheap. At Goslar, in Germany, Caspar Neumann (1683–1737) saw zinc being scraped off the furnace walls after the smelting of lead. It was called 'furnace cadmia', and sold for use in brass manufacture (Dawkins, 1950). Watson (1786) credits Henckel as being the first European to deliberately produce zinc from calamine, quoting from his publications in 1721 and 1737. Dawkins (1950) observes that Marggraf published the process in 1743.

In England, William Champion of Bristol patented the process for the production of zinc in 1738 (no. 564). Watson (1786), however, says that it was Dr Isaac Lawson who first produced the metal whilst working for Champion, and suggests that Lawson had visited China. As pointed out by Day (1990), however, derivation from India is more likely since Champion employed a downward distillation, as used in India, as opposed to the upward distillation traditional in China. Champion went on to produce zinc commercially, but traders to the East Indies lowered their prices in an attempt to force him out of business. He reduced production, and moved to the Warmley works where, according to Watson, he was still producing brass by the calamine process in 1748. His patent was extended to 1750, although this was opposed by metal workers, on the grounds that it would affect production of 'brass, copper, Prince's metal, and Bath metal'. He tried to expand and monopolize the industry but was eventually forced to close as a result of rising competition from the more forward-looking Birmingham brassworks, and went bankrupt in 1769.

John Champion patented in 1758 the process of producing brass and zinc from the common ore of *zincblende* (ZnS) or *black jack* (patent number 726). This ore had previously been considered worthless, but even with this advance it is obvious that metallic zinc was still far too expensive to use in brass production by direct mixing, a process which was patented by James Emerson in 1781 (no. 1297). Watson (1786), however, describes the use of zinc in the production of high quality gilding brasses such as *pinchbeck, tomback,* and *Mannheim gold* (see below).

By the 1780s, English production was dominated by the Birmingham brasshouses, producing mainly cast and stamped objects. Before this time, Birmingham had been a centre of highly skilled craftsmen producing finished goods, but the advent of the canals and the steam engine allowed the production of raw materials to move to the area. In 1786, Watson noted that '*great quantities of good brass are made by most realms in Europe, as well as by the English, but the English brass is more adapted to the*

Birmingham manufactures than any other sort is. The manner of mixing different sorts of brass, so as to make the mixture fit for particular manufactures is not known to foreigners; though this is a circumstance of great importance'. Elsewhere, he records that German calamine brass contains 29% Zn (64 lb copper yielding 90 lb brass), but English brass contains about 34% Zn (45 lb copper giving 60–70 lb brass, but typically 68 lb). In addition, Bohemian calamine was said to contain iron, whilst English calamine contained lead. In a further reference to English brass, Watson states that $1\frac{1}{2}$ tons of brass is made from 1 ton of copper, and that it is better than that of foreign manufacture due to the use of pure calamine and granulated copper. We must assume that this figure of 50% Zn in calamine brass, even allowing for granulation of the copper, is something of an exaggeration, although our understanding of the process is far from complete.

Watson describes the brass produced by Emerson (*i.e.*, by the direct mixing of metals) as 'the purest and finest brass in the world', saying that it is free from iron and therefore good for use in compasses. Despite this praise, Emerson went bankrupt in 1803. By the turn of the century, however, large scale zinc production was underway in Europe, with factories at Carinthia (1799), Ruhberg (1799–1800), Liege (1809), and the Vielle Montagne Company in 1837 (Lones, 1919). In addition, an alteration to English import tariffs in 1839 made it cheaper to import zinc, so the declining price of the metal resulted in the slow phasing out of the calamine process. This happened around 1830 in Cheadle and about 1840 in Bristol. The last brasshouse in Birmingham to produce brass by the calamine method closed in 1866 (Aitken, 1866), whilst Hamilton (1967) records that calamine furnaces were still in operation in South Wales in 1858. According to Lones (1919), spelter (direct) brass had completely replaced calamine brass by 1870, but 'the old process yielded the best qualities of brass.'

In the 19th Century, according to Aitken (1866), the practical brassworker 'could distinguish the difference between calamine brass, and the brass now made by direct mixture, in the peculiar appearance of its polished surface'. He also asserts that calamine brass was superior. Day (1973) records that, in Bristol, the direct method used refined copper and metallic zinc plus scrap brass of high quality ('*shruff*' – offcuts from other processes plus material purchased from outside). It was carefully tested and sorted to enable the correct percentages of copper and zinc to be obtained in the finished brass, allowing for volatilization. This was the responsibility of the 'mixer', who kept records of the various qualities and quantities of scrap brass required to give the desired product when mixed with new metals. As regards the calamine brass, however, Aitken states that '*it is worthy of note, that though various calamines were used – dug from*

Derbyshire, Flintshire, Somersetshire and Yorkshire – from long experience and care, the percentage of zinc in the different qualities of brass produced was secured with as much certainty as it is by the method of direct mixtures now practised.'

In the closing years of the production of calamine brass, the following grades of brass are recorded by Aitken (1866):

BB – made of best copper. '*Latten*' or sheet brass with 33% Zn 'as nearly as possible'.
BC – inferior copper, slightly more zinc.
AM – some copper and calamine – tainted with lead (?).
YY – made of ash metal and other inferior materials (?).

More helpfully, the various qualities of gilding metals in use during the 18th and 19th Centuries are listed by Lones (1919):

Tomback – derived from a Malay word – 86% Cu, 14% Zn + a little tin;
Pinchbeck – named after an 18th Century London clockmaker – 88% Cu, 12% Zn;
Mannheim gold – 80% Cu, 20% Zn;
Leaf gold or *'Dutch metal'* – 84% Cu, 16% Zn.

Two other metals are referred to by various authors, but there is some disagreement about the composition:

Prince's metal – said to have been invented by Prince Rupert in 1680 for casting guns – 73% Cu, 27% Zn (Lones), but 'equal quantities copper and zinc' (Aitken).
Bath metal – 83% Cu, 17% Zn (Hamilton): 55% Cu, 45% Zn (Lones).

By the 19th Century, several grades of brass seem to have been in common use, although there is sometimes confusion between the terms employed (from Day and Tylecote, 1991, with additions):

62% Cu/37% Zn/1% Sn	*Naval Brass*
63% Cu/37% Zn	*Common Brass*
67% Cu/33% Zn	*Stamping Brass*
67% Cu/32% Zn/1% Pb	*Clock* or *engraving metal*
70% Cu/30% Zn	*Cartridge Brass*
75% Cu/25% Zn	*Sheet Brass* (Hamilton), *Rolling Brass, Drawing Brass* ('as free from impurities as possible') (Lones).

In the later 19th Century, the demand for hot working and special brasses resulted in a wide range of named alloys:

Muntz metal (*yellow metal, patent brass*) – 60% Cu, 40% Zn (patented 1832)
Sterro metal – 60% Cu, 38% Zn, 2% Fe
Delta metal – 56–58% Cu, 40–42% Zn, 1–2% Fe + Pb, Mn (1883)
Duranna metal – 65% Cu, 30% Zn, 1.75% Sb
Collins Yellow Sheathing Metal – 56% Cu, 44% Zn

This brief review has concentrated on the development of the production methods of brass in Europe, but it must be remembered that a number of other dates are important in the history of the various manufacturing processes, such as *ca.* 1697 AD for the introduction of the rolling mill to replace the battery process, which, however, continued in use in some places to the end of the 18th Century.

THE CHEMICAL ANALYSIS OF METAL OBJECTS

The analysis of metal objects of the type discussed here invariably poses a number of questions to the chemical analyst. There are two major sources of difficulty:

(i) most Medieval brass objects are inherently valuable – the majority of the objects are either museum samples, or in the collection of an individual, who, understandably, requires that any analysis should be as near 'non-destructive' as possible;
(ii) copper alloy objects are inherently inhomogeneous, particularly if the lead content is more than a fraction of a percent. This can be exacerbated if the object is from an excavated context, since it is well-known that electrochemical processes in water can cause severe loss of zinc (*de-zincification*) from the surface of brass objects (Finnegan *et al.*, 1981; Polunin *et al.*, 1982; Trethewey and Pinwill, 1987).

Many of the analytical techniques described in Chapter 2 can be regarded as largely fulfilling these requirements, but two (atomic absorption spectrometry and X-ray fluorescence) have been used for the majority of the work published to date on European Medieval museum objects made of brass. The experimental technique of atomic absorption has been well-described in the literature (Hughes *et al.*, 1976) and only requires a fine diameter drilling (approximately 1 mm) to be made into the metal, yielding a sample of around 10–20 mg. This is sufficient to

measure around ten elements with a coefficient of variation (one standard deviation) of between 1 and 4% for the major and minor elements. The main drawback with this method is the possibility of erroneous results due to sampling if the metal being analysed is inhomogeneous. This is particularly a problem if the object is cast and contains a large amount of lead (as exemplified by Hughes *et al.*, 1982, but not on Medieval brass), although the problems of electrochemical surface modification due to corrosion are minimized if the first few turns of the drilling are discarded and only bright metal turnings are used.

The damage as a result of the drilling may be unacceptable to museum curators if complete artefacts (or small coins) are being analysed. The analytical method chosen to cause the absolute minimum of damage to museum objects has been energy dispersive X-ray fluorescence, which, with specially designed instrumentation to allow large objects to be positioned in front of the detector, is a rapid and virtually non-destructive means of analysing whole objects without sampling. The main disadvantages are the relatively poor sensitivity to the trace elements, and the necessity to clean an area large enough for the *X*-ray beam (typically a couple of millimetres in diameter) in order to reduce the problems of surface enrichment (in the case of excavated objects) or surface lacquering on museum objects. In the case of thin metal objects (such as coins, or sheet metal plates) it is possible simply to clean the edge and to allow for the reduced area irradiated by the primary beam by normalizing the analytical results to 100%. Providing the methodology is (as should always be the case) validated using 'secondary' standards – standards with certified analyses not used in the primary calibration, which give an independent check on the method – this gives an acceptable compromise between analytical rigour and object preservation. In the analytical work reported here on scientific instruments and brass tokens (carried out in Oxford), the following peak intensities were measured – Fe K_α, Ni K_α, Cu K_α, Zn K_β, Pb L_α + As K_α, As K_β, Pb L_β, Ag K_α, Sn K_α, and Sb K_β. Because of severe peak overlap between Cu K_β and Zn K_α, Pb L_α and As K_α, and Sn K_β and Sb K_α, the Zn K_β, As K_β, and Sb K_β peaks were used to quantify Zn, As, and Sb, which necessarily gives reduced sensitivity to these elements. In the calibration procedure, all the measured intensities were ratioed to the Cu K_α intensity to reduce the problems of long term drift and to minimize geometry and surface condition problems. Each instrument to be analysed was dismantled as much as possible to allow the analysis of all the major components, but small easily-replaceable items such as screws were generally ignored. In almost all cases the instruments were not corroded, and sample preparation was limited to the removal of surface

Table 6.1 *Details of the XRF analyses of brass instruments and jettons*

Minimum detectable levels (wt%):								
Fe	Ni	Cu	Zn	As	Pb	Ag	Sn	Sb
0.05	0.05	–	0.25	0.18	0.05	0.05	0.22	0.10

Coefficients of variation (1σ):

Cu, Zn	1–2%
Sn, Pb	5–10%
Rest	10–20%

tarnish and lacquer with a sharp scalpel or a glass fibre brush. Experiments on a series of brass sheets treated with various chemical and mechanical cleaning agents also showed that this preparation was sufficient to remove any traces of the likely conservation treatments, and also that replicate analyses reproduced the bulk composition on a range of modern brass standards (Pollard, 1983a). Table 6.1 shows the estimated coefficients of variation (1 s.d.) and minimum detectable levels for all the elements measured.

THE CHEMICAL STUDY OF EUROPEAN BRASS TOKENS AND COINS

The Oxford University Research Laboratory for Archaeology and the History of Art has had a long standing interest in the analysis of metals, and also in the authentication of metal objects. An example of this was the work on the 'Drake Plate' (Hedges, 1979), which is a small sheet of brass (20 × 14 × 0.3 cm) found in 1936 in the San Francisco Bay area. It is inscribed with a title claim to the land in the name of Elizabeth I, signed by Francis Drake and dated June 17 1579 (Michel and Asaro, 1979). The plate was found on analysis to have a zinc content of 34.8 (±0.4) %, together with negligible levels of Pb, Sb, and Sn (<0.05%). Comparative analyses were undertaken of 18 brass scientific instruments (from the History of Science Museum, Oxford) and four memorial brasses (from St. John's College Chapel, Oxford) covering the date range 1540 to 1720 AD. These, together with the compilation by Cameron (1974) of data from English memorial brasses between the 12th and 16th Centuries AD showed an average composition for Elizabethan brass of around 20% for Zn with between 0.5 and 1% of both tin and lead. On this basis, combined with the very important fact that the sheet was found to be of a thickness consistent with No. 8 gauge brass of the American Wire Gage standards used in the 1930s, it was concluded that the sheet of metal was unlikely to be of Elizabethan manufacture.

Although this particular piece of forensic work was, for a number of reasons, conclusive in condemning the Plate of Brass, it did reveal some problems with regard to our knowledge of Medieval and later European brassmaking. It is apparent that the argument is somewhat circular, because very few of the objects used to establish the criteria for authenticity have a comprehensively known history, and can therefore be assumed to be completely genuine. This can be a particular problem with scientific instruments or clocks, which can often have parts replaced as they wear out or are lost. It is clear that the compositional trends really need confirmation from some other well-dated brass objects of the same period, to compare with both the instruments and other published analyses. The published analyses of European brass objects of the period are unfortunately of little help in this situation, because they are mainly of cast objects such as candlesticks and statuettes, which contain much more lead than items made from sheet brass (Werner, 1977, 1980, 1982; Brownsword and Pitt, 1983). The trace element data contained in these analyses are, however, of interest and can be compared with those from sheet brass. The most relevant published data are those of Cameron (1974) on English monumental brasses, discussed below. Potentially, suitable objects for study are the small brass tokens known as *jettons*, described in detail by Barnard (1916). They are coin-like tokens of no monetary significance, although they are usually considered as numismatic items. They were used as reckoning or gaming counters, and, being of little value, were unlikely to have been extensively copied or forged. By the same argument, however, they were also unlikely to have been made of the best quality brass, or minted under strict regulation, although Barnard (1916) claims that many of the French tokens were made in the Royal Mints, with a large degree of regulation. Despite this, they can be dated and provenanced, at least to a country of origin, and as such are useful as indicators of the state of brass production at a particular place and time.

A large number of tokens (and some copper coinage) dating from *ca.* 1280–1900 AD was analysed in the Oxford Laboratory during the 1980s, from England (Mitchiner *et al.*, 1985; 1987a), Nuremburg (Mitchiner *et al.*, 1987b), France (Mitchiner and Pollard, 1988), and other European countries (Mitchiner *et al.*, 1988). The total number of jetton and coin analyses carried out under this programme exceeded 600 (including approximately 300 from Nuremburg, 160 from England, and 100 from France). Complete details and analyses are given in the publications cited above, but a brief summary of the major findings which have relevance to the technology of brass manufacture is given here.

A preliminary examination of the entire data shows that the results fall naturally into two main groups – an earlier group (before approximately

1450 AD) characterized by high tin and low nickel, and a later group with low tin and high nickel. The early group has an average composition of 85.7% Cu, 8.0% Zn, 3.7% Sn, and 1.4% Pb, with 0.39% Fe, 0.23 % As, 0.14% Ag, 0.30% Sb, and <0.05% Ni. The (larger) later group, which predominantly consists of Nuremburg jettons, but also contains some later French and Tournai examples, has an average composition of 75.7% Cu, 22.6% Zn, and 0.66% Pb, with 0.25% Ni, 0.17% Fe, <0.2% Sn and As, and <0.1% Ag and Sb. These results are illustrated in Figures 6.3 and 6.4, where the tin and nickel contents are plotted against time. Pre-1450, the majority of jettons have a tin content in excess of 2%, whereas after that date very few have more than 1%. Before 1400, almost all jettons have a nickel content of less than 0.1%, whereas later examples have up to 0.4%. The 'spike' in the nickel concentration between 1600 and 1650 is very characteristic, and is reflected to a lesser extent in similar plots of the arsenic and antimony results, and may reflect the temporary use of some inferior copper ore. Technologically, however, the most important diagram is that of zinc, shown in Figure 6.5, which plots all jettons with more than 20% zinc. This group is dominated by the Nuremburg samples, and so the ensuing remarks may only apply to that production centre. Following the

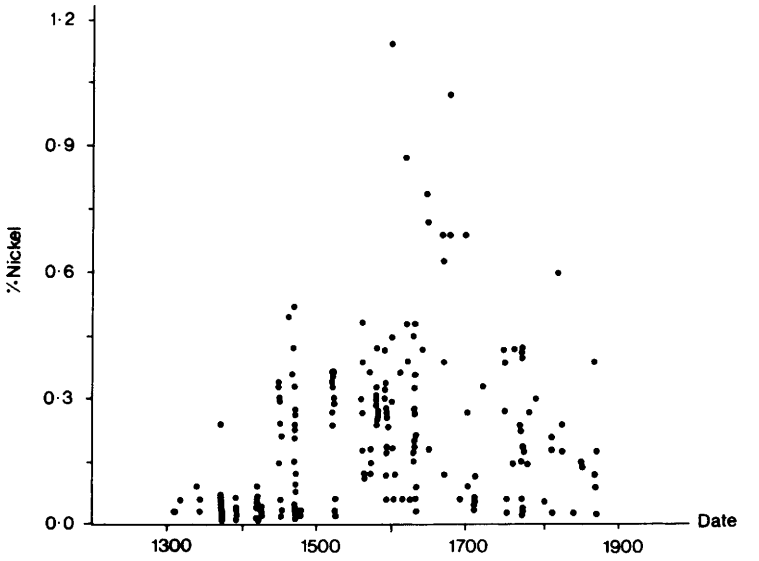


Figure 6.3 *Nickel content of European brass jettons, ca. 1300–1850 AD*

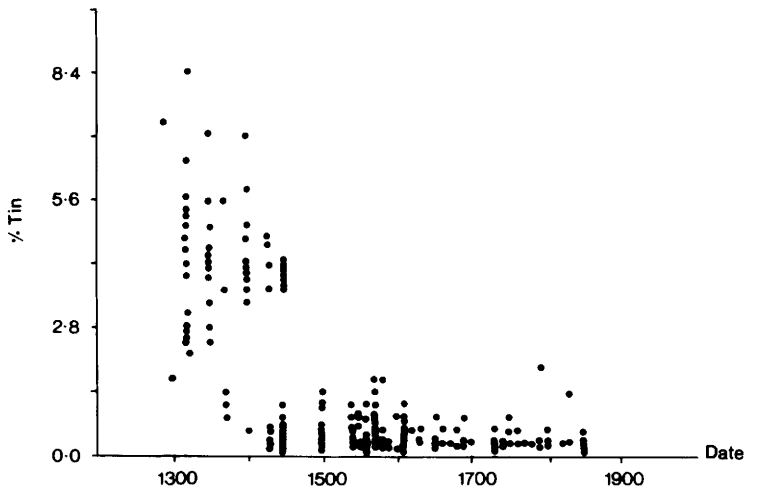


Figure 6.4 *Tin content of European brass jettons,* ca. *1300–1850 AD*

documentary sources and the analytical work of Craddock and others, the two 'critical' concentration levels of 28 and 33% are marked (relating to the maximum uptake by the 'classical' and the 'granulated' calamine processes), showing that the 28% level is reached as early as 1450, and exceeded by 1560 (160 years before the patent awarded to Nehemiah Champion for the granulation process in England). Indeed, a couple of examples as early as 1565 appear to exceed 33%, but a clear step is seen at 1675, when 33% is routinely exceeded, suggesting manufacture by direct mixing. Even taking the later date, this again is 100 years before the patenting of the direct process in England. On the assumption that the technological interpretations are correct (see above), we must assume that the improvements to the calamine process patented in England during the 18th Century were known on the Continent some 100 years earlier. The study of these relatively insignificant little artefacts clearly may have an important influence on our interpretation of European brass manufacturing.

One further point can be made from this overview by considering Figure 6.6, the lead content of the later jettons in the series. From 1450 to 1750, the overwhelming majority of jettons contain less than 1% lead, but around 1760 a small group (8) of high-lead examples emerge (with more than 1.5% lead). The average composition of these tokens is 63.1% Cu, 33.9% Zn, and 2.2% Pb, where the high zinc concentration may suggest manufacture by some form of direct mixing. Possibly the

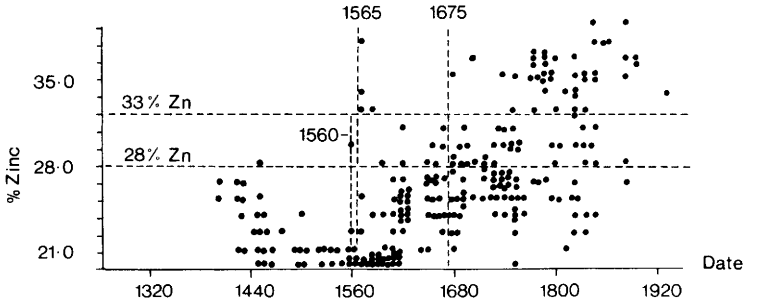


Figure 6.5 *Zinc content of European brass jettons with more than 20% Zn,* ca. *1400–1850 AD*

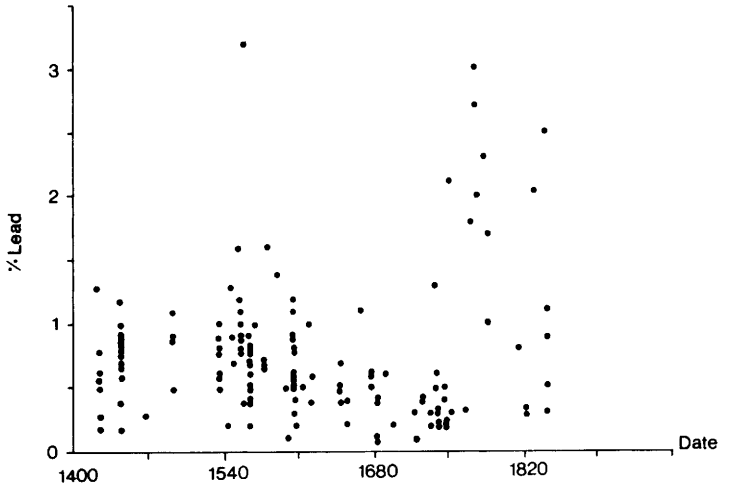


Figure 6.6 *Lead content of later European brass jettons,* ca. *1400–1850 AD*

increased lead is being added along with the zinc, or perhaps the copper being used for this process is less well refined.

The technological information derived from a study of the early jettons can be summarized quite simply – before *ca.* 1450 AD, European jettons contained on average around 4% tin and 1–2% lead, with a low zinc content, typically less than 10%. Traces of antimony, silver, and arsenic are often present at 0.1–0.2%. Nickel is very low, usually less than 0.05%. After 1450, the zinc content rises quickly to over 20%, with a corresponding fall in the lead (less than 1%) and tin (less than 0.2%). The trace elements are usually less than 0.1%, with the exception of nickel,

present up to 0.5%. The combined copper plus zinc total is usually greater than 97%.

There is probably a twofold explanation for this change during the 15th Century. One is a change in the copper ore source supplying much of northwest Europe, from an arsenic – antimony – silver rich ore to one containing more nickel, and the second is a change in manufacturing technique from one in which scrap bronze is normally included in the calamine process (which almost invariably used some scrap in the recipe), to one where only scrap brass or copper was allowed. It is interesting to speculate at this point as to whether this change can be related to the economic shifts occurring in northwest Europe. Before 1400 AD, the Hanseatic League dominated trade along the north coast of Europe, and supplied copper from the Falun mine in Sweden, which opened around 1200 AD (Tylecote, 1976). The League declined during the 15th Century and their position in the European copper trade was taken over by the Fuggers of Augsberg, supplying copper from Hungary and the Tyrol. By the early 16th Century Jakob Fugger virtually monopolized the copper supply, as well as that of lead and silver, and actually obtained a monopoly in mercury supplied from Spain. Political problems in Spain, however, brought about the decline of the Fugger family, resulting in the loss of the Hungarian mines by 1546. There is some additional independent evidence for this scenario, in the form of the analyses published by Werner (1982), which also draws on previous compilations by von Bibra and others. In his Table 11.4.1, Werner gives the analyses of 11 coppers extracted from Austro-Hungarian ores, which shows an average of 0.48% Ni, with only traces of As (0.09%), and traces of Sb. Table 11.10a gives the analyses of 10 Swedish copper artefacts of the 11th–13th Century AD, showing an average Ni content of 0.036%, with 0.16% As and 0.13% Sb. Although far from conclusive, these analyses suggest that the characteristics observed in the jettons do indeed reflect a change in the copper in circulation from sources in Sweden to somewhere in central Europe, some time around 1450 AD.

In view of the documentary evidence for the introduction of new brass-making processes into England in the 18th Century, and the analytical evidence presented above for the earlier invention of these processes in Europe, it is worth re-examining the jetton and coin evidence for the later period in England (Mitchiner *et al.*, 1985, 1987a). A plot of the zinc content of around 50 tokens minted in England between 1600 and 1850 AD shows that the 33% limit is 'breached' almost immediately (Figure 6.7). These tokens (two from Bridgwater, Somerset, dated 1654 and bearing the name William Sealy, with zinc contents of 34.2% and 34.6%, plus two with slightly less zinc – 32.4% in one from Taunton dated 1667

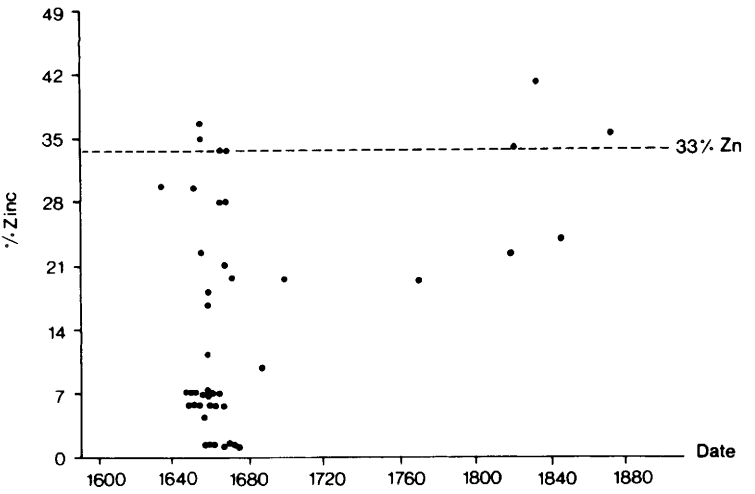


Figure 6.7 *Zinc content of later English jettons,* ca. *1600–1850 AD*

and 31.5% in one from Great Yarmouth, also dated 1667) suggest that high zinc brass, possibly made using some metallic zinc, was available in England by the mid-17th Century. As discussed above, zinc was first produced on a commercial scale in Europe in 1738 and the manufacture of brass by direct mixing (but still using some calamine) was patented in England in 1781. Zinc was, however, available on a small scale some time before the 18th Century – the Dutch East India Company (established 1602) had by the mid-17th Century replaced the Portuguese as the major trading company with the East, and were certainly importing zinc, if not brass. Smith and Gnudi, in their annotated translation of Vannoccio Biringuccio's *Pirotechnia* (1959) suggest that Glauber (in his *De Prosperitate Germanias*, Amsterdam, 1656) was the first European to realize that brass was an alloy of copper and zinc, and that calamine was a zinc ore. Prince Rupert, with whom Glauber was associated, is said to have made a harder metal than that obtainable from the calamine process by adding zinc to the alloy. This has the advantage of being cheaper than making a high-zinc brass entirely from pure copper and zinc. Watson (1786) confirms that cost was a major reason for the continuation of the calamine process, and says that the use of zinc was restricted to the manufacture of low-zinc gilding metals. On the basis of this, we feel reasonably confident in identifying Prince Rupert's metal as a brass with more than 34% zinc, available shortly after 1650, well in advance of the English patent for the full 'direct' process.

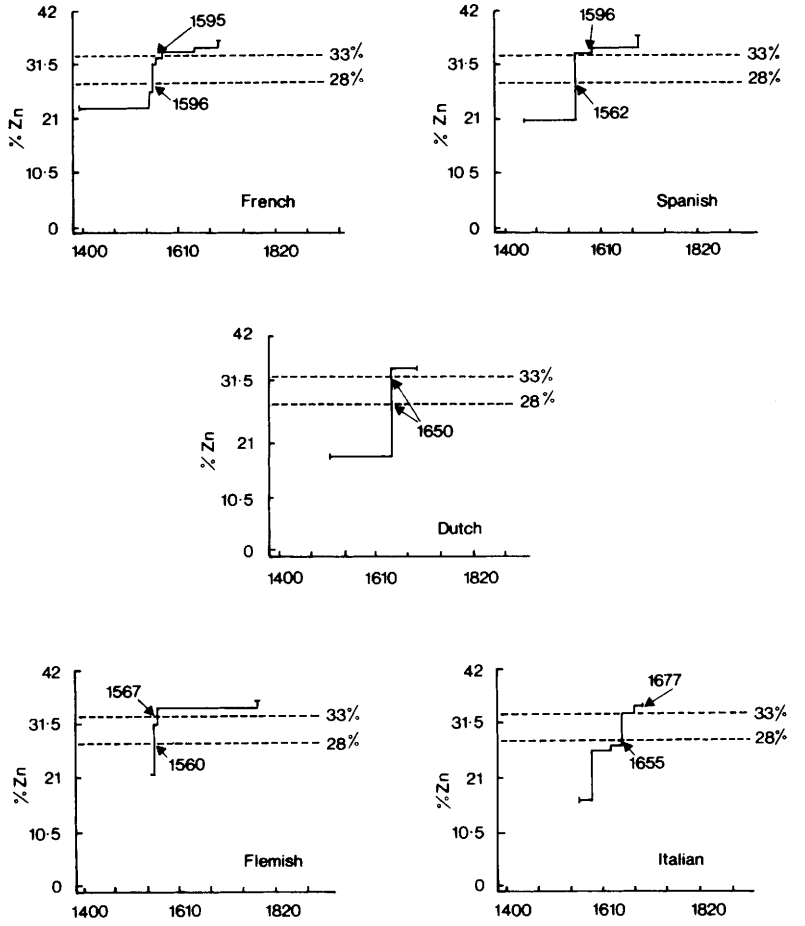


Figure 6.8 *Summary of the maximum zinc content of European brass scientific instruments (ca. 1400–1800 AD), classified by country of origin (also see opposite page)* (Redrawn from Pollard (1983b), and Mortimer (1989; Figure 1) by permission of Elsevier Science and the author)

THE ANALYSIS OF EUROPEAN BRASS SCIENTIFIC INSTRUMENTS

Having accumulated a great deal of knowledge, both documentary and analytical, about the later history of European brassmaking, it seems reasonable to turn cautiously from brass jettons to other objects also

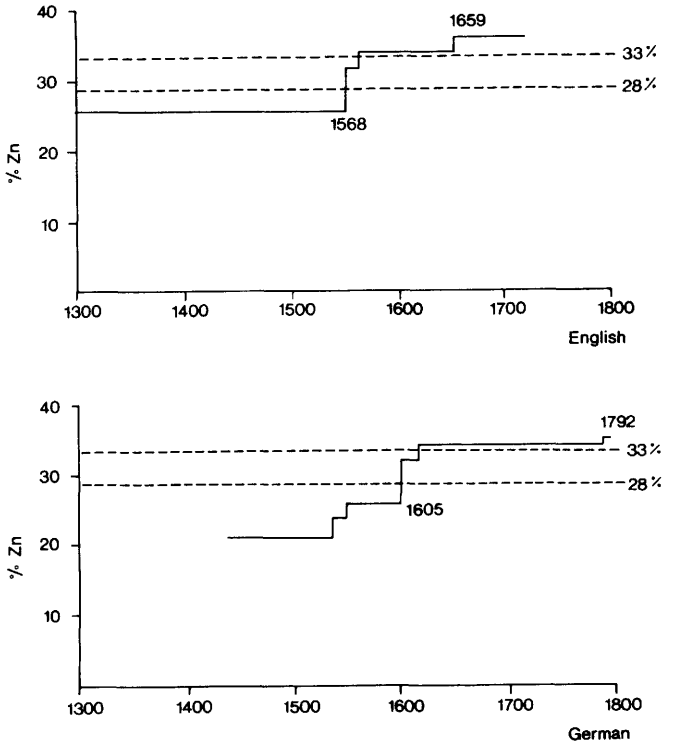


Figure 6.8 *continued*

made of brass but with a much higher monetary value, such as scientific instruments and clocks, the authenticity of which is a constant cause for concern. A number of problems may be anticipated, however, mostly related to the nature of the analytical evidence from the jettons – how representative was this metal of that available to the instrument or clock makers, and how geographically representative is it? It is clear that a large programme of analysis of 'authentic' scientific instruments is still required to corroborate the combined literary and analytical evidence for brassmaking practices in Europe from 1400 onwards. It is also appreciated that chemical information can only contribute to the overall study of these instruments, and is unlikely to be conclusive on its own. Chemical analysis of metals can, after all, only give evidence of compatibility with known objects of the same age. With the co-operation of the History of Science Museum, Oxford, therefore, the Research Laboratory for Archaeology began a programme of analysis of dated

instruments, which resulted in the initial accumulation of 285 individual analyses from 65 instruments, ranging in date from *ca.* 1400–1770 AD. By country, there were fifteen instruments each from Italy, France, and Germany, ten from Holland and five each from Spain and Flanders. A summary of these results was presented at a United Kingdom Institute of Conservation meeting in 1983 (Pollard, 1983b), which was only published in abstract form, and so a summary of the main findings is reproduced here. A further 69 English and German instruments were subsequently analysed, but again were only published in summary form (Mortimer, 1989). It is hoped that the complete analyses can be published in the future.

The main technical interest as before is the maximum zinc content observed in the instruments, which is summarized in Figure 6.8, showing only the maximum zinc content as a function of date, classified by country of origin. For the purposes of this summary, all the analytical results have been accepted as valid, with the exception of one Dutch analysis with 42% zinc, which was rejected as modern, and the attributions have been assumed to be correct. Four countries apparently started using brass with more than 28% zinc at around the same time (*ca.* 1560

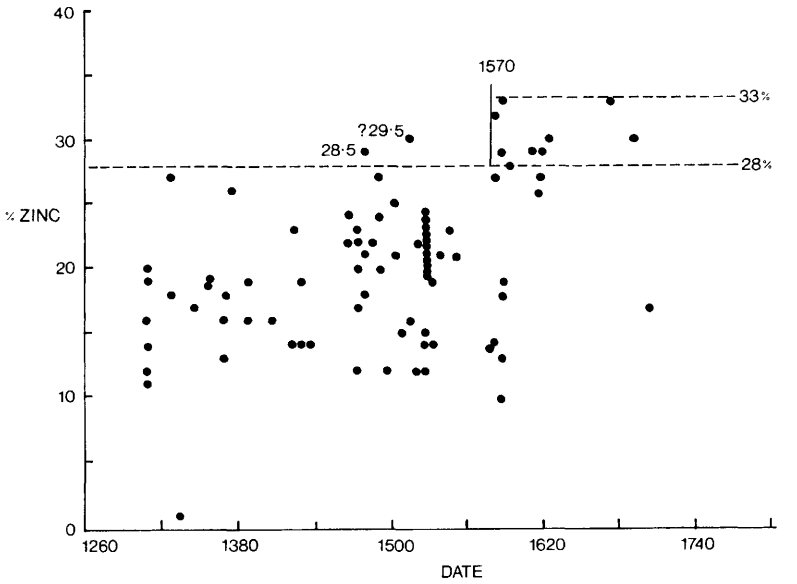


Figure 6.9 *Zinc content of English monumental brasses,* ca. *1250–1700 AD* (Redrawn from Cameron, 1974; Figure 3, by permission of the Royal Archaeological Institute)

AD) when the zinc content of English, French, Spanish, and Flemish instruments jumps simultaneously from below 24% to about 33%. Politically, Spain and the Low Countries (the Spanish Netherlands) were united under Charles V and later Philip II, although rebellion resulted in the establishment of the Protestant Dutch Republic around 1600. The wars with France ended with a treaty in 1559, so we may surmise that after 1560 conditions for free trade in raw materials or finished metals were relatively stable. Nuremburg, on the other hand, was by this time part of the Protestant states of Germany, and was certainly rising in importance as a brass producer, and therefore may not have imported any of this new alloy.

An early date of around 1560 is therefore claimed for the introduction of a process capable of producing brass with more than 28% zinc (presumably a modified calamine process using granulated copper) into the Catholic countries of Spain, France, and Flanders, and also in the brass available in England. The data on the jettons offer good support for this, coming up with a similar date of 1560 for the introduction of this metal. Further support can be found for this model in the work of Cameron on English monumental brasses (1974). Figure 6.9 shows a plot of zinc content against date taken from Cameron's results. Neglecting one point, which is a suspect 19th Century analysis, the first brasses with more than 28.5% zinc are found from 1570 onwards. The early trade between England and Flanders in finished monumental brasses is well-known (Page-Phillips, 1972; Chapter 4), thought to be due to the Hanseatic League. After the glut of re-usable material supplied by the Reformation had dried up, Page-Phillips suggests that tomb makers were again turning to the continent for supplies of sheet brass, some time after 1560. Table 6.2 lists the average composition of post-1560 scientific instrument components with more than 28% zinc from Spain, France, and Flanders, plus the relevant results from Cameron mentioned above. There is clearly a high degree of similarity between all four sets of figures, and it is not, therefore, unreasonable to suggest a common source for this brass – presumably Flanders, in view of the dominant position of Flemish metalworkers at this time.

Table 6.3 (from Pollard, 1983b) lists the average composition of low zinc brasses ($<28\%$ Zn) used for instruments, classified again by country of manufacture. The majority are pre-1560, but some later instruments have been included where they appear to be made from the same type of metal. Also included is the average for Cameron's analyses of early monumental brasses. There is some similarity between the French, Spanish, and Dutch instruments, particularly in the low nickel concentration, suggesting trade of raw materials or finished products through

Table 6.2 *Average analyses of post 1560 AD brass scientific instruments from France, Spain, and Flanders with more than 28% zinc*

		Fe	*Ni*	*Cu*	*Zn*	*As*	*Pb*	*Ag*	*Sn*	*Sb*
Flemish Instruments	$\bar{x}$	0.24	0.21	68.0	30.0	<0.2	0.85	<0.1	0.41	<0.1
1560–1770	s	0.06	0.16	3.0	2.9		0.15		0.10	
$n = 16$										
Spanish Instruments	$\bar{x}$	0.27	0.25	65.1	32.2	<0.2	1.43	<0.1	0.40	<0.1
1562–1596	s	0.27	0.07	2.3	2.3		0.87		0.24	
$n = 5$										
French Instruments	$\bar{x}$	0.22	0.20	66.8	31.2	<0.2	0.86	<0.1	0.30	<0.1
1560–1652	s	0.12	0.10	3.2	3.5		0.41		0.41	
$n = 33$										
English Monumental Brasses	$\bar{x}$	(0.4)	–	68.8	29.8	–	1.0	–	0.2	–
(Cameron) 1571–1678	s			2.4	2.1		0.5		0.2	
$n = 11$										

(– = not reported)

(n = number of analyses)

Table 6.3 *Average analyses of low zinc brasses (< 28% Zn) classified by country. The majority date from before 1560 AD*

		Fe	*Ni*	*Cu*	*Zn*	*As*	*Pb*	*Ag*	*Sn*	*Sb*
French Instruments	$\bar{x}$	0.25	<0.05	78.7	18.6	<0.2	0.66	0.12	1.37	<0.1
1400–1700	s	0.08		2.5	2.5		0.32	0.05	0.46	
$n = 46$										
German Instruments	$\bar{x}$	0.16	0.45	79.8	18.1	<0.2	0.72	<0.1	0.39	<0.1
1490–1677	s	0.08	0.33	3.1	3.0		0.68		0.19	
$n = 20$										
Spanish Instruments	$\bar{x}$	0.16	<0.05	76.5	20.3	<0.2	0.62	<0.1	1.81	<0.1
ca. 1450–1598	s	0.08		3.4	3.9		0.32		1.10	
$n = 10$										
Dutch Instruments	$\bar{x}$	0.15	<0.05	79.3	18.2	<0.2	0.23	<0.1	1.96	<0.1
1516										
$n = 2$										
Italian Instruments	$\bar{x}$	0.13	0.46	82.7	15.5	0.45	<0.1	0.08	0.54	<0.1
1580–1636	s	0.02	0.11	1.5	1.5	0.20		0.04	0.07	
$n = 9$										
Italian Instruments	$\bar{x}$	0.08	<0.05	71.3	25.6	<0.2	2.60	<0.1	<0.2	<0.1
1588–1694	s	0.04		1.8	1.6		0.50			
$n = 14$										
English Monumental	$\bar{x}$	–	–	75.9	18.9	–	3.2	–	1.4	–
Brasses (Cameron) 1300–1691	s			4.5	4.5		2.4		1.4	
$n = 70$										

(– = not reported) (n = number of analyses)

the Hanseatic League, as discussed for the jettons. The German instruments are quite different, having higher nickel and lower tin. The early Italian instruments are interesting, forming two distinct compositional groups, one with high nickel, but low zinc and lead, and the other with low nickel and tin, but high zinc and lead. It may be that these two represent northern and southern Italy – the Kingdom of Naples being a Spanish stronghold, the north having economic ties with Germany, but a more detailed analysis of the data is required to justify this suggestion.

The first appearance of high zinc brasses in scientific instruments – more than 33% zinc – is slightly more difficult to ascertain. Table 6.4 lists the thirteen instruments found during the first programme of analyses which have components made from a metal containing more than 34% zinc. The figure of 34% was chosen to allow a small margin for error, either in the analytical results or the accuracy of the boundary figure, but where an instrument consisted of one part with more than 34% zinc, parts with more than 33% were included in the average. The fraction in brackets after the accession number represents the number of results included in the average compared with the total number of parts analysed on that particular instrument. The three earliest examples (dated 1567, 1595, and 1596) are all slightly different from the later examples, in that they all have a relatively high nickel and tin content (greater than 0.1 and 0.2% respectively). A summary of the difference between these pre-1650 analyses and those post-dating 1650 is given at the bottom of Table 6.4, together with an average of the compositions of the high zinc jettons discussed above for comparison. A closer examination of these early instruments (together with those tokens in Figure 6.5, p. 217, dating to before 1675 with more than 34% zinc) is called for, since these may represent either the very maximum achievable by the calamine process, or they may be the first evidence for the direct use of zinc in Europe, dating from around 1570 onwards. On present evidence, it is tempting to suggest that there is a close correlation between the date attributed by Smith and Gnudi (1959) to the work of Glauber and the first appearance of the low nickel, low tin, high zinc brass instruments, and that 1650 is therefore the most likely date for the first widespread production of high zinc brasses (*i.e.*, made with the deliberate addition of some metallic zinc) in Europe. It is also likely that this alloy was known as Prince Rupert's metal.

THE ANALYTICAL AUTHENTICATION OF BRASS INSTRUMENTS

As an example of the use of this type of information for helping in the study of scientific instruments, we present two previously unpublished

Table 6.4 *Analyses of scientific instruments with more than 34% zinc*

		Fe	*Ni*	*Cu*	*Zn*	*As*	*Pb*	*Ag*	*Sn*	*Sb*
Astronomical ring. Flemish, 1567 57–84/24 (2/3)		0.16	0.29	63.5	34.1	<0.2	1.37	<0.1	0.33	<0.1
Astrolabe. French, 1595 IC 211 (5/6)		0.25	0.24	64.4	33.7	<0.2	0.98	<0.1	0.22	<0.1
Equinoctial Dial. Spanish, 1596 S7 (2/2)		0.14	0.21	63.9	34.1	<0.2	0.81	<0.1	0.55	<0.1
Perpetual Calendar. Dutch, 1650 2497 (3/5)		0.13	<0.05	64.0	34.2	<0.2	1.25	<0.1	<0.2	<0.1
Circumferentor. French, 1652 57–84/253 (3/4)		0.23	0.12	63.9	34.6	<0.2	0.82	<0.1	<0.2	<0.1
Quadrant. Italian, 1677 – (1/3)		0.27	<0.05	63.7	34.7	<0.2	0.94	0.10	<0.2	<0.1
Equinoctial Dial. Spanish, 1695 S10 (4/5)		0.11	<0.05	63.1	35.5	<0.2	0.96	<0.1	<0.2	0.11
Scaphe. Italian, 1697 I.98 (1/3)		0.16	<0.05	64.6	34.2	<0.2	0.45	<0.1	0.24	<0.1
Equinoctial Dial. French, 1700 F.208 (2/5)		0.16	0.21	63.1	35.7	<0.2	0.52	<0.1	<0.2	<0.1
Equinoctial Dial. Dutch, 1700 G.126 (2/6)		0.18	0.34	63.1	35.4	<0.2	0.59	<0.1	<0.2	<0.1
Dialling Instrument. Dutch, 1705 28 (2/3)		0.14	<0.05	63.4	35.0	<0.2	1.15	<0.1	<0.2	<0.1
Astronomical Compendium. Dutch, 1709 M3 & M29 (7/14)		0.15	<0.05	64.2	34.3	<0.2	1.03	<0.1	<0.2	0.10
Solar & Lunar Dial. Flemish, 1770 M18 (8/13)		0.14	<0.05	63.6	34.6	<0.2	1.29	<0.1	<0.2	<0.1
Pre-1650 average	$\bar{x}$	0.18	0.25	63.9	34.0	<0.2	1.05	<0.1	0.37	<0.1
$n = 3$	s	0.06	0.04	0.5	0.2		0.29		0.17	
Post-1650 average	$\bar{x}$	0.17	<0.1	63.7	34.8	<0.2	0.90	<0.1	<0.2	<0.1
$n = 10$	s	0.05		0.5	0.6		0.30			
High zinc jettons	$\bar{x}$	<0.1	<0.1	62.0	36.3	<0.2	1.11	<0.1	<0.2	<0.1
$n = 35$	s			3.3	3.2		0.82			

case studies. One relates to some instruments from the Barberini Collection, now in the National Maritime Museum, London, and one relates to two 18th Century English clocks by famous makers (Tompion and Graham) from a private collection.

The Barberini Collection comprises 24 various scientific instruments, supposedly put together by the Barberini family in Florence – particularly Maffeo Barberino (1568–1644), later Pope Urban VIII, and more particularly Cardinal Francesco Barberino (1597–1679). Unfortunately, only one piece of this magnificent collection can be definitely traced back to Francesco Barberino – a large concave burning mirror. The first complete list of the collection dates only to the First World War, and it was purchased by Michel and Landau of Paris, in the 1930s, and some of it was exhibited at the Descartes Exhibition at the Bibliothèque Nationale in Paris in 1936. The complete collection was finally purchased by the National Maritime Museum in 1949 (Shaw, 1973).

Only part of the collection was analysed at the Research Laboratory for Archaeology in Oxford – five examples were chosen, as follows (descriptions from Shaw, 1973):

6. A.47/46–230C (860). Large Astrolabe, probably Italian. 41.5 cm diameter, in original case. Unsigned and undated, ca. 1630.
8. Cl/S.10/47–216C (883). Sector. Brass, 19 cm length. Signed *I Galli*^{S} *fecit.* Undated, ca. 1600.
14. D.359/47–223C (867). Bowl sundial, 32 cm diam., 50 cm high. Signed on back of bowl *I. Galli.* Undated, ca. 1600.
16. T.63/47–210C (870). Trigomètre. Barberini cypher on back. Brass, 37 cm length folded. Tooled morocco case embossed with Barberini cypher. Signed *Philipus Danfrie: F : Anno Domini 1580.*
17. T.64/47–211C (871). Companion scale to trigomètre, with sights. Barberini cypher on back. Brass, 33 cm length folded. Tooled morocco case embossed with Barberini cypher similar to above.

A summary of the analyses of these five instruments is given in Table 6.5. The figure in brackets gives the total number of separate parts of the instrument analysed (the denominator), and the number of analyses included in the average figure (the numerator).

The large astrolabe (number 6) was dismantled into sixteen parts, and was found to be made from broadly similar (but definitely not identical) metal. The average zinc content is around 13%, but the range on the sixteen parts analysed is from 9.0 to 20.9%; hence the relatively large standard deviation. Similarly, the lead levels vary from 0.2 to 4.6%, and six of the components analysed had measurable amounts of As

Table 6.5. *Average analyses of five instruments in the Barberini Collection*

Description			*Fe*	*Ni*	*Cu*	*Zn*	*As*	*Pb*	*Ag*	*Sn*	*Sb*
6. Astrolabe	(16/16)	$\bar{x}$	0.34	0.17	82.0	13.2	<0.2	1.8	0.12	2.1	0.23
		s	0.13	0.06	2.7	3.4		1.2	0.06	0.8	0.20
8. Sector	(3/3)	$\bar{x}$	0.23	0.10	75.1	21.9	<0.2	0.97	0.09	1.53	0.17
		s	0.02	0.02	2.1	2.3		0.12	0.02	0.12	0.08
14 . Bowl sundial	(7/8)	$\bar{x}$	0.35	0.27	77.7	17.8	<0.2	2.5	<0.1	1.0	0.21
		s	0.11	0.13	2.6	3.5		1.0		0.6	0.18
(bowl)	(1/8)	$\bar{x}$	0.24	1.54	69.6	20.5	<0.2	2.9	0.06	4.4	0.58
16. Trigomètre	(8/15)	$\bar{x}$	0.23	0.16	77.4	19.2	<0.2	0.97	<0.1	1.88	<0.1
		s	0.05	0.05	2.6	2.8		0.14		0.23	
	(7/15)	$\bar{x}$	0.32	0.24	65.6	31.9	<0.2	1.43	<0.1	0.33	<0.1
		s	0.12	0.06	1.1	1.0		0.35		0.24	
(case hook)			0.07	<0.05	55.4	43.3	<0.2	1.1	0.07	<0.2	<0.1
17. Scale	(5/10)	$\bar{x}$	0.41	0.23	74.5	22.6	<0.2	0.98	<0.1	0.9	<0.1
		s	0.04	0.06	2.4	2.9		0.13		0.6	
	(5/10)	$\bar{x}$	0.49	0.23	66.7	30.9	<0.2	1.04	<0.1	0.4	<0.1
		s	0.11	0.06	1.2	0.7		0.15		0.4	
(case hooks)	(2/2)	$\bar{x}$	0.18	<0.05	68.7	29.6	<0.2	0.4	<0.1	1.08	<0.1
		s	0.01		0.5	0.1		0.2		0.16	

(0.2–0.3%). There is clearly some evidence that the instrument was not made from uniform material – whether this is significant in the context of the authenticity of an instrument is difficult to say. Simply taking the average analysis and comparing it with Italian instruments of a similar period (Table 6.3) shows that the analysis of the Barberini piece differs in detail from either of the two Italian groups identified ('high' and 'low nickel'), but that the composition is not itself unreasonable for the proposed date. The sector signed by Galli (number 8) is relatively homogeneous, in that the three pieces analysed are compositionally similar. It does not match any of the other instrument groups exactly, but again its composition is not unlikely for the period. The same comments apply to the majority of pieces (seven out of eight) of the bowl sundial (number 14), except to note that the most important part of the sundial – the bowl itself – has a relatively unusual composition, with 1.54% nickel, which is as high as anything analysed during the whole programme.

The trigomètre and its companion scale (items 16 and 17) are both similar, in that they appear to be made from two qualities of brass – one approximately 75% Cu, 20% Zn with 1–2% each of Pb and Sn, and one about 65% Cu, 30% Zn, with 0.5–1.5% Pb and Sn. The lower zinc composition is similar to (but not identical with) contemporary instruments from parts of Europe, but the higher zinc composition, on the jetton evidence, would be unusual (but not impossible) for the period after 1560. It must be noted that the summary in Table 6.5 gives equal weight to all components of the instruments, ranging from major components such as the arms of the trigomètre down to smaller items such as sites and mount plates. Only a detailed study of the analyses in relation to each component can be expected to give a true picture of the history of the instrument.

In summary, the analyses of these five instruments was essentially inconclusive. All of the compositional groups obtained were plausible for the period claimed (therefore making it impossible to declare any object a 'fake'), but the inconsistency of the analyses when compared with the larger data set of other scientific instruments of the period leaves open the possibility that some at least of the instruments are not entirely genuine – particularly the trigomètre and its companion scale, which show two very distinct compositional groups, one of which is unusual for the late 16th Century. As a corollary to this study, two samples from the wooden cases from items 16 and 17 were subjected to radiocarbon dating by accelerator mass spectrometry in Oxford. The results were 275 ± 75 BP (OxA-536) and 310 ± 80 BP (OxA-537) respectively. Unfortunately, when these are calibrated, although the central dates are

1640 AD and 1630 AD, the 95% confidence intervals on these dates extend right up to 1950 AD in each case, because of the flatness of the calibration curve from 1650 onwards. Although strictly inconclusive, it is possible to wring a little more information from the dates, by noting that there is an 84% probability that the true age lies before one standard deviation above the mean. On this basis, there is an 84% chance that the date of the first sample is before 1800 AD, and the second before 1660 AD (Gowlett, pers. comm.). Albeit in a somewhat unsatisfactory manner, it is possible to suggest on this evidence that the cases at least are likely to be genuine.

A more convincing case study is provided by the chemical analyses of the brass components of two privately-owned 18th Century English clocks by famous makers – Thomas Tompion (1639–1713, working in Fleet Street, London) and George Graham (1673–1751, also working in London). As can be imagined, the analysis of the moving parts of clocks of this quality (and value) places the highest constraints on the analytical procedures in terms of minimizing damage – even the removal of very fine drillings would have been unacceptable, since it would have destroyed the balance of the parts. At the time when these analyses were carried out (1983), surface analyses (with minimal cleaning using a solvent to remove lacquer) by XRF provided the only acceptable means of analysis, although now one might consider laser ablation ICP-MS as a suitable alternative (see Chapter 2). The XRF analyses are summarized in Table 6.6. The Tompion (dated 1709) was dismantled into thirteen pieces, and proved to have a remarkably uniform analysis, which is entirely consistent with English scientific instruments of a similar date (Mortimer, personal communication). The Graham, dated to 1722, was analysed in fifteen separate components, and these were less homogeneous. The larger proportion of the components (nine out of fifteen) had an analysis very similar to that of the Tompion, and therefore also consistent with an early 18th Century date. A further three elements had a slightly higher lead content (2.73%) than the previous set, but were not too dissimilar from one component of the Tompion, which had a lead content of 2.24%. Two elements of the Graham (the escapement and the pendulum) had a much higher zinc content (32%) – not impossible for the early 18th Century, but more likely to be later. One part (a 49 mm diameter solid wheel) was very similar to the bulk of the clock, but had a significantly higher tin content (3.12%). The conclusion drawn from this work was that the Tompion was entirely consistent with having been made in 1709, and all the parts analysed were likely to be original. The Graham was less internally homogeneous, but was also consistent with a date of 1722, with the possibility that the escapement and pendulum had been replaced at a later date.

Table 6.6 *Summary of clock analyses*

Description			*Fe*	*Ni*	*Cu*	*Zn*	*As*	*Pb*	*Ag*	*Sn*	*Sb*
Tompion, 1709	13/13	$\bar{x}$	0.26	<0.05	74.1	22.9	<0.2	1.5	<0.1	0.97	<0.1
		s	0.10		1.2	1.3		0.4		0.20	
Graham, 1722	9/15	$\bar{x}$	0.31	<0.05	72.4	25.1	<0.2	1.3	<0.1	0.62	<0.1
		s	0.06		1.0	1.3		0.3		0.15	
	3/15	$\bar{x}$	0.41	<0.05	74.2	21.4	0.2	2.73	~0.1	0.70	0.23
		s	0.06		0.3	0.3	0.2	0.09		0.07	0.03
	2/15	$\bar{x}$	0.16	<0.05	65.9	32.1	<0.2	1.6	<0.1	<0.2	<0.1
		s	0.04		0.4	1.1		0.5			
	1/15		0.29	<0.05	72.5	21.6	<0.2	2.30	<0.1	3.12	<0.1

SUMMARY

On the basis of a preliminary appraisal of the data, several interesting points have emerged. From both the instrument and the jetton data, evidence has been obtained for two variations on the production of brass in later Medieval Europe via the calamine process – one, largely dating to before 1450 AD, yielding a brass with appreciable levels of tin, which virtually disappears after this date. This may be the result of using scrap bronze in the melt (perhaps because of the high value placed on raw copper), which was discontinued in the later period (or perhaps only scrap copper and brass were subsequently allowed). The latter process results in a pure brass, with the combined copper and zinc total exceeding 97%. The data on the jettons, showing an increase in the nickel content of the later tokens, suggest that this might be roughly coincident with a change in ore source in addition to the use of an improved manufacturing technique. The early period probably represents the phase of domination of the copper trade, and also that in finished metal, by the Hanseatic League, supplying copper from Sweden. The decline of the Hanseatic League resulted in the rise of the Fuggers of Augsberg, trading in central European copper, some of which has characteristically higher nickel. It is worth noting, however, that a change in trace element concentrations does not automatically mean a change in ore source – a change in foundry practice may well also affect the trace element levels (Pollard *et al.*, 1991). Nevertheless, these observations are in harmony with our knowledge of the economic history of Europe at the time.

Contemporary literary evidence suggests that two dates should mark improvements in the manufacturing process of English brass. One is the early 18th Century, when the use of granulated copper is said to have increased the zinc uptake of the calamine process from 28 to 33%, and the second is the introduction of direct mixing of metals as a way of making brass in the late 18th Century. Interpretation of the analytical data on these points is always difficult, because apparent anachronisms can be interpreted either as earlier introductions of these methods, or as evidence of the objects themselves being later copies. There is now a reasonable amount of evidence, however, to suggest that brass compatible with the granulated calamine process was being used in Europe by around 1560. The published analyses of Cameron (1974) also give figures in this category dated to *ca.* 1570, supporting the view that granulated copper was used in the calamine process well before the early 18th Century, if the technological interpretation of the reason for this increase in zinc content is correct.

The date of the earliest appearance of brass made by direct mixing of

metals (not including gilding metals, with low zinc contents) is rather more difficult to ascertain. From Figure 6.5, it could be as early as around 1565, but all the evidence points to its widespread use by 1675. 18th Century documentary evidence is quite plainly against the use of zinc for brasses other than gilding metals on account of cost, and hence the survival of the calamine process into the 19th Century. Craddock (1981; 16) suggests that the iron content should distinguish calamine brass from the later type, with calamine brass containing more iron (*i.e.*, 0.2–0.5%, possibly as high as 'several percent'). Watson (1786) also says that brass produced by the direct method is free from iron, and is therefore good for compasses. The data discussed here shows relatively little iron for all the groups discussed above, with the majority of individual results lying between 0.1 and 0.4% Fe. These are within the range expected for calamine brass, and no significant decrease is observed in the high zinc brasses. In this case the iron therefore appears to give no clear indication of manufacturing technique.

The problem of using this information to detect later copies of scientific instruments and other brass artefacts is one which still requires considerable attention. It cannot be claimed that chemical analysis can be used to date a piece of brass – all that can be done is to compare its composition with those of known genuine pieces. This makes it possible that anachronisms will be found, but the chemical evidence should be considered along with stylistic and other technical reports, simply as part of the overall assessment. Werner (1980) used his vast data bank of analyses of cast brass objects in an attempt to detect known modern castings – he did not succeed in every case. It is always possible, therefore, that some copies will meet the chemical requirements, either by accident or from a knowledge of the metallurgical background. It is also possible that perfectly genuine instruments will have analyses outside the expected range for a particular country and date, but the maximum zinc content is believed to be an absolute thermodynamic (or kinetic) limitation in calamine brass – hence the emphasis on this aspect of the analyses in this work. Much more data needs to be considered before anything more than general statements about 'consistency with published analyses' can be made, but a good start has been made.

REFERENCES

Aitken, W.C. (1866). Brass and brass manufactures. In *The Resources, Products, and Industrial History of Birmingham and the Midland Hardware District*, ed. Timmins, S., Hardwicke, London, pp. 225–281.

Barnard, F.P. (1916). *The Casting-counter and the Counting Board.* Clarendon, Oxford.

Bayley, J. (1990). The production of brass in antiquity with particular reference to Roman Britain. In *2000 Years of Zinc and Brass*, ed. Craddock, P.T., British Museum Occasional Paper No. 50, British Museum, London, pp. 7–27.

Bowman, S.G.E., Cowell, M.R. and Cribb, J. (1989). Two thousand years of coinage in China: an analytical survey. *Historical Metallurgy* **23** 25–30.

Brownsword, R. and Pitt, E.E.H. (1983). Alloy composition of some cast 'latten' objects of the 15/16th centuries. *Historical Metallurgy* **17** 44–49.

Budd, P., Haggerty, R., Pollard, A.M., Scaife, B. and Thomas, R.G. (in press). Rethinking the quest for provenance. *Antiquity.*

Caley, E.R. (1964). *Orichalcum and Related Ancient Alloys, Origins, Composition and Manufacture with Special Reference to the Coinage of the Roman Empire.* American Numismatic Society Notes and Monographs No. 151, New York.

Caley, E.R. (1967). Investigations on the origin and manufacture of orichalcum. In *Archaeological Chemistry*, ed. Levey, M., American Chemical Society, Division of the History of Chemistry, University of Pennsylvania Press, pp. 59– 74.

Cameron, H.K. (1974). Technical aspects of monumental brasses. *Archaeological Journal* **131** 215–237.

Craddock, P.T. (1978). The composition of the copper alloys used by the Greek, Etruscan and Roman civilizations. 3, The origins and early use of brass. *Journal of Archaeological Science* **5** 1–16.

Craddock, P.T., (1981). The copper alloys of Tibet and their background. In *Aspects of Tibetan Metallurgy*, eds. Oddy, W.A. and Zwalf, W., British Museum Occasional Paper No. 15, British Museum, London, pp. 1–32.

Craddock, P.T. (1990). Zinc in classical antiquity. In *2000 Years of Zinc and Brass*, ed. Craddock, P.T., British Museum Occasional Paper No. 50, British Museum, London, pp. 1–6.

Craddock, P.T. (ed.) (1990). *2000 Years of Zinc and Brass*, British Museum Occasional Paper No. 50, British Museum, London.

Craddock, P.T. (1995). *Early Metal Mining and Production.* Edinburgh University Press, Edinburgh.

Craddock, P.T., Burnett, A.M. and Preston, K. (1980). Hellenistic copper base coinage and the origins of brass. In *Scientific Studies in Numismatics*, ed. Oddy, W.A., British Museum Occasional Paper No. 18, British Museum, London, pp. 53–64.

Craddock, P.T., Freestone, I.C., Gurjar, L.K., Middleton, A.P. and

Willies, L. (1990). Zinc in India. In *2000 Years of Zinc and Brass*, ed. Craddock, P.T., British Museum Occasional Paper No. 50, British Museum, London, pp. 29–72.

Dawkins, J.M. (1950). *Zinc and Spelter.* Zinc Development Agency, Oxford.

Day, J. (1973). *Bristol Brass: A History of the Industry.* David and Charles, Newton Abbot.

Day, J. (1990). Brass and zinc in Europe from the Middle Ages until the 19th Century. In *2000 Years of Zinc and Brass*, ed. Craddock, P.T., British Museum Occasional Paper No. 50, British Museum, London, pp. 123–150.

Day, J. and Tylecote, R.F. (1991). *The Industrial Revolution in Metals.* The Institute of Metals, London.

Finnegan, J.E., Hummel, R.E. and Verink, E.D. (1981). Optical studies of dezincification in alpha-brass. *Corrosion* **37** 256–261.

Grant, E. (ed.) (1974). *A Sourcebook on Medieval Science.* Harvard University Press, Harvard.

Haedecke K. (1973). Gleichgewichtsverhältnisse bei der messingherstellung nach dem Galmeiverfahren. *Erzmetall* **26** 229–233.

Hamilton, H. (1967). *The English Brass and Copper Industries to 1800.* Cass, London (2nd edn.).

Hawthorne, J.G. and Smith, C.S. (trans.) (1979). *Theophilus, De Diversis Artibus.* Dover Publications, New York.

Hedges, R.E.M. (1979). Analysis of the Drake plate: comparison with the composition of Elizabethan brass. *Archaeometry* **21** 21–26.

Hegde, K.T.M. (1989). Zinc and brass production in Ancient India. *Interdisciplinary Science Reviews* **14** 86–96.

Holmyard, E.J. (1957). *Alchemy.* Penguin Books, Harmondsworth.

Hoover, H.C. and Hoover, H.L. (trans.) (1950). *Agricola, De Re Metallica.* Dover Publications, New York.

Hughes, M.J., Cowell, M.R. and Craddock, P.T. (1976). Atomic absorption techniques in archaeology. *Archaeometry* **18** 19–37.

Hughes, M.J., Northover, J.P. and Staniaszek, B.E.P. (1982). Problems in the analysis of leaded bronze alloys in ancient artefacts. *Oxford Journal of Archaeology* **1** 359–363.

Lones. T.E. (1919). *Zinc and its Alloys.* Pitmans, London.

Michel, H.V. and Asaro, F. (1979). Chemical study of the Plate of Brass. *Archaeometry* **21** 3–19.

Mitchiner, M.B., Mortimer, C. and Pollard, A.M. (1985). The chemical compositions of English seventeeth-century base metal coins and tokens. *British Numismatic Journal* **55** 144–163.

Mitchiner, M.B., Mortimer, C. and Pollard, A.M. (1987a). The chemical

compositions of nineteenth-century copper-base English jettons. *British Numismatic Journal* **57** 77–88.

Mitchiner, M.B., Mortimer, C. and Pollard, A.M. (1987b). Nuremberg and its jettons. c.1475 to 1888: chemical compositions of the alloys. *Numismatic Chronicle* **147** 114–155.

Mitchiner, M.B., Mortimer, C. and Pollard, A.M. (1988). The alloys of Continental copper-base jettons (Nuremburg and Medieval France excepted). *Numismatic Chronicle* **148** 117–128.

Mitchiner, M.B. and Pollard, M. (1988). Reckoning counters: patterns of evolution in their chemical composition. In *Metallurgy in Numismatics 2*, ed. Oddy, W.A., Royal Numismatic Society, London, pp. 105–126.

Mortimer, C. (1989). *X*-ray fluorescence analysis of early scientific instruments. In *Archaeometry: Proceedings of the 25th International Symposium*, ed. Maniatis, Y., Elsevier, Amsterdam, pp. 311–317.

Needham, J. (1974). *Science and Civilisation in China, Vol. 5 Chemistry and Chemical Technology. Part II: Spagyrical Discovery and Invention: Magisteries of Gold and Immortality*. Cambridge University Press, Cambridge.

Page-Phillips, J., (1972). *Macklin's Monumental Brasses*. Allen and Unwin, London.

Partington, J.R. (1961). *A History of Chemistry*. Macmillan, London.

Pollard, A.M. (1983a). Authenticity of brass objects by major element analysis? Paper presented at *Symposium on Archaeometry*, Castel Dell'Ovo, Naples, 18th–23rd April 1983.

Pollard, A.M. (1983b). An investigation of the brass used in Medieval and later European scientific instruments. Paper presented at *The Preservation of Historical Scientific Materials*, UKIC Meeting, Geological Museum, London, 14th November 1983.

Pollard, A.M. (1988). Alchemy – a history of early technology. *School Science Review*, June 1988, 701–712.

Pollard, A.M., Thomas, R.G. and Williams, P.A. (1990). Experimental smelting of arsenical copper ores: implications for Early Bronze Age copper production. In *Early Mining in the British Isles*, eds. Crew S. and Crew P., Plas Tan y Bwlch Occasional Paper No. 1, Plas Tan y Bwlch, pp. 72–74.

Pollard, A.M., Thomas, R.G. and Williams, P.A. (1991). Some experiments concerning the smelting of arsenical copper. In *Archaeological Sciences 1989 Proceedings of Conference, University of Bradford, September 1989*, eds. Budd, P., Chapman, B., Jackson, C., Janaway, R. and Ottaway, B., Oxbow Monograph 9, Oxbow Books, Oxford, pp. 169–174.

Polunin, A.V., Pchelnikov, A.P., Losev, V.V. and Marshakov, I.K.

(1982). Electrochemical studies of the kinetics and mechanism of brass dezincification. *Electrochimica Acta* **27** 465–475.

Preston, K. (1980). *Roman Brass Coinage: Its Origins and the Influence of its Introduction on First Century A.D. Provincial Mints.* Unpublished M.A. dissertation, Department of Archaeological Sciences, University of Bradford.

Rollason, E.C. (1973). *Metallurgy for Engineers.* Edward Arnold, London (4th edn.).

Scott, D.A. (1991). *Metallography and Microstructure of Ancient and Historic Metals.* Getty Conservation Institute, Los Angeles.

Shaw, K. (1973). *Section 34 – Barberini Collection,* Department of Astronomy, National Maritime Museum, London.

Smith, C.S., and Gnudi, M.T. (trans.) (1959). *Biringuccio, Pirotechnia.* Basic Books, New York.

Trethewey, K.R. and Pinwill, I. (1987). The dezincification of free-machining brasses in seawater. *Surface and Coatings Technology* **30** 289–307.

Tylecote, R. (1976). *A History of Metallurgy.* Metals Society, London.

Watson, R. (1786). *Chemical Essays* Vol. 4. Evans, London.

Werner O. (1970). Über das vorkommen von zink und messing im altertum und im mittelalter. *Erzmetall* **23** 259–269.

Werner, O. (1977). Analysen Mittelalterlicher bronzen und messinge I. *Archäologie und Naturwissenschaften* **1** 144–220.

Werner, O. (1980). Composition of recent reproduction castings and forgeries of Mediaeval brasses and bronzes. *Berliner Beitrage zur Archaeometrie* **5** 11–35.

Werner, O. (1982). Analysen Mittelalterlicher bronzen und messinge II und III. *Archäologie und Naturwissenschaften* **2** 106–170.

Xu Li (1990). Traditional zinc-smelting technology in the Guma district of Hezhang County. In *2000 Years of Zinc and Brass*, ed. Craddock, P.T., British Museum Occasional Paper No. 50, British Museum, London, pp. 103–121.

Chapter 7

The Chemistry and Use of Resinous Substances

INTRODUCTION

Naturally occurring organic molecules are as important in the contemporary world as they were in the past. Although the synthetic chemical industry plays a major role through the production of a vast range of molecules tailored to particular needs, the study of natural products in the living world continues unabated. For example, chemical exploration of the constituents of plant and animal tissues has major pharmacological implications. The pharmacological activities of a single albeit diverse group of molecules known as alkaloids has long been known: '*since early times selected plant products (many containing alkaloids) have been used as poisons for hunting, murder and euthanasia; as euphoriants, psychedelics, and stimulants* (e.g., *morphine and cocaine); or as medicines* (e.g., *ephedrine, for respiratory problems)*' (Mann *et al.* 1994; 389). Other naturally occurring plant and animal tissues can be used as adhesives, disinfectants, sealants, dyestuffs, perfumes, incenses, waterproofing agents, and so on. Direct archaeological evidence for such activities in the past is difficult to obtain and many researchers rely on textual sources, the behaviour of contemporary industrial and non-industrial peoples, and whatever fragmentary evidence the archaeologist can place into the equation in terms of botanical remains, material culture (through the presence of specific artefacts, and so on), and occasional chemical investigations. In contrast to the limited chemical evidence for the use of alkaloids, another class of naturally occurring chemical compounds, the terpenoids, have been found to survive in a large number of archaeological contexts from around the world. Terpenoids are the major constituents of resins. The contrast in abundance may be due to the survivability and visibility of the latter compound class in a range of burial environments, although the fact that

few explicit methodologies have been applied to explore the survival of alkaloids in archaeological contexts is significant.

The aim of this chapter is to review generally the potential for the study of the organic chemistry of amorphous deposits associated with artefacts (such as pottery vessels, stone tool surfaces, and other materials such as basketry). Artefacts such as pottery vessels were used frequently to prepare, store, transport, serve, and consume foodstuffs and other natural products in the past. Consequently traces of these substances may be preserved on the surface of the vessel. Similarly, organic molecules can occlude in the permeable ceramic matrix. These chemical remnants offer valuable clues to the use of pottery and other artefacts and may provide novel identifications of food items and other organic substances in the archaeological record. For the sake of brevity, only one category of natural product is reviewed: higher plant resins and related substances. The aim is to give examples of identification of aged samples and to consider the range of uses to which these substances were put. The case study focuses on the Neolithic of northern Europe. The summary then includes a broader statement on the potential for study of other organic substances.

RESINS: DEFINITION AND USES

A *resin* is one of a number of natural products defined as a *plant exudate.* Other exudates include latex, gum, and kino. Resins are non-cellular, water-insoluble substances and serve to protect higher plants, if wounded, from excessive water loss and the invasion of microorganisms. Resins often comprise both volatile and nonvolatile fractions. Gianno (1990; 5) distinguishes exudates from *extractives* (the latter include tannins, alkaloids, and essential oils and require, for example, a solvent to enable their isolation from plant fluids). The derivatives produced by heating resin, as well as resinous wood are collectively referred to as *pyroligeneous* substances. These include tar, the initial pyrolysate and pitch, a thicker substance derived from further heating of the tar to drive off remaining volatiles. Tar and pitch are technological terms not often used by botanists and chemists (Gianno, 1990; 8).

Resin-producing trees are found over vast areas, encompassing much of the torrid and temperate parts of the world. Resin preserves reasonably well and possesses a wide array of functional attributes. The properties of *'adhesiveness, insolubility in water, inflammability, healing and poisoning properties, fragrance, plasticity, vitreosity, colorability, pigment mediability, and resistance to spoilage are qualities that apply, to a greater or lesser degree, to all resins.'* (Gianno, 1990; 1). As such, these natural substances have played a role in most communities. In the modern world, synthetic chemicals,

including those obtained from petroleum or coal, have replaced many of the uses of natural resins.

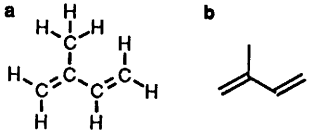
In the past, natural resins served as universal adhesives for fixing stone points and blades to hafts of wood or antler, to glue feathers to arrow shafts, and for repairing materials such as broken pottery and bone combs. As sealants, resins were used to coat the surfaces of pottery and basketry in order to provide an impermeable lining. Resins were also used to waterproof hunting equipment and fishing nets, as well as canoes and ships from Noah's Ark (Genesis 6:14) to the Mary Rose (Robinson *et al.*, 1987). A wide range of less visible uses have been documented historically and/or ethnographically. The inflammability of resins was exploited as a source of light and even liquid fire. Burning resin could be used as incense and its disinfectant properties found use in wine (notably retsina) and embalming. Resins also contributed to medicinal preparations. Finally, resins were undoubtedly the earliest chewing gums; mastic from *Pistacia lentiscus*, birch bark tar, pine resins, and many other resinous substances were chewed to alleviate toothache and sore throats.

CHEMISTRY OF RESINS

An up-to-date and comprehensive survey of the chemistry of natural resins relevant to art historical and archaeological contexts is given in Mills and White (1994; 95–128) and is not repeated here. Rather those aspects of structure and chemistry relating to molecular transformation and to the composition of aged resins encountered in archaeological contexts will be emphasized.

Terpenes, or terpenoids, are distributed widely in plants from marine and terrestrial sources. The term terpene derives from *'terpen'* and is attributed to Kekulé who used it to describe $C_{10}H_{16}$ hydrocarbons in turpentine oil. Terpenes designate molecules made up of isoprene units and are collectively referred to as secondary metabolites, since it has long been considered that these molecules do not play a role in primary metabolism. However, this view is being eroded as an increasing number have been found to play active roles in ecological interactions and as defence and attack chemicals (Dev, 1989; 790–791).

Isoprene (2-methylbuta-1,3-diene; [7.1a and b]) is a C_5 unit. Structure 7.1a shows the full structural formula where each line between the atoms represents two shared electrons in a covalent bond. In the case of more complex molecules, skeletal structures are used as in 7.1b where carbon atoms are normally represented by an intersection of bonds. Carbon to hydrogen bonds are not shown although all other atoms (O, N, P, and so on), if present, are indicated. Stereochemistry is indicated by bold lines



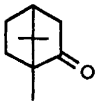
Structures 7.1a and 7.1b *2-Methylbutadiene (isoprene). a) The full structural formula, and b) the abbreviated skeletal structure*

or wedges (β-bonds, coming out of the plane of the paper) and broken lines (α-bonds, going behind the plane of the paper).

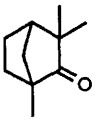
The term *terpenoid* is now the preferred generic name for this class of natural products. Many thousands of individual terpenoid molecules are known, belonging to a range of skeletal types. These are classified in terms of the number of isoprene units. For the purposes of this chapter the groups noted in the following section are relevant (see Dev, 1989 and Banthorpe, 1994 for recent comprehensive reviews). Trivial names for terpenoids and a great many other natural products remain much used since the systematic names are cumbersome and are seldom used unless the structures are sufficiently simple.

Monoterpenoids and Sesquiterpenoids

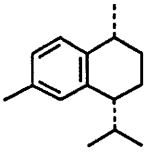
Mono- and sesquiterpenoids are the usual constituents of essential oils, although they are also found in algae. Monoterpenoids are C_{10} compounds which are distributed widely in the plant kingdom. Many structures are known, including a number of oxygenated molecules with alcohol, ketone, and ether groups. The actual chemical properties are determined by the nature of the functional group. The volatile fraction obtained by distilling pine resin is known as *oil of turpentine* and is composed largely of monoterpenoids. Monoterpenoids and sesquiterpenoids (C_{15} compounds), although volatile, do survive over considerable timespans in favourable preservation environments. In fossil resins, such as amber, combined gas chromatography–mass spectrometry (GC-MS) has demonstrated the survival of monoterpenoids (such as camphor [Structure 7.2] and fenchone [Structure 7.3]) and sesquiterpenoids held within the natural polymer (Mills *et al.*, 1984/85). The survival of the sesquiterpenoids calamenene [Structure 7.4] and cadelene [Structure 7.5] has been demonstrated in the softwood (probably pine) tar samples that had poured out of transport vessels (amphoras) at the site of the wreck of a 6–7th Century BC Etruscan boat located off the coast of the Italian island of Giglio (Robinson *et al.*, 1987). These molecules will survive if protected from oxidation within a solidified mass of resinous material.



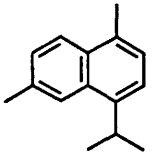
Structure 7.2 *Camphor*



Structure 7.3 *Fenchone*



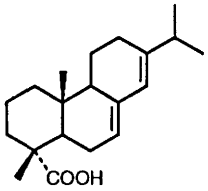
Structure 7.4 *Calamenene*



Structure 7.5 *Cadelene*

Diterpenoids

Diterpenoids (C_{20} compounds) comprise the bulk composition of resins from the families Coniferae (encompassing Pinaceae, Cupressaceae, and Araucariaceae) and Leguminosae. The most abundant source of resin in temperate regions are trees of the genus *Pinus*. Diterpenoid compounds possess mainly abietane, pimarane, and labdane skeletons. In 'soft' resins (*i.e.*, those containing no polymerized structures), such as those derived from Pinaceae, abietane and pimarane compounds are predominant. The Pinaceae, and especially *Pinus*, generally have resins with a high content of abietic acid [Structure 7.6], a tricyclic molecule and a small number of abietane isomers (Mills and White, 1994; 98–102), but resins from *Abies* and *Picea* species also contain large amounts of labdanes, as do Cupressaceae resins. Some *Pinus* resins are rich in pimarane-type acids (Fox *et al.*, 1995).



Structure 7.6 *Abietic acid*

The double bonds in abietane acids are conjugated and in fresh *Pinus* resins will undergo significant modification during treatment. Warming of the resin (*e.g.*, as is the case during distillation to remove oil of turpentine) induces isomerization reactions leading to a mixture

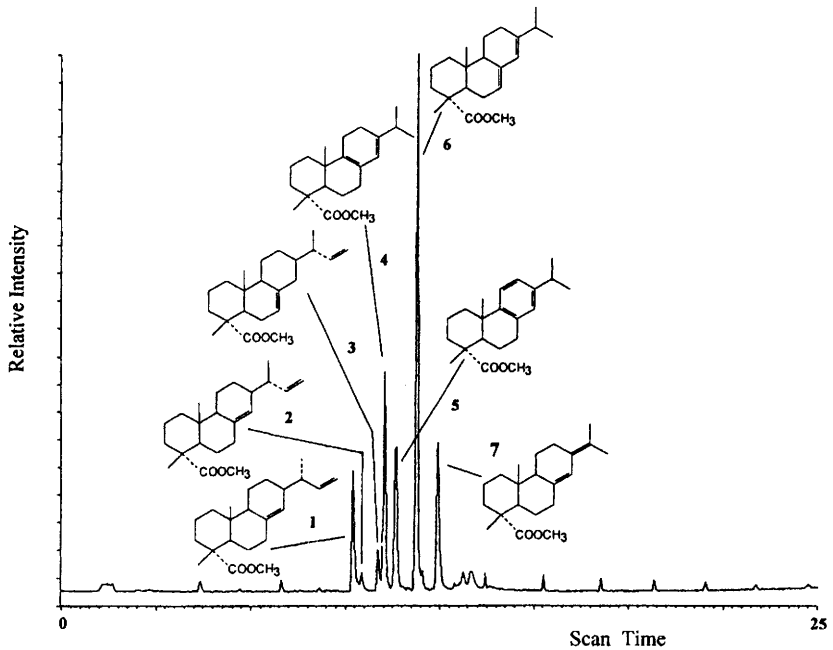
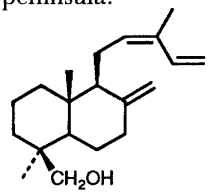


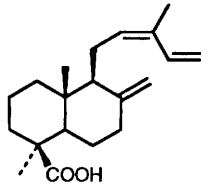
Figure 7.1 *Total ion current (TIC) chromatogram obtained by GC-MS analysis of a resin* (Pinus sylvestris). *The diterpenoid resin acids were methylated (using diazomethane) to improve chromatographic performance. Peak identities: 1, Methyl pimarate; 2, Methyl sandaracopimarate; 3, Methyl isopimarate; 4, Methyl palustrate; 5, Methyl dehydroabietate; 6, Methyl abietate; 7, Methyl neoabietate. For GC-MS operating conditions, see Heron and Pollard (1988)*

enriched in abietic acid at the expense of other abietane molecules (Mills and White, 1977; 14; see Figure 7.1). The solid product remaining is referred to as *rosin* or *colophony*.

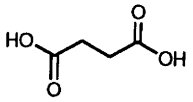
The fossil resin, amber, comprises a complex mixture of molecules based primarily on diterpenoid and monoterpenoid structures. Amber is a hard resin (2–2.5 on the Moh scale) and behaves largely as a polymer. This fraction is formed from the esterification of a polyvalent alcohol with a dibasic acid (Mills and White, 1994; 110). The polyvalent alcohol is the co-polymer of communol [Structure 7.7] and communic acid [Structure 7.8] (only the *cis* forms are shown), as found in kauri resin from *Agathis australis* (Araucariaceae), although additional modifications may have occurred as a result of geological conditions. The dibasic acid is succinic acid [Structure 7.9], the molecule used by Otto Helm as the basis of his scheme for separating Baltic amber (succinite) from other fossil resins in the late 19th Century (see Chapter 1). The solvent (ether) soluble portion of amber contains hundreds of individual molecular species, many of which can be identified using GC-MS (Mills *et al.*, 1984/85). Amber occurs naturally over large areas of northern Europe, although the richest sources remain the east Baltic coast and the west Jutland peninsula.



Structure 7.7 cis-*Communol*



Structure 7.8 cis-*Communic acid*



Structure 7.9 *Succinic acid*

Triterpenoids

Triterpenoids (C_{30} compounds) are the most ubiquitous of the terpenoids and are found in both terrestrial and marine flora and fauna (Mahato *et al.*, 1992). Diterpenoids and triterpenoids rarely occur together in the same tissue. In higher plants, triterpenoid resins are found in *'numerous genera of broad-leaved trees, predominantly but not exclusively tropical.'* (Mills and White, 1994; 105). They show considerable diversity in the carbon skeleton (both tetracyclic and pentacyclic structures are

found) which occur in nature either in the free state or as glycosides, although many have either a keto or a hydroxyl group at C-3, with possible further functional groups and/or double bonds in the side-chains. Tetracyclic triterpenoids belong primarily to the steroids.

A number of important resins are composed of triterpenoids, including the dammar resins derived from trees of a sub-family of the family Dipterocarpaceae. Dammar resins are fluid, balsamic oleoresins highly suited for caulking and waterproofing. Frankincense (olibanum) is a gum-resin collected from various *Boswellia* spp. and contains amyrin epimers and triterpenoid acids. The gum component is polysaccharide in origin and is water soluble. The Anacardiaceae family contains the genus *Pistacia* (Mills and White, 1977; 21; Mills and White, 1989).

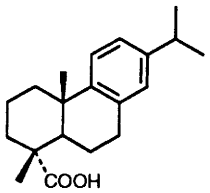
ANALYSIS OF RESINS IN ARCHAEOLOGICAL CONTEXTS

Preservation of resin is generally favoured in anaerobic or near-anaerobic locations (as in the case of waterlogged deposits as well as marine and lacustrine areas), under permafrost or in arid environments. These conditions provide protection against atmospheric oxidation and photoxidation and reduce the activity of microorganisms. The extent of post-depositional chemical alteration will also be dictated by the physical state of the resin (whether it is present as a thin film or as a thick deposit of material). Chemical analysis of putative resin samples usually proceeds with the aim of identifying the botanical source, although precise species-specific identification of aged samples is problematic. Geographical origin is important if it is suspected that resins may have been transported. Since resins can be observed on artefact (ceramic, stone, bone, and wood) surfaces, analysis offers opportunities for assessing the ways in which artefacts were used. The analysis of ancient resins is also relevant to procurement and methods of preparation as well as in determining the range of uses that these substances were put to. Identification of ancient resin samples is not a straightforward task. Although ancient resins have been investigated for many decades, their chemical complexity has hindered confident assignments. Visual characteristics and examination of simple chemical or physical properties may offer little or no clue as to the identity of resin samples, whether ancient or modern. Consequently, chemical analysis must be performed in order to characterize which molecular species are present.

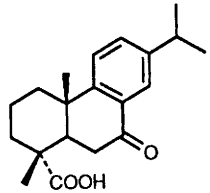
The development of sensitive and specific analytical techniques has improved significantly the opportunities for identifying organic molecules. In particular, GC and GC-MS (see Chapter 2) are ideal for the

separation and characterization of individual molecular species. Characterization generally relies on the principle of chemotaxonomy, where the presence of a specific compound or distribution of compounds in the ancient sample is matched with its presence in a contemporary authentic substance. The use of such '*molecular markers*' is not without its problems, since many compounds are widely distributed in a range of materials, and the composition of ancient samples may have been altered significantly during preparation, use, and subsequent burial. Other spectroscopic techniques offer valuable complementary information. For example, infrared (IR) spectroscopy and ^{13}C nuclear magnetic resonance (NMR) spectroscopy have also been applied.

Terpenoids are susceptible to a number of alterations mediated by oxidation and reduction reactions as well as changes brought about by microbial activity. For example, the most abundant molecule in aged *Pinus* samples is commonly dehydroabietic acid [Structure 7.10], a monoaromatic diterpenoid based on the abietane skeleton which occurs in fresh (bleed) resins only as a minor component. This molecule forms during oxidative dehydrogenation of abietic acid, which predominates in rosins. Further atmospheric oxidation (autoxidation) leads to the formation of 7-oxodehydroabietic acid [Structure 7.11]. This molecule has been identified in many aged coniferous resins such as those used to line transport vessels in the Roman period (Heron and Pollard, 1988; Beck *et al.*, 1989), in thinly spread resins used in paint media (Mills and White, 1994; 172–74), and as a component of resin recovered from Egyptian mummy wrappings (Proefke and Rinehart, 1992).



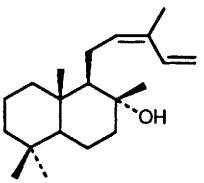
Structure 7.10 *Dehydroabietic acid*



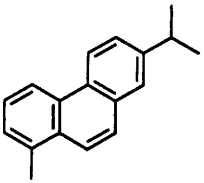
Structure 7.11 *7-Oxodehydroabietic acid*

Distinguishing between genera, for example, of larch, fir, and pine resins should be possible even though they contain many of the same acids. In addition to abietane and pimarane acids, larch resins contain labdane alcohols. Fir resins have large amounts of a labdane alcohol (*cis*-abienol [Structure 7.12]) which although susceptible to alteration should give rise to recognizable oxidation products (White, 1992; 8). Given the changes in fresh resins during preparation, use, and deposition, linking an aged sample to a species-specific origin may not be feasible. However,

White (1992; 7) suggests that in certain cases the ratio of pimaradiene (pimarane, sandaracopimaric, and isopimaric) acids might be instructive. Since *Pinus* spp. are the most abundant source of resin, the assumption is frequently made that a pine source is suspected.



Structure 7.12 cis-*Abienol*



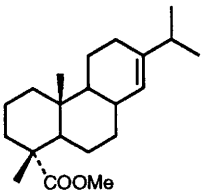
Structure 7.13 *Retene*

When resin or resinous wood is heated strongly, significant changes in resin composition occur leading to the formation of tar and pitch. Chemical changes include thermal dehydrogenation, decarboxylation, and demethylation which gives rise to a large number of potential alteration products of varying aromaticity. Stable end-products of these reaction pathways include retene [Structure 7.13], a triaromatic defunctionalized diterpenoid with the formula $C_{18}H_{18}$. Intermediates include dehydroabietane, dehydroabietin, simonellite, the *nor*-abietatrienes, and tetrahydroretene, although some of these molecules are present in low abundance in relatively fresh bleed resins.

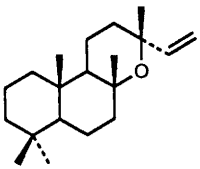
Recently, Beck *et al.* (in press) used GC-MS to monitor the increase in the proportion of retene in pine tars with increasing temperature and ^{13}C NMR to monitor the increased aromatic signal resulting from dehydrogenation reactions. Diterpenoid molecules of probable pine origin have been detected in many archaeological contexts and some detailed compositional studies have appeared (Robinson *et al.*, 1987; Heron and Pollard, 1988; Beck *et al.*, 1989; Reunanen *et al.*, 1989; Beck *et al.*, 1994).

The potential complexity of softwood tars is exemplified by the analysis of a black resinous substance which filled a ceramic container recovered from a late Roman shipwreck dating to the late 4th Century AD (Beck *et al.*, 1994). It is thought that the tar was used on board the ship as naval stores. The tar, separated into acid and neutral fractions, was analysed using GC-MS. The methylated acid fraction comprised 54 components with the unusual methyl abiet-13-en-18-oate [Structure 7.14] present in greater abundance than methyl dehydroabietate. In the unmethylated neutral fraction, comprising 61 recognizable peaks, sesquiterpenoids, diterpenoid hydrocarbons, and methyl dehydroabietate were identified. Whereas compositional data may inform on the specific

nature of the production process (precise source of the tar, temperature, presence of admixtures and so on), there *'remains much to be learned about the relationship between the present composition and the history, both pre- and post-depositional, of tars and pitches.'* (Beck *et al.*, 1994; 119).



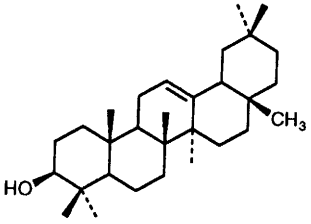
Structure 7.14 *Methyl abiet-13-en-18-oate*



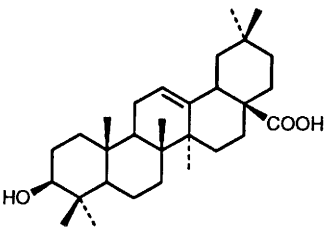
Structure 7.15 *Manoyl oxide*

In another study carried out by Biers and co-workers (Biers *et al.*, 1994), the residual contents of intact Corinthian 'plastic' vases of the 7–6th Centuries BC were analysed non-destructively by pouring solvent into the vessels and decanting. A large number of mono-, sesqui-, and diterpenoids were identified in the solvent washes. The diterpenoid, manoyl oxide [Structure 7.15] was identified in 16 vases. This molecule is found in the bark of *Pinus/Abies* spp. and in the essential oils of the Cupressaceae family, including *Juniperus oxycedrus.*

In general, triterpenoids identified in the archaeological record have not been subject to significant structural alteration from fresh authentic samples although these have occurred in favourable preserving environments, such as marine contexts. The identity of resin preserved in around 100 Canaanite amphoras from the near anaerobic conditions of the Late Bronze Age wreck at Ulu Burun in southern Turkey has been demonstrated using GC-MS (Mills and White, 1989). Triterpenoids congruent with the tree *Pistacia atlantica* Desf. which grows extensively in the eastern Mediterranean, the Near East and North Africa were identified. A number of neutral and acidic triterpenoids were identified, including β-amyrin [Structure 7.16] and oleanonic acid (as its methyl derivative; [Structure 7.17]). This resin, also known as *Chios turpentine*, was possibly



Structure 7.16 *β-Amyrin*



Structure 7.17 *Oleanonic acid*

used as incense. This is an important find since it represents a substantial cargo of resin (and myriad other materials) and indicates trade in resins during the Late Bronze Age in the eastern Mediterranean. Chemical characterization of other ancient resins, in particular the gum-resins, such as frankincense from *Boswellia* spp. and myrrh from *Commiphora* spp. has yet to be demonstrated using GC-MS. However, there is ample documentary evidence for trade and use of these substances.

A study of 15–17th Century AD shipwrecks in southeast Asia has provided another example of the trade in resins (Gianno *et al.*, 1990). Resin composed largely of triterpenoids, possibly from a tree of the family Dipterocarpaceae, was identified by FTIR and GC-MS. Three other samples of resin found within jars that were still sealed have been interpreted as representing a valuable commodity which was being traded; an interpretation which is supported by their identification as being derived from a *Styrax* spp. tree, possibly *Styrax benzoin.* This is known from historical and ethnographic sources as a valued incense and medicinal resin, and is believed to have been traded over a wide area since the 13th Century AD.

Terpenoids occur widely in the sedimentary record such as deep-sea sediments, fossil resin, coal, and so on (Simoneit *et al.*, 1986). Diterpenoids serve as valuable marker compounds of terrigenous resinous plants while sesquiterpenoids can have either a higher plant or an algal origin. In sediments, organic molecules are subjected to increases in temperature and pressure and considerable structural modifications, rearrangements, and polymerization reactions may occur. Recently, ten Haven *et al.*, (1992) have suggested a scheme for the geological fate of terrigenous triterpenoids during the earliest stages of diagenesis. The extent of preservation of individual molecules depends on their discrete chemical structure (presence of double bonds, labile side-chains, and so on). Aromatization of higher plant diterpenoids is favoured where carbon–carbon double bonds are present in a six carbon ring, although hydrogenation (reduction) reactions may also occur. Indeed the degree of functionality retained in terpenoids in the sedimentary record provides a means for considering the extent of diagenesis (Grimalt *et al.*, 1989). Generally, free terpenoids are converted into a similar suite of defunctionalized molecules found in wood tars. Polyterpenoids based on labdane structures are generally referred to as resinite. These 'fossilized' products of resins are widespread and abundant in the geosphere. Recently, pyrolysis GC-MS techniques have been applied with a view to characterizing the source resin acid composition (Anderson and Winans, 1991).

NEOLITHIC TAR

On Thursday 19 September 1991, tourists Erika and Helmut Simon discovered a body protruding from glacial ice below the Hauslabjoch in the Tyrolean (Ötzaler) Alps between Austria and Italy. The realization that the body (later named Ötzi) had lain there for several thousand years took some time and a remarkable account has been written by Konrad Spindler (Spindler, 1993). Radiocarbon dates now available suggest the most likely date of the body to be around 3330 BC. The finds associated with the body are as astounding; in total 20 different items of equipment, including weapons, tools, and containers were recovered. One of these was an almost pure copper axe attached to a haft of yew glued and bound with leather or hide. The glue spread over 4.3 cm, bonding the axe to the wood. Also, in a quiver, trimmed halves of feathers were found attached to arrows made from long shoots of the wayfaring tree (*Viburnum lantana*) with flint tips attached to the arrows using adhesive material (Spindler, 1993; 123–4). The analysis of the glue was carried out by a group of chemists from Vienna Technical University (Sauter *et al.*, 1992) who concluded that the tar was produced by heating bark of the family Betulaceae, possibly *Betula pendula* (also referred to as *B. alba* or *B. verrucosa* in other parts of Europe) in a sealed vessel, thereby limiting the amount of oxygen.

The preparation of bark and wood tars represents one of the earliest chemical/technological processes. Indeed, the archaeological literature devoted to prehistoric Europe contains many references to black, sticky substances and lumps of 'glue' recovered from excavations of sites from the Mesolithic onwards, but only recently has scientific analysis become common as a means of identification. These analyses have stimulated a wider evaluation of the uses of tars and of the artefacts with which these substances are associated. The evidence suggests that the preparation and use of birch bark tar in particular is a pan-European phenomenon at least as early as the Neolithic. In recent years, samples have been identified on artefacts recovered from sites in a number of countries, including Finland, Sweden, France, Germany, Austria, Italy, and Slovenia. Evidence dates back to the Mesolithic in some areas. Birch bark tar has also been identified in later (Roman and Medieval) periods. In certain regions of Europe, birch bark tar survived as an important natural product until the onset of the synthetic chemical industry, and can still be found today in commodities such as birch bark tar soap. The specific origin of the tar from bark is, from chemical evidence alone, an assumption since detailed compositional studies of birch wood and its

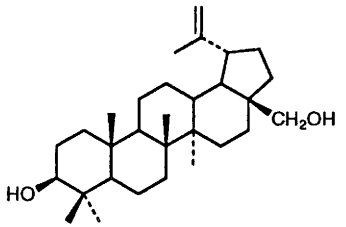
ability to produce tar with the same properties as that obtained from the bark have yet to appear in print. However, the assumption that bark itself is a rich source of 'tar-forming compounds' is backed up by the ethnohistorical record and experimental simulations.

Birch is widespread in the cooler regions of Europe. Although the wood was not valued as timber and was rarely used for building, the bark was much exploited (Vogt, 1949). It could be detached readily in thin layers and was easy to work and sew. Its properties as a waterproofing and insulating material ensured widespread use in roofing and flooring, for making containers which were efficient for food storage and preparation by both indirect and direct heating methods (Vogt, 1949; Dimbleby, 1978; 36) and even for making canoes and tents. Two birch bark containers were found in Ötzi's tool kit, although one was inadvertently destroyed soon after the body was found.

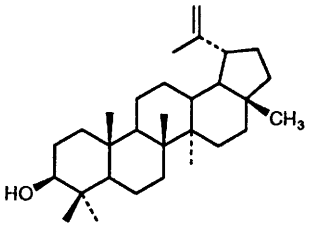
The Chemistry of Birch Bark and Birch Bark Tars

The bark of *B. pendula*, especially the outer part, contains a high level of triterpenoid compounds, in particular betulin or betulinol (lup-20(29)-ene-3β,28-diol [Structure 7.18]; Hayek *et al.*, 1989; Ekman, 1983; Ukkonen and Erä, 1979; O'Connell *et al.*, 1988; Pokhilo *et al.*, 1990; Reunanen *et al.*, 1993; Binder *et al.*, 1990), which contributes more than 70% of the total triterpenoids. Betulin is a white crystalline compound which gives silver birch bark its characteristic colouration. High proportions (> 25% dry weight of the solvent soluble fraction) of betulin are known to occur in the outer bark of *Betula pendula*. The other triterpenoid components include lupeol (lup-20(29)-en-3β-ol; [Structure 7.19]), lupenone (lup-20(29)-en-3-one; [Structure 7.20]) and occasionally, betulinic acid (lup-20(29)-en-3β-ol-28-oic acid; [Structure 7.21]). The bark of young trees contains much larger quantities of terpenoids than older trees (Ukkonen and Erä, 1979; 217), which appears to confirm the statement by Rajewski (1970) that historically only the bark of young trees was selected for the preparation of tar. Triterpenoids in black birch (*Betula lenta*) are much reduced in comparison. In one study, the total triterpenoids amounted to only 0.8% dry weight of the bark with lupeol as the most abundant of the terpenoids. (Cole *et al.*, 1991).

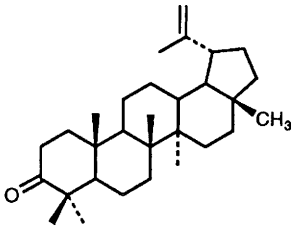
In addition to triterpenoids in the bark of *Betula* spp., suberin (a biopolyester comprising primarily hydroxy, epoxy, and dicarboxylic acids) is also present, as demonstrated by GC-MS of modern birch bark extracts following alkaline hydrolysis (Ekman, 1983). Suberin monomers include 9,10-epoxy-18-hydroxyoctadecanoic acid [Structure 7.22] and ω-hydroxyacids (*e.g.*, 22-hydroxydocosanoic acid; [Structure 7.23]).



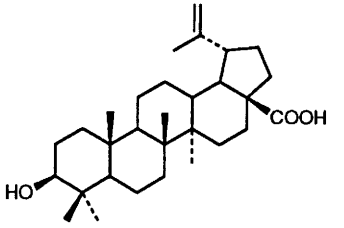
Structure 7.18
Betulin (lup-20(29)-ene-3β,28-diol)



Structure 7.19
Lupeol (lup-20(29)-en-3β-ol)



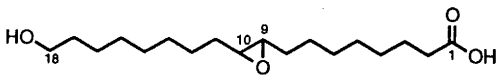
Structure 7.20
Lupenone (lup-20(29)-en-3-one)



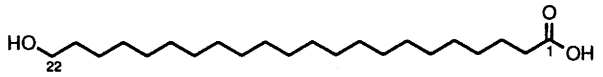
Structure 7.21
Betulinic acid (lup-20(29)-en-3β-ol-28-oic acid)

Suberin is quite resistant to biodegradation by fungi, but, unless protected, extracellular enzymes of fungi will gradually degrade it.

The connection of hafting adhesives with birch bark subjected to a process of 'distillation' has been suggested for a long time. For example, Grewingk (1882;25) and Lidén (1938; 25–8 and 1942; 99–100; both quoted in Clark 1975; 210) considered birch bark to be the source of prehistoric tar. Experimental studies have demonstrated that dry distillation of birch bark in a sealed vessel readily produces a viscous tar at temperatures between 250 and 350 °C (Charters *et al.*, 1993a). According to these early investigations birch bark tar could be modified through the addition of beeswax and animal fat in order to lower its melting point. Grewingk also suggested possible admixtures of pine resin and iron oxide. The chemist Berlin proposed that archaeological samples were birch bark tar mixed with amber (Ruthenberg, in press). Very little information is given regarding the criteria used to characterize these



Structure 7.22 *9,10-Epoxy-18-hydroxyoctadecanoic acid*



Structure 7.23 *22-Hydroxydocosanoic acid*

materials. Ruthenberg (in press) considers that smell during combustion, ('*Juchtengeruch*') and certain distillation properties formed the basis for differentiation. Consequently, these results must be considered with caution and at best should serve only as possible guides towards identification.

Spectroscopic and chromatographic investigations of ancient resinous substances began during the 1960s. In 1965, Sandermann, using IR spectroscopy, concluded that birch bark derivatives were present on potsherds of Neolithic/Bronze Age date from Germany. Similarly, an adhesive used to mend Iron Age pottery was ascribed a birch origin by analysis of carbon, hydrogen, and nitrogen elemental ratios and IR spectroscopy (Sauter, 1967). Using thin layer chromatography (TLC), a birch bark origin has been attributed, on the basis of betulin detection, to adhesives used to haft flint implements (Funke, 1969) and to lumps of tar with human tooth impressions (Rottländer, 1981). It should be emphasized that detections of betulin using TLC alone are questionable because very many triterpenoid molecules exhibit similar R_f values (Hayek *et al.*, 1989; 2229). Infrared and NMR spectroscopy were used to identify the origin of resinous material on lithics and ceramics from northern Yugoslavia (Hadzi and Orel, 1978). The ancient samples resembled spectra from freshly prepared birch bark tar. For detailed compositional information GC and GC-MS are the required techniques.

Analysis of tars of archaeological date by combined GC-MS (Hayek *et al.*, 1990; 1991; Heron *et al.*, 1991a; Reunanen *et al.*, 1993; Charters *et al.*, 1993a) has demonstrated that betulin is retained as a major component in those ascribed a birch origin (Figure 7.2). This is to be expected, given its high thermal stability (Hayek *et al.*, 1989). Betulin seems to be highly resistant to fungal attack as is indicated by exceptional preservation of artefacts fabricated from birch bark, although biodegradation products of betulin have been reported in the natural products literature (*e.g.*, Fuchino *et al.*, 1994).

Hayek *et al.* (1990) reported the identification of 14 samples of birch bark tar of Chalcolithic to Early Iron Age date from sites in Austria and Denmark. In this study, authentic tars from bark samples of a number of different species (*e.g.*, birch, oak, alder, hazel, elm, and so on) were produced under laboratory conditions (using Kugelrohr distillation) and

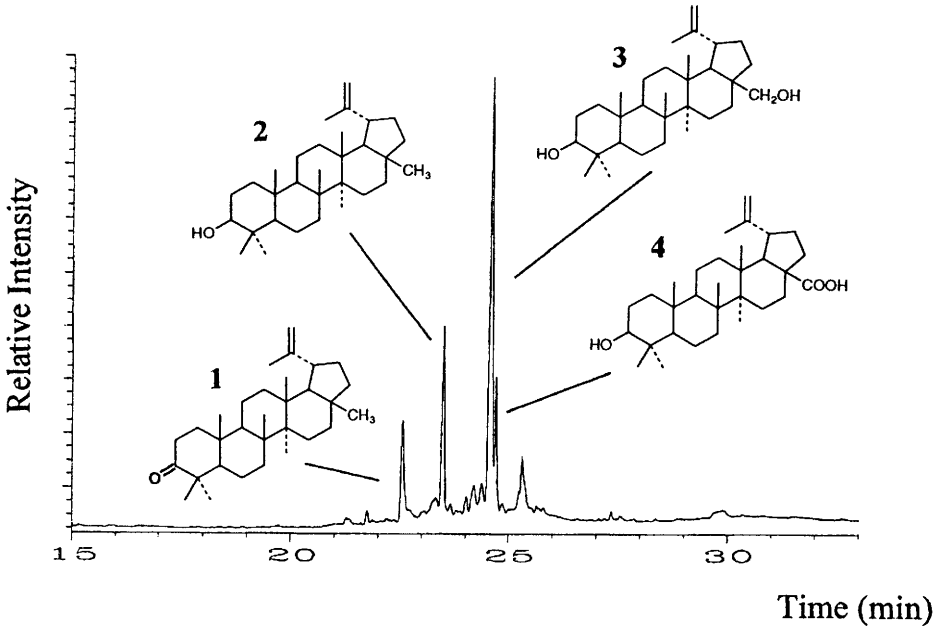
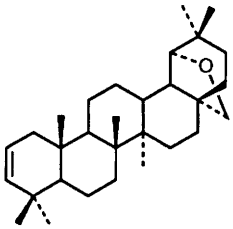


Figure 7.2 *Partial capillary gas chromatogram displaying the principal lipid solvent soluble constituents of a black tarry deposit on the surface of a Neolithic potsherd from Ergolding Fischergasse, Bavaria, Germany. The sample was trimethylsilylated prior to GC (for further details see Bonfield* et al., *in press). Peak identities (confirmed by GC-MS analysis): 1, lupenone; 2, lupeol; 3, betulin; 4, betulinic acid*

compared with the aged samples using GC-MS. Principal components analysis was used to compare aged and modern samples. In each case, the compositional data supported a birch bark origin. However, it was also found that the bark and tar of many species (including beech) contain triterpenoids such as betulin although relative abundances are not given.

The chemical composition of birch bark tar is dependent on the temperature at which tar is produced. In producing simulated tars in the laboratory for comparison with an adhesive used to repair a Roman jar from Stanwick, Northamptonshire, Charters *et al.* (1993a) found that tars prepared at 350 °C displayed an increase in triterpenoid hydrocarbons as well as unresolved components presumably resulting from pyrolysis although the precise nature of these molecules has not been elucidated. Binder *et al.* (1990) and Charters *et al.* (1993a) also report the presence of allobetul-2-ene [Structure 7.24] in aged birch bark tars. Since this molecule has not been reported in extracts from fresh birch bark, it is likely to be formed during heating to produce the tar.



Structure 7.24 *Allobetul-2-ene*

Analysis of a large number of amorphous deposits surviving on artefact surfaces at the Neolithic waterlogged settlement at Ergolding Fischergasse, Germany (mid-4th Millennium BC; Ottaway, 1995) has confirmed the abundance of birch bark tar (Bonfield *et al.*, in press). Although the majority of finds sampled are from potsherds, similar deposits have been found on other artefacts from the site, including stone implements (sickle blades and whetstones) and worked bone. Evaluation of the position of the deposits on the potsherds suggested the most common application of the tar was as a layer on the inner surface, but layers on the outside wall are also found, including an unusual example where a smoothed layer appears to have been applied deliberately to the outer surface of the base of a pot. Interior linings may have served to seal the permeable fabric allowing storage of liquids, although the pottery was well fired and of apparent low porosity. Alternatively, Schiffer (1990) has shown that interior resin-lined pottery possesses excellent heating effectiveness, so

these linings could represent use for direct boiling containers. Ethnographic evidence exists for the application of resin to the outer surfaces of pots (*e.g.*, Foster, 1956), which may have improved transmission of heat. In truth, a great diversity of organic substances have been used to coat and seal pottery (*e.g.*, Rice, 1987; 163–164). For example, traditional potters of sub-Saharan Africa coat their pots with aqueous extracts of stem bark from a number of plants, including *Bridelia furruginea*; a recent phytochemical study has investigated the procyanidins thought to be important in the process of providing an impermeable layer (Diallo *et al.*, 1995). Tar may also have been applied in order to facilitate handling when the vessel was full and heavy, as is suggested for the characteristic rough exteriors to the vessels from Ergolding Fischergasse.

Deposits on three sherds of pottery from Ergolding Fischergasse were identified as softwood tar. Resin acids found in abundance in bleed resins are absent although this is not unexpected in strongly heated samples. A high proportion of dehydroabietic acid and defunctionalized diterpenoids such as retene suggest that these samples are wood tars. The composition of these samples is very similar to those reported for the putative pine tars by Robinson *et al.* (1987) used on the Mary Rose. In one study, Beck and Borromeo (1990) have suggested that the presence of methyl esters in a tar indicates that it is produced from the wood since methanol in the resinous wood reacts with the free resin acids during heating to form the methyl derivative. Tars formed by heating resin alone (obtained by tapping) comprise resin acids but no esters. Of course, if methylation is the chosen method of derivatization prior to GC separation then it is not possible to distinguish whether there are any methyl esters originally present in the tar. However, by separating the soluble fraction into acid and neutral constituents, the presence of methyl esters in the non-methylated neutral fraction can be used to assess the production method. Alternatively, if the samples are trimethylsilylated then no fractionation is required.

The Production and Uses of Neolithic Tars

The precise methods for the preparation of tar in prehistory are not well understood. Archaeological and documentary evidence exists for sophisticated and large-scale methods of manufacture in the Medieval period (Kurzweil and Todtenhaupt, 1991). A number of simpler methods involving heating in sealed pottery vessels have been proposed and tested (Kurzweil and Todtenhaupt, 1990). However, these processes would leave little trace in the archaeological record. Evidence for the use of tars appears to pre-date the introduction of pottery by some millennia, so

questions remain as to the full range of production processes. Experiments simulating the preparation of tar without the use of pottery vessels have yet to meet with complete success (Czarnowski and Neubauer, 1991). As a footnote to the ubiquitous production and use of birch bark tar in prehistoric Europe the subject is not complete without reference to its familiar smell. Recently, in an investigation of the physiological response to various odours (birch tar, galbanum, heliotropine, jasmine, lavender, lemon, and peppermint) as measured by electroencephalography, birch tar was among the most consistently unpleasant smells recorded (Klemm *et al.*, 1992).

In the Neolithic of northern Europe, birch bark and other tars served many of the functions noted in the earlier section on resins: definition and uses. As adhesives, tars could be used to mend broken combs or pottery vessels (Sauter, 1967; Abb. 3), to reattach handles and to fix lithic implements to their hafts (Funke, 1969; Larsson, 1983). Some pottery vessels display possible adhesive around the rim, plausibly to fix a skin or lid over the top of the pot. In some areas of Switzerland and southern Germany, tar was used to attach cut out decorations of birch bark to the outside of pots (Vogt, 1949). Certain finds confirm that birch bark tar was chewed. Nine lumps of tar with human tooth impressions have been found at the Neolithic lake dwelling at Hornstaad-Hörnle I, in southern Germany (Rottländer, 1981; Schlichtherle and Wahlster, 1986; 92) and others are known from Mesolithic bog sites in Scandinavia (Larsson, 1983; 75–6). It is plausible that birch bark tar served a medicinal purpose. A more prosaic interpretation is that chewing the tar rendered it more ductile for use. Further uses of tars are documented much later in the Medieval period. Tar was also painted over doors or beds where it was thought to ward off evil spirits, perhaps due to its pungent smell. Rajewski (1970) observes that references to tar can be encountered in many proverbs of at least Medieval date in both rural and urban populations. This remarkable ethnohistoric record has its roots in much earlier populations of northern Europe.

Alternatives to Birch Bark and Softwood Tar

Whereas the range of uses of resin exudates and tar/pitch extractives has been emphasized, some of these functions could have been plausibly served by a range of other natural products. For example, knowledge of other prehistoric adhesives could be biased by the poor survival of protein-based (such as fish, bone, and horn) glues. However, it should be remembered that these substances are water-soluble and in many cases would not have been particularly useful. A single burnt residue on a

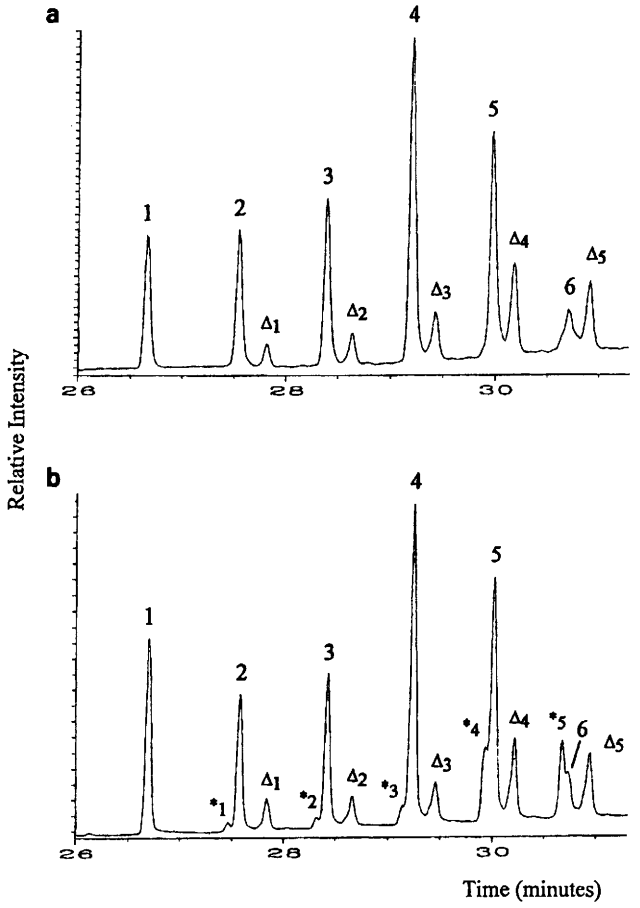
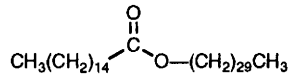


Figure 7.3 *Partial gas chromatograms comparing the wax ester distribution in (a) Neolithic sample from Ergolding Fischergasse with (b) authentic beeswax (*Apis mellifera*). Peak identities: Peaks 1–6 are wax esters in the range C_{40} (peak 1) to C_{50} (peak 6) comprising hexadecanoic (palmitic) acid esterified with alcohols of increasing chain length (C_{24} to C_{34}). Peaks Δ_1 to Δ_5 may represent co-elution of two molecular species one of which is a series of hydroxyesters. Peaks $*_1$ to $*_5$ are present only in the authentic sample and represent wax esters comprising an unsaturated (octadecenoyl) fatty acid moiety. Their absence in the ancient sample is not unexpected given the susceptibility of the double bond to oxidation or reduction reactions. For further details see Heron* et al. *(1994)* (Reproduced by permission of Springer-Verlag)



Structure 7.25 *Triacontyl hexadecanoate*

potsherd from the Neolithic site at Ergolding Fischergasse, Germany, has been identified as beeswax (Heron *et al.*, 1994). GC analysis showed the presence of wax esters in the range of 40–50 carbon atoms. The partial high temperature gas chromatograms displayed in Figure 7.3 compare the wax ester distribution in authentic beeswax (*A. mellifera*) with the lipid extract of the surface deposit coating the Neolithic sherd. Confirmation of the peak identities was provided by GC-MS. The principal wax esters comprise even-carbon number aliphatic chains of saturated alcohols and fatty carboxylic acids with total carbon numbers in the range from C_{40} to C_{52} with the C_{46} wax ester (triacontyl hexadecanoate [Structure 7.25]) the most abundant. Beeswax is produced by various *Apis* spp. The European honeybee (*Apis mellifera*) produces honeycombs of almost pure wax. The principal constituents are 14% odd-carbon-number n-alkanes with a range of 21–33 carbon atoms (with heptacosane $C_{27}H_{56}$ as the most abundant alkane); 35% C_{40}–C_{52} monoesters; 14% C_{56}–C_{66} diesters; 3% triesters; 4% C_{24}–C_{34} hydroxymonoesters; 8% hydroxypolyesters; 12% free acids; 1% C_{16}–C_{20} acid monoesters; 2% acid polyesters, and 7% unidentified material (Schulten *et al.*, 1987). In the Neolithic sample, the n-alkane fraction was severely depleted suggesting their combustion when the beeswax was burned. The sealing and waterproofing properties of beeswax would have made it a useful commodity in prehistoric Europe. The presence of wax on the inner surface of the potsherd suggests that a thin layer of hot wax may have sealed the permeable fabric of the vessel, enabling the storage of liquids. Alternatively, beeswax could have been stored in the vessel, awaiting use in a number of activities. The early occurrence of beeswax also implies the availability of honey to Neolithic communities in Europe. In the Early Bronze Age, the identification of an adhesive used to fix a knife blade to a double T-shaped handle (from Xanten-Wardt, Germany) as a bituminous substance, either earth wax or mineral pitch (Koller and Baumer, 1993) also suggests occasional use of other substances.

SUMMARY: EVIDENCE FOR OTHER ORGANIC SUBSTANCES

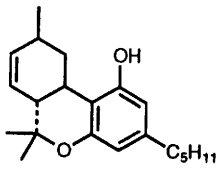
Aside from the characterization of birch bark tar, higher plant resins and their heated derivatives, and beeswax, what potential remains for the

identification of other organic substances used in prehistoric Europe? This short summary intends to whet the appetite by reviewing briefly recent investigations and considering possible future avenues for research. The association between pottery vessels and organic substances has been emphasized. Determining how pottery was actually used in the past is a key area for current and future research (Rice, 1987; 207–243). Archaeologists commonly use the form of reconstructed vessels to infer possible use. For example, in the early Neolithic of southern England, Thomas (1991; 89–90) uses size and shape criteria for distinguishing vessels used primarily for consumption, perhaps associated with feasting activity, and those used for storage and other domestic activities and notes the predominance of either one or the other in assemblages at certain site types. Such suggestions could benefit from chemical investigations aimed at defining pottery use (see Heron and Evershed (1993) for a review). However, as Hodder has lamented, few systematic studies have been carried out on Neolithic pottery (Hodder, 1990; 204) and this point has been reiterated in Vencl's review (Vencl, 1994) of the possible identification of liquids, such as milk, mead, beer, and wine in prehistoric Europe.

The situation is not irretrievable. At the Late Mesolithic site of Tybrind Vig in Denmark, stable isotope analysis of a burnt food deposit from a characteristic pointed-bottomed vessel of the Ertebølle culture was combined with identification of microscopic fragments in the char (Andersen and Malmros, 1984). Low $\delta^{13}C$ values suggested a terrestrial origin for the burnt food debris; a finding supported by charred remains of monocotyledonous plants, probably grass. However, bones, scales, and fin rays of cod (*Gadus morrhua*) were also identified. At Bradford, we have shown that a wide variety of lipid molecules survive in Neolithic pottery and that it is possible to identify both animal and plant substances (Bonfield and Heron, unpublished results). The most comprehensive study, employing GC and GC-MS analysis of lipid residues, albeit of Late Saxon/Early Medieval pottery of the 9th–11th Centuries AD from Raunds, Northamptonshire, UK, has characterized animal fats, including dairy products, and epicuticular leaf waxes of *Brassica* sp., from either cabbage or turnip (Evershed *et al.*, 1990; Heron *et al.*, 1991b; Evershed *et al.*, 1991; Evershed *et al.*, 1992). This has enabled correlations between lipid composition and pot type (Charters, 1993b).

The relationship between pottery vessels and foodstuffs may be a close one but there are other possibilities as this chapter has demonstrated with the association of resins surviving in pottery vessels. At Franchthi Cave in southern Greece, Vitelli (1993; 215) has considered that the very low number of Neolithic vessels produced at the site as well as their unusual shapes and absence of exterior sooting suggests that they were

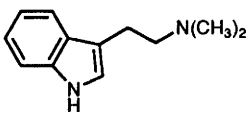
not used to prepare food. Rather she suggests that on account of the interior wear patterns on small saucers and large basins in the assemblage that they could have been used to burn incense, aromatic gums, and perhaps substances with psychoactive properties; the wear resulting from repeated scraping of the interior wall to remove residual contents. Such a hypothesis is in line with suggestions made by Sherratt (1991) for use of specific pottery vessel types in later Neolithic Europe for burning or infusing psychoactive substances, in particular opium (*Papaver somniferum*) and cannabis (*Cannabis sativa*). This is supported by botanical evidence from a number of sites. According to Sherratt (1991; 52), *'Any account of prehistoric Europe which omits a consideration of such substances is likely to be incomplete'*. Specific chemical evidence from the examples presented is again lacking although recognition of these substances in later contexts has been proposed. There is unambiguous archaeological evidence for burning cannabis seeds in the Iron Age of the Steppe region. The remarkable finds from the frozen tombs of Pazyryk (4th Century BC) in southern Siberia include paraphernalia for inhaling cannabis smoke – a fur bag filled with cannabis seed, a censer filled with stones, and a hexapod frame of an inhalation tent. Possible medicinal use of cannabis has been reported from the Middle East (Zias *et al.*, 1993). The remains of a female aged about 14 were found in a family burial tomb near Jerusalem. Coins recovered from the tomb suggest it was in use during the 4th Century AD. In the pelvic area of the girl lay the skeletal remains of a 40-week foetus. It was suggested that the girl may have died during pregnancy. In the abdominal area of the skeleton, nearly 7 g of a grey, carbonized material was recovered. Combined GC-MS and NMR spectroscopy confirmed preliminary suggestions that the substance was cannabis. The principal constituent was identified as Δ^6–Tetrahydrocannabinol (THC; [Structure 7.26]). Although a minor constituent of cannabis, its presence indicated the conversion of the major constituents such as Δ^1-THC and cannabidiol into Δ^6-THC; a conversion apparently known to occur during the burning process. The burnt residue was assumed to be cannabis administered to the girl as an inhalant to facilitate the birth process by alleviating pain.



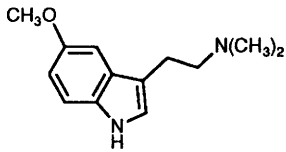
Structure 7.26 *Δ^6-Tetrahydrocannabinol (THC)*

Possible opium use has been associated with Bronze Age Cypriot Base-Ring I juglets which resemble in shape an upturned opium-poppy capsule's head and may have been used to hold opium for export to Egypt where these vessels are also found (Merrillees, 1962, Knapp, 1991; 25–26). Despite attempts and claims to the characterization of opium alkaloids in these vessels (the most recent being Merrillees with Evans, 1989), this observation remains contentious.

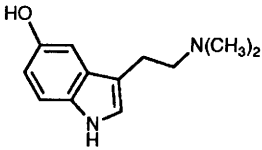
In other areas of the world, a number of recent investigations underline the potential for providing new evidence for the early use of such substances. For example, analysis of snuff samples from the grave of a mummified body in Northern Chile dating to around the end of the 8th century AD identified the psychoactive alkaloids dimethyltryptamine [Structure 7.27], 5-methoxydimethyltryptamine [Structure 7.28], and 5-hydroxy-*N*,*N*-dimethyltryptamine (bufotenine; [Structure 7.29]), suggesting the source of the snuff was a species of the genus *Anadenanthera* (Torres *et al.*, 1991). The snuff was part of kit which comprised wooden snuff trays, a snuffing tube, usually of wood or bone, a spoon or spatula and a mortar and pestle. Elsewhere, the alkaloids caffeine [Structure 7.30] and theobromine [Structure 7.31] have been identified using liquid chromatographic techniques in Mayan pottery vessels used to consume cocoa (Hurst *et al.*, 1989) and lactones congruent with the intoxicating drink kava (prepared from the roots of *Piper methysticum*) have been identified in residues adhering to archaeological ceramics from Fiji



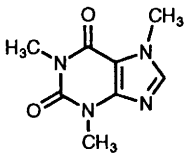
Structure 7.27
Dimethyltryptamine



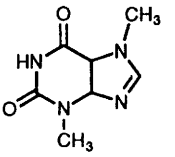
Structure 7.28
5-Methoxydimethyltryptamine



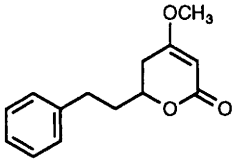
Structure 7.29
5-Hydroxy-N,N-*dimethyltryptamine*



Structure 7.30 *Caffeine*



Structure 7.31 *Theobromine*



Structure 7.32 *7,8-Dihydrokawain*

(Hocart *et al.*, 1993). A wide range of molecules were characterized including 7,8-dihydrokawain [Structure 7.32].

Clearly there is considerable potential for using archaeological chemistry for identification of a wide range of naturally occurring organic molecules used in the past. There are undoubted benefits to the contribution of our understanding of the relationship between human populations and their use of plant and animal resources as well as in the determination of the myriad ways in which artefacts were used.

REFERENCES

Anderson, K.B. and Winans, R.E. (1991). Nature and fate of natural resins in the geosphere I. Evaluation of pyrolysis-gas chromatography/mass spectrometry for the analysis of natural resins and resinites. *Analytical Chemistry* **63** 2901–2908.

Andersen S.H. and Malmros, C. (1984). 'Madskorpe' på Ertebøller fra Tybrind Vig. *Aaboger for Nordisk Oldkyndighed of Historie* 78–95.

Banthorpe, D.V. (1994). Terpenoids. In *Natural Products: Their Chemistry and Biological Significance*, eds. Mann, J., Davidson, R.S., Hobbs, J.B., Banthorpe, D.V. and Harborne, J.B., Longman, London, pp. 289–359.

Beck, C.W. and Borromeo, C. (1990). Ancient pine pitch: Technological perspectives from a Hellenistic shipwreck. In *Organic Contents of Ancient Vessels: Materials Analysis and Archaeological Investigation*, eds. Biers, W.R. and McGovern, P.E., MASCA Research Papers in Science and

Archaeology, Volume 7, The University Museum of Archaeology and Anthropology, University of Pennsylvania, Philadelphia, pp. 51–58.

Beck, C.W., Smart, C.J. and Ossenkop, D.J. (1989). Residues and linings in ancient Mediterranean transport amphoras. In *Archaeological Chemistry IV*, ed. Allen, R.O., Advances in Chemistry Series 220, American Chemical Society, Washington, pp. 369–80.

Beck, C.W., Stewart, D.R. and Stout, E.C. (1994). Analysis of naval stores from the Late Roman ship. In *Deep Water Archaeology: A Late Roman Ship from Carthage and an Ancient Trade Route Near Skerki Bank off Northwest Sicily*, eds. McCann, A.M. and Freed, J., Journal of Roman Archaeology, Supplementary Series No. 13, Michigan, Ann Arbor, pp. 109–121.

Beck, C.W., Stout, E.C. and Jánne, P.A. (in press). The pyrotechnology of pine tar and pitch inferred from quantitative analyses by gas chromatography/mass spectrometry and carbon-13 nuclear magnetic resonance spectroscopy. *Proceedings of the First International Symposium on Wood Tar and Pitch, Biskupin, Poland.*

Biers, W.R., Gerhardt, K.O. and Braniff, R.A. (1994). *Lost Scents: Investigations of Corinthian 'Plastic' Vases by Gas Chromatography-Mass Spectrometry*. MASCA Research Papers in Science and Archaeology Volume 11, University of Pennsylvania Museum of Archaeology and Anthropology, Philadelphia, Pennsylvania.

Binder, D., Bourgeois, G., Benoist, F. and Vitry, C. (1990). Identification de brai de bouleau (*Betula*) dans le Néolithique de Giribaldi (Nice, France) par la spectrométrie de masse. *Revue d'Archéométrie* **14** 37–42.

Bonfield, K.M., Heron, C. and Nemcek, N. (in press). The chemical characterization of wood tars in prehistoric Europe: A case study of the Neolithic of southern Germany. *Proceedings of the First International Symposium on Wood Tar and Pitch, Biskupin, Poland.*

Charters, S., Evershed, R.P., Goad, L.J., Heron, C. and Blinkhorn, P. (1993a). Identification of an adhesive used to repair a Roman jar. *Archaeometry* **35** 91–101.

Charters, S., Evershed, R.P., Goad, L.J., Leyden, A., Blinkhorn, P. and Denham, V. (1993b). Quantification and distribution of lipids in archaeological ceramics: Implications for sampling potsherds for organic residue analysis and the class of vessel use. *Archaeometry* **35** 211–223.

Clark, J.G.D. (1975). *The Earlier Stone Age Settlement of Scandinavia*. Cambridge University Press, Cambridge.

Cole, B.J.W., Bentley, M.D. and Hua, Y. (1991). Triterpenoid extractives in the outer bark of *Betula lenta* (black birch). *Holzforschung* **45** 265–268.

Czarnowski, E. and Neubauer, D. (1990). Aspekte zur produktion und verarbeitung von birkenpech. *Acta Praehistorica et Archaeologica* **23** 11–14.

Dev, S. (1989). Terpenoids. In *Natural Products of Woody Plants II: Chemicals Extraneous to the Lignocellulosic Cell Wall*, ed. Rowe, J.W., Springer-Verlag, Berlin, pp. 691–807.

Diallo, B., Vanhaelen, M. and Gosselain, O.P. (1995). Plant constituents involved in coating practices among traditional African potters. *Experientia* **51** 95–97.

Dimbleby, G.W. (1978). *Plants and Archaeology*. John Baker, London (2nd edn.).

Ekman, R. (1983). The suberin monomers and triterpenoids from the outer bark of *Betula verrucosa* Ehrh. *Holzforschung* **37** 205–211.

Evershed, R.P., Heron, C. and Goad, L.J. (1990). Analysis of organic residues of archaeological origin by high-temperature gas chromatography and gas chromatography/mass spectrometry. *Analyst* **115** 1339–1342.

Evershed, R.P., Heron, C. and Goad, L.J. (1991). Epicuticular wax components preserved in potsherds as chemical indicators of leafy vegetables in ancient diets. *Antiquity* **65** 540–544.

Evershed, R.P., Heron, C., Charters, S. and Goad, L.J. (1992). The survival of food residues: New methods of analysis, interpretation and application. In *New Developments in Archaeological Science*, ed. Pollard, A.M., Proceeding of the British Academy **77** 187–209.

Foster, G.M. (1956). Resin-coated pottery in the Philippines. *American Anthropologist* **58** 732–733.

Fox, A.F, Heron, C. and Sutton, M.Q. (1995). Characterization of natural products on Native American archaeological and ethnographic materials from the Great Basin region, USA: a preliminary study. *Archaeometry* **37** 363–375.

Fuchino, H., Sou, K., Imai, H., Wada, H. and Tanaka, N. (1994). A biodegradation product of betulin. *Chemical and Pharmaceutical Bulletin* **42** 379–81.

Funke, H. (1969). *Chemische-analytische Untersuchungen verschiedener archäologische Funde*. Dissertation Hamburg.

Gianno, R. (1990). *Semelai Culture and Resin Technology*. Memoirs of the Connecticut Academy of Arts and Sciences, Volume XXII, Newhaven, Connecticut.

Gianno, R., Erhardt, D., von Endt, D.W., Hopwood, W. and Baker, M.T. (1990). Archaeological resins from shipwrecks off the coasts of Saipan and Thailand. In *Organic Contents of Ancient Vessels: Materials Analysis and Archaeological Investigation*, eds. Biers, W.R. and McGovern,

P.E., MASCA Research Papers in Science and Archaeology, Volume 7, The University Museum of Archaeology and Anthropology, University of Pennsylvania, Philadelphia, pp. 59–67.

Grimalt, J.O., Simoneit, B.R.T. and Hatcher, P.G. (1989). Chemical affinities between the solvent extractable and bulk organic matter of fossil resin associated with an extinct podocarpaceae. *Phytochemistry* **28** 1167–1171.

Hadzi, D. and Orel, B. (1978). Spektrometricne raziskave izvora jantarja in smol iz prazgodovinskih najdisc na Slovenskem. *Vestnik Sloveneskega Kemitskega Drustva* **25** 51–62.

Hayek, E.W.H., Jordis, U., Moche, W. and Sauter, F. (1989). A bicentennial of betulin. *Phytochemistry* **28** 2229–2242.

Hayek, E.W.H., Krenmayr, P., Lohninger, H., Jordis, U., Moche, W. and Sauter, F. (1990). Identification of archaeological and recent wood tar pitches using gas chromatography/mass spectrometry and pattern recognition. *Analytical Chemistry* **62** 2038–2043.

Hayek, E.W.H., Krenmayr, P., Lohninger, H., Jordis, U., Sauter, F. and Moche, W. (1991). GC/MS and chemometrics in archaeometry; investigation of glue on copper age arrowheads. *Fresenius Journal of Analytical Chemistry* **340** 153–156.

Heron, C. and Evershed, R.P. (1993). The analysis of organic residues and the study of pottery use. In *Archaeological Method and Theory V*, ed. Schiffer, M.B., University of Arizona Press, Tucson, pp. 247–286.

Heron, C., Evershed, R.P., Chapman, B.C.G. and Pollard, A.M. (1991a). Glue, disinfectant and 'chewing gum' in prehistory. In *Archaeological Sciences, Bradford 1989*, eds. Budd, P., Chapman, B.C.G., Jackson, C., Janaway, R.C. and Ottaway, B.S., Oxbow Publications, Oxford, pp. 325–331.

Heron, C., Evershed. R.P. and Goad, L.J. (1991b). Effects of migration of soil lipids on organic residues associated with buried potsherds. *Journal of Archaeological Science* **18** 641–659.

Heron, C. and Pollard, A.M. (1988). The analysis of natural resinous materials from Roman amphoras. In *Science & Archaeology, Glasgow 1987*, eds. Slater, E.A. and Tate, J.O., Proceedings of a conference on the application of scientific methods to archaeology, Oxford, British Archaeological Reports, British Series 196, pp. 429–447.

Heron, C., Nemcek, N., Bonfield, K.M., Dixon, D. and Ottaway, B.S. (1994). The chemistry of Neolithic beeswax. *Naturwissenschaften* **81** 266–269.

Hocart, C.H., Fankhauser, B. and Buckle, D.W. (1993). Chemical archaeology of kava, a potent brew. *Rapid Communications in Mass Spectrometry* **7** 219–224.

Hodder, I. (1990). *The Domestication of Europe.* Blackwell, Oxford.

Hurst, J, Martin, R., Tarka, S. and Hall, G. (1989). Authentication of cocoa in Maya vessels using high performance liquid chromatographic techniques. *Journal of Chromatography* **466** 279–289.

Klemm, W.R., Lutes, S.D., Hendrix, D.V. and Warrenburg, S. (1992). Topographical EEG maps of human response to odors. *Chemical Senses* **17** 347–361.

Knapp, A.B. (1991). Spice, drugs, grain and grog: Organic goods in Bronze Age East Mediterranean trade. In *Bronze Age Trade in the Mediterranean,* ed. Gale, N.H., Paul Åströms Förlag, Jonsered, pp. 21–68.

Kurzweil, A. and Todtenhaupt, D. (1990). *Teer aus Holz,* Berlin: Museumsdorf Düppel.

Kurzweil, A. and Todtenhaupt, D. (1991). Technologie der Holzteergewinnung. *Acta Praehistorica et Archaeologica* **23** 63–79.

Koller, J. and Baumer, U. (1993). Analyse einer Kittprobe aus dem Griff des Messers von Xanten-Wardt. In Koschik, H., Messer aus dem Kies: Zu zwei Messern der jüngeren Bronzezeit aus dem Rhein bei Xanten-Wardt und aus der Weser bei Petershagen-Hävern. *Acta Praehistorica et Archaeologica* **25** 117–131.

Larsson, L. (1983). Ageröd V: An Atlantic Bog Site in Central Scania. *Acta Archaeologica Lundensia 12,* Lund, Sweden.

Mahato, S.B., Ashoke, K.N. and Roy, G. (1992). Triterpenoids. *Phytochemistry* **31** 2199–2249.

Mann, J., Davidson, R.S., Hobbs, J.B., Banthorpe, D.V. and Harborne, J.B. (eds.) (1994). *Natural Products: Their Chemistry and Biological Significance.* Longman, London.

Merrillees, R.S. (1962). Opium trade in the Bronze Age Levant. *Antiquity* **36** 287–292.

Merrillees, R.S. with Evans, J. (1989). Highs and lows in the Holy Land: opium in Biblical times. In *Yigael Yadin Memorial Volume,* eds. Ben-Tor, A., Greenfield, J. and Malamat, A., Eretz Israel 20, Israel Exploration Society, Hebrew University, Jerusalem, pp. 148–154.

Mills, J.S. and White, R. (1977). Natural resins of art and archaeology. *Studies in Conservation* **22** 12–31.

Mills, J.S. and White, R. (1989). The identity of resins from the Late Bronze Age shipwreck at Ulu Burun (Kaş). *Archaeometry* **31** 37–44.

Mills, J.S. and White, R. (1994). *The Organic Chemistry of Museum Objects,* Butterworth–Heinemann, London (2nd edn.).

Mills, J.S., White, R. and Gough, L.J. (1984/85). The chemical composition of Baltic amber. *Chemistry and Geology* **47** 15–39.

O'Connell, M.M., Bentley, M.D., Campbell, C.S., Cole, B.J.W. (1988).

Betulin and lupeol in bark from four white-barked birches. *Phytochemistry* **27** 2175–2176.

Ottaway, B.S. (1995). *Ergolding Fischergasse: Eine feuchtbodensiedlung der Altheimer kultur in Niederbayern.* Verlag Michael Laßleben, Kallmünz.

Pokhilo, N.D., Makhenev, A.K., Demenkova, L.I. and Uvarova, N.I. (1990). Composition of the triterpene fraction of outer bark extracts of *Betula pendula* and *Betula pubescens* (in Russian). *Khimiya Drevesiny* **6** 74–77.

Proefke, M.L. and Rinehart, K.L. (1992). Analysis of an Egyptian mummy resin by mass spectrometry. *Journal of the American Mass Spectrometry Society* **3** 582–589.

Rajewski, Z. (1970). Pech und teer bei den Slawen. *Zeitschrift für Archäologie* **4** 46–53.

Reunanen, M., Ekman, R. and Heinonen, M. (1989). Analysis of Finnish pine tar and tar from the wreck of frigate St. Nikolai. *Holzforschung* **43** 33–39.

Reunanen, M., Holmbom, B. and Edgren, T. (1993). Analysis of archaeological birch bark pitches. *Holzforschung* **47** 175–177.

Rice, P.M. (1987). *Pottery Analysis: a sourcebook.* University of Chicago Press, Chicago.

Robinson, N., Evershed, R.P., Higgs, W.J., Jerman, K. and Eglinton, G. (1987). Proof of a pine wood origin for pitch from Tudor (Mary Rose) and Etruscan shipwrecks: Application of analytical organic chemistry in archaeology. *Analyst* **112** 637–644.

Rottländer, R.C.A. (1981). A Neolithic 'chewing gum'. Poster presentation, 21st Archaeometry Symposium, Brookhaven, USA (abstract only).

Ruthenberg, K. (in press). Historical development and comparison of analytical methods for the identification of tar and pitch. *Proceedings of the First International Symposium on Wood Tar and Pitch, Biskupin, Poland.*

Sandermann, W. (1965). Untersuchung vorgeschichtlicher 'Gräberharze' und Kitte. *Technische Beiträge zur Archäologie* **2** 58–73.

Sauter, F. (1967). Chemische untersuchung von 'Harzüberzügen' auf hallstattzeitlicher keramik. *Archaeologia Austriaca* **41** 25–36.

Sauter, F., Jordis, U. and Hayek, E. (1992). Chemsiche untersunchungen der Kittschäftungs-materialien. In *Der Mann im Eis, Band 1, Bericht über das Internationale Symposium 1992 in Innsbruck*, eds. Höpfel, F., Platzer, W. and Spindler, K., Eigenverlag der Universität Innsbruck, Innsbruck, pp. 435–441.

Schiffer, M.B. (1990). The influence of surface treatment on heating effectiveness of ceramic vessels. *Journal of Archaeological Science* **17** 373–381.

Schlichtherle, H. and Wahlster, B. (1986). *Archäologie in Seen und Mooren – Den Pfahlbauten auf der Spur.* Konrad Theiss Verlag, Stuttgart.

Schulten, H.-R., Murray, K.E. and Simmleit, N. (1987). Natural waxes investigated by soft ionization mass spectrometry. *Zeitschrift für Naturforschung C: Biosciences* **42** 178–190.

Sherratt, A. (1991). Sacred and profane substances: The ritual use of narcotics in Later Neolithic Europe. In *Sacred and Profane*, eds. Garwood, P., Jennings, D., Skeates, R. and Toms, J., Oxford Committee for Archaeology, Oxford, pp. 50–64.

Simoneit, B.R.T., Grimalt, J.O., Wang, T.G., Cox, R.E., Hatcher, P.G. and Nissenbaum, A. (1986). Cyclic terpenoids of contemporary resinous plant detritus and of fossil woods, ambers and coals. *Organic Geochemistry* **10** 877–889.

Spindler, K. (1993). *The Man in the Ice.* Weidenfeld and Nicholson, London.

ten Haven, T.L., Peakman, T.M. and Rullkötter, J. (1992). Δ^2–Triterpenes: Early intermediates in the diagenesis of terrigenous triterpenoids. *Geochimica et Cosmochimica Acta* **56** 1993–2000.

Thomas, J. (1991). *Rethinking the Neolithic.* Cambridge University Press, Cambridge.

Torres, C.M., Repke, D.B., Chan, K., McKenna, A.L. and Schultes, R.E. (1991). Snuff powders from Pre-Hispanic San Pedro de Atacama: Chemical and contextual analysis. *Current Anthropology* **32** 640–649.

Ukkonen, K. and Erä, V. (1979). Birch-bark extractives. *Kemia-Kemi* **6** 217–220.

Vencl, S. (1994). The archaeology of thirst. *Journal of European Archaeology* **2** 299–326.

Vitelli, K.D. (1993). *Franchthi Neolithic Pottery: Volume 1: Classification and Ceramic Phases 1 and 2.* Indiana University Press, Bloomington.

Vogt, E. (1949). The birch as a source of raw material during the Stone Age. *Proceedings of the Prehistoric Society* **5** 50–51.

White, R. (1992). A brief introduction to the chemistry of natural products in archaeology. In *Organic Residues in Archaeology: Their Analysis and Identification*, eds. White, R. and Page, H., UK Institute for Conservation Archaeology Section, London, pp. 5–10.

Zias, J., Stark, H., Seligman, J., Levy, R., Werker, E., Breuer, A. and Mechoulam, R. (1993). Early medical use of cannabis. *Nature* **363** 215.

Chapter 8

Amino Acid Stereochemistry and the First Americans

INTRODUCTION

In 1989, D.J. Meltzer published an article in the journal *American Antiquity* entitled 'Why don't we know when the first people came to North America?' (Meltzer, 1989). He noted that, despite over one hundred years of intense debate, the question had failed to be resolved. There are two polarized positions which are so far apart that the only point of agreement is that humans first entered North America from Siberia. One school maintains that this took place at the end of the last glaciation (sometime between 14 000 and 12 000 uncalibrated radiocarbon (^{14}C) years before present – signified as BP), and that the vast Continent was empty (in human terms) at this time. The second school, citing a number of sites with radiocarbon dates which appear to signify human habitation before this time, insist that the colonization took place some time before 15 000 BP, possibly as early as 35 000 BP. The 'late' school reject the various sites cited as 'early' on grounds such as poor dating, or poor evidence of human occupation. They point to the apparent explosion of human activity starting some time between 11 500 and 11 000 BP, characterized by stone tools known as Clovis fluted projectile points, and interpret this as evidence for the spread of a vigorous population of efficient large mammal hunters – the 'Clovis folk' – who spread from North to South America in a few thousand years, and gave rise to 'the Clovis horizon' in American prehistory (Fagan, 1991; Dillehay and Meltzer, 1991). This short chronology model is also viewed favourably by those palaeontologists seeking to explain the sudden demise of American megafauna (mammoths, camelids, sloths, *etc.*) at approximately the same time, on the grounds that human 'overkill' is a simple explanation supported by historical observations of similar extinction

events as humans have colonized unpopulated islands (*e.g.*, the extinction of the New Zealand Moa after the arrival of the Maori about 1000 years ago: Martin, 1984). Supporters of the 'long' chronology dispute this 'blitzkrieg' model of human expansion and megafaunal extinction, and point to a handful of sites scattered throughout the Continent which they say yield substantial evidence for pre-Clovis occupation – sites referred to by J.M. Adovasio (one of the major supporters of an early date for the migration) as 'the ones that will not go away' (Adovasio, 1993).

Much of the debate centres around the exact dating of the archaeological sites, and the quality of the association between the evidence for human occupation and the material dated (Haynes, 1992). The backbone for the chronology of all late Quaternary geology and archaeology (back to around 30–40 000 BP) is radiocarbon dating (Bowman, 1990). In the early days, radiocarbon required relatively large samples (either of charcoal, wood, or bone – in the latter case, up to 100 g was required) from which to purify the carbon and measure the residual activity of the ^{14}C atoms. Developments in nuclear accelerator physics during the 1970s resulted in the availability of an alternative mass spectrometric technique for radiocarbon dating, known as *accelerator mass spectrometry* (AMS: Aitken, 1990). One of the major advantages of AMS ^{14}C dating is that the size of the sample is reduced from hundreds of grams to a few milligrams, making it possible to date unique objects (such as the Turin shroud), as well as human skeletal remains, without their total destruction. This has the twofold advantage that rare human remains can be dated directly (using the older methods, such material was often dated by contextual association with something less valuable, such as charcoal), and that the chemical purification process can be carried out much more rigorously, to reduce the chances of error due to contamination.

Before AMS ^{14}C dating in the 1980s, there was always some room to doubt the quality of radiocarbon dates produced on human bone, either on the grounds of poor association between the bone and the material actually dated (if this was not bone), or, if the bone itself was used, because of the actual procedure employed. In the early stages of the debate, many of the dates were produced from the 'whole bone' (mineral plus organic phase), which may include secondary carbonates from the mineral fraction of the bone, which are therefore liable to be heavily contaminated by circulating groundwaters during burial. The announcement, therefore, in the early 1970s of another technique, apparently independent of radiocarbon, which dated bone directly using minute samples, was greeted with great enthusiasm by archaeologists, especially in North America, since it seemed that the long-running colonization debate could be solved once and for all by dating directly a number of

the putatively early Paleoindian bones. This technique is known as *amino acid racemization dating* (AAR), which relies on the fact that the amino acids which constitute bone protein are formed *in vivo* as the left-handed (L) enantiomer (see below), but, after death, racemize slowly to produce measurable quantities of the right handed (D) form (Masters, 1986a). One particular molecule – aspartic acid – was singled out for extensive study, and finally apparently yielded conclusive evidence of the antiquity of humans in North America when it was announced that some of the Californian Paleoindian bones were as old as 70 000 years BP (Bada *et al.*, 1974). For a moment, it seemed as if the debate had been settled in favour of the longer chronology.

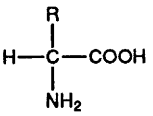
The sceptics were, however, still unconvinced – partly because the skeletons that were given such old dates didn't *look* that old. Morphologically, they were deemed to be identical to modern Native American skeletons, and, by analogy with Europe, skeletons which were that old should show some archaic traits. In Europe, skulls dating before about 40 000 BP show features such as brow ridges, which are associated with Neanderthal populations. This debate coincided with the first archaeological ^{14}C measurements using the new accelerator technique, and it was subsequently confirmed that an error had been made, and that none of the bones were as old as had been suggested . In fact, none appeared to be older than mid-Holocene (*i.e.*, around 5–6000 BP: Taylor *et al.*, 1985). As a result of this, doubt was inevitably cast on the validity of the AAR method, but subsequent investigations showed that this was not the prime cause of the problem. Because of the relationship between ambient temperature and the rate of a chemical reaction (in this case, the racemization process), it had been decided to determine the racemization reaction constant using the measured D/L aspartic acid ratio in a bone of known date – known, that is, from radiocarbon dating. It turns out that this calibration date was in serious error, and, if the racemization measurements are re-calculated using the corrected date, the ages predicted by AAR are reasonably in line with those produced by other techniques (Bada, 1985). Nevertheless, this cameo within the wider debate on the origins of North American Paleoindians has had a lasting effect on the reputation of the applications of amino acid stereochemistry in archaeology. As a dating technique, it has largely been superseded by AMS ^{14}C dating, which, for the last 35 000 years at least, now offers a more established technique on similar sized samples. The measurement of enantiomeric ratios of various amino acids is now, however, being considered as an indicator of age at death in forensic and archaeological investigations, since not all tissues maintain a uniformly asymmetric ratio *in vivo* (Gillard *et al.*, 1990). Measurement of amino acid racemization in

the protein fraction of mollusc shells is also widely used to give stratigraphic relationships and dating evidence in quaternary geochronology (Bada, 1991).There is still considerable scientific interest in the phenomenon of racemization of amino acids in proteins contained in calcified tissue, and how this process is affected by interaction with a burial environment over geological time periods. In this chapter we give a brief introduction to the structure of mammalian bone protein, and to the process of amino acid racemization. The factors surrounding the Californian Paleoindian controversy are reviewed and brought up to date, followed by a summary of some of the current uses of amino acid stereochemistry in archaeological science.

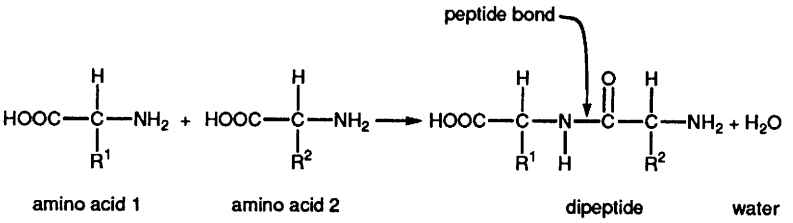
THE STRUCTURE OF BONE COLLAGEN

Bone is a natural composite material, with an average composition for dry compact human bone of 70% (by weight) insoluble inorganic matter, 20% organic matter, and 10% water (*i.e.*, water lost below 105 °C). It has a structure consisting of a fibrous organic matrix (collagen) which contains within it the finely crystalline inorganic mineral phase. The principal components of the inorganic phase are calcium and phosphate ions. It used to be described as poorly mineralized calcium hydroxyapatite ($Ca_{10}(PO_4)_6(OH)_2$), but more recent work using transmission electron microscopy has revealed that the mineral takes the form of thin plates or tablets, with a length and breadth of a few hundred Ångstrom units but a thickness of only 20–30 Å (Weiner and Traub, 1992). (The ångstrom is a non-SI unit corresponding to 10^{-10} m, but it is still widely used in mineral chemistry. The 'correct' units are nm (10^{-9} m), giving dimensions of 2–3 nm in this case). The mineral phase itself is now best described as *dahllite*, a carbonate hydroxyapatite mineral with an approximate stoichiometry $Ca_5(OH)\{[PO_4]_{0.5-0.9}[(CO_3)(OH)]_{0.1-0.5}\}_3$. These mineral plates are embedded in the collagen fibres (see below) in a highly ordered way. As a living tissue, human bone is constantly being destroyed and reformed and contains a number of specialized cells responsible for the production of bone tissue (osteoblasts), the maintenance of tissue (osteocytes), and the removal of bone (osteoclasts). Two types of bone structure are distinguished – *cancellous* or spongy bone and compact or *cortical* bone. Cancellous bone is characterized by a porous structure, consisting of a network of trabeculae. The distribution of cortical and cancellous bone throughout the skeleton is governed largely by biomechanical considerations, which require the unique combination of lightness and strength that is a characteristic of bone (McLean and Urist, 1968).

Approximately 90% (by weight) of the organic fraction in bone is made up of a fibrous structural protein of the *collagen* family. The remaining 10% are either other proteins (collectively termed '*non-collagenous proteins*') or various lipids (fats). Collagens of several types are widely distributed throughout the connective tissue of the body, and are largely responsible for the strength and elasticity of such tissues. Proteins are biopolymers consisting of one or more chains of amino acids, linked together by peptide bonds. Amino acids are a family of organic compounds with the general formula:



where R is one of a number of organic groups, termed *radicals* – the simplest being the hydrogen atom (H), making the amino acid glycine (gly) which is the smallest member of the amino acid family. The most important aspect of the chemistry of the amino acids as a family is that they can link together to form long chains via the peptide bond. A peptide bond is formed when the amine (NH_2) radical of one amino acid links to the acid (COOH) radical of the next, with the elimination of a water molecule:



These elimination reactions can continue until a protein chain has been formed containing many hundreds of amino acids all linked together – each amino acid in the chain is termed an amino acid *residue*. The protein that is formed is characterized by the specific sequence of amino acids, labelling from the nitrogen-containing radical end of the chain. There are 23 'natural' amino acids found in animal proteins, but collagens typically only contain about 17 of these – a full list of the structure of these amino acids can be found in most text books on organic chemistry or biochemistry, such as Morrison and Boyd (1983; Table 30.1). The dominant sequence found in collagens is the repeated unit:

–gly–X–Y–gly–X–Y–

where gly signifies the amino acid glycine, and X and Y are any of the other amino acids found in collagen. X and Y are often the amino acids proline (pro) and hydroxyproline (hypro). This results in an average composition in most kinds of human collagen (by number of residues) of 33–35% glycine, 7–13% proline, and 9–13% hydroxyproline, with the other amino acids making up less than half of the total residues. Hydroxyproline is an unusual constituent in that it only appears to be a significant component of collagen and does not generally occur in other proteins. This fact has relevance for radiocarbon dating when discussing the purity of potentially contaminated collagens. Aspartic acid (where R = CH_2COOH) is the amino acid most commonly used in archaeological racemization studies. Normal bone collagen contains about 4–5% of aspartic acid residues by number (Miller and Gay, 1982; Table 3).

Collagen is made up of a rope-like structure consisting of three of these polypeptide chains, twisted together in a right hand helix. Each individual amino acid chain is twisted to the left, one turn per three residues (thus aligning the glycine molecules above each other at every third residue), with ten turns of each chain per turn of the triple helix. There are seven or eight common sequences of amino acid chains in collagens, and collagens in different tissues are made up of different combinations. The most common collagen (Type I) makes up approximately 90% of body collagen, and occurs in bone, tendon, dentine, cornea, soft tissue, and scar tissue. It is made up of two chains of the same type [labelled $\alpha 1(I)$], plus one different chain [$\alpha 2(I)$], and is therefore described as $\alpha 1(I)_2\alpha 2(I)$. The majority of the rest of the collagens found in human tissue are made up of three identical chains, such as Type III, occurring in blood vessel walls, which consists of $\alpha 1(III)_3$ (Miller and Gay, 1982).

The individual collagen 'ropes' are called fibrils, and have an average molecular weight of around 300 000 Daltons (atomic mass units), with a length of 260 nm and a diameter of 1.4–2.0 nm. In the collagen fibres which can be seen in bone, the individual fibrils are aligned head to tail, with a gap of 40 nm between fibrils, and a stagger of 65 nm (one quarter of the length of the molecule) between adjacent rows. This gives rise to the characteristic 65 nm banding that is visible in electron microscope photographs of collagen. Mature collagen is insoluble in water because of the covalent cross-links between adjacent polypeptides in the 'rope' helix. Solubility increases as the protein is denatured – *i.e.*, as the overall molecular weight is reduced.

As noted above, the other organic components of bone are often grouped together as non-collagenous proteins (ncp). Common in tooth

dentine (but not in bone) are phosphoproteins, with an unusual amino acid composition of 50% serine and 40% aspartic acid, and a total phosphorus content of 26 weight percent. Other proteins include osteocalcin in which the glutamic acid side-chain has been carboxylated. These so-called *gla proteins* are the major component of the ncp fraction in bone. The lipid component in dentine and bone makes up only about 0.1 wt% of the tissue: in bone, three-quarters of this is triglyceride (triacylglycerol), with the rest being predominantly cholesterol (Williams and Elliott, 1989; 366).

STEREOCHEMISTRY OF AMINO ACIDS

One of the interesting properties of amino acids (except glycine, the simplest member of the family) is known as *chirality*, which gives rise to the subject of *stereochemistry*. In the general formula for the amino acids given above, the molecule is drawn as planar, with a central carbon atom having four different chemical groups (radicals) attached to the bonds, each of which projects at right angles to its neighbour. In reality, the molecule is not flat, and the carbon atom is at the centre of a three-dimensional molecule in which each of the attached radicals projects outwards in a tetrahedral structure, as shown in Figure 8.1. Because each radical is different, there are two chemically identical but structurally different isomers of this molecule, related to each other in the same way as an object and its mirror image (Figure 8.1). Just as our own left and right hands are (in principle) identical but different (in that they cannot be superimposed on each other by rotation), these two molecules can be regarded as the left and the right handed forms of the

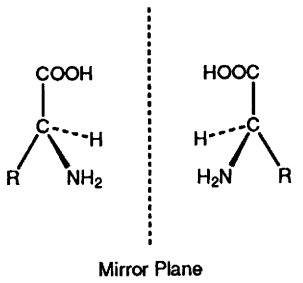


Figure 8.1 *Stereo pair of amino acids*

amino acid (termed *enantiomorphs*). They cannot be distinguished from each other using simple tests such as melting point or refractive index, but they do differ in their power (when in aqueous solution) to rotate plane polarized light – a property known as *optical activity*, and measured by *polarimetry*. In the older chemical notation for these molecules, the two enantiomorphs are termed D-and L- optical isomers, because a solution of one isomer in water will rotate plane polarized light to the right (the D-, or dextro-rotatory form), and the other to the left (the L-, or laevo-rotatory form). The nomenclature D- and L- are termed *relative configurations*. Strictly speaking, we should refer to (+) and (−) instead of D- and L-, since plus and minus refer directly to the direction of rotation of plane polarized light. It was Fischer in the 19th Century who arbitrarily gave the notation D- to that isomer of glyceraldehyde which gave rise to positive (right-handed) rotation. All compounds which could be converted into this particular enantiomorph by degradation or synthesis were termed 'D' compounds (Dawber and Moore, 1980; 51).

Unfortunately, it is not possible to predict from the molecular structure which way plane polarized light will be rotated. Worse, it is sometimes possible to alter the direction of rotation of plane polarized light by changing the structure of the molecule away from the chiral centre (in the case of amino acids, the central carbon atom), without actually altering the bonding at this centre. To deal with this, the modern (*absolute*) notation is systematically based on the molecular configuration, and is known as the *Cahn-Ingold-Prelog* convention. To use this, the four radicals attached to the chiral carbon atom are labelled a, b, c, and d in decreasing order of atomic weight (*i.e.*, the heaviest group is labelled a, in sequence to the lightest which is d). The molecule is then imagined as if the lightest group (d) is pointing towards the observer, and the orientation of the remaining groups is noted. If it has the sequence a–b–c in a clockwise fashion, the absolute configuration of the molecule is *R*. If it is anti-clockwise, it is *S*. For most naturally-occurring amino acids in proteins, the absolute configuration of L-amino acids is usually *S* (Morrison and Boyd, 1983; 343). Although absolute configurations are becoming the norm in organic chemistry, the majority of the literature on amino acid stereochemistry in archaeology, geology, and forensic sciences still uses the relative terminology, and we have therefore retained that convention for the remaining discussion.

Further complications may arise with the larger amino acids such as isoleucine, where the R side-chain itself contains a chiral carbon atom ($R = CH_3CH_2C^*H(CH_3)$, where the asterisk denotes the second chiral centre). This molecule is an example of a *diastereomer* – a molecule with

more than one chiral centre. Diastereomers have different physical and chemical properties, and their interconversion is more complicated, and is termed *epimerization.*

RACEMIZATION OF AMINO ACIDS

Racemization is the process by which one enantiomer converts to the other. In normal circumstances, organic compounds which may exhibit optical activity (such as simple sugars) exist in solution as equal numbers of D- and L- forms. The overall optical activity of such a solution is zero, since equal but opposite rotations cancel out, and the solution is termed *racemic.* It is in fact in a state of dynamic equilibrium, with the rate of conversion (racemization) of the D- to L- form the same as the conversion of the L- to D- form – thus there is no change overall. The situation in biochemical reactions is quite different, particularly in relation to the amino acids which make up proteins such as collagen. Biochemical syntheses (*e.g.*, the production of biomolecules by the human body) are often asymmetric, in that they favour the production of one optical isomer over the other. In fact, the manufacture of amino acids for bone collagen is so asymmetric that *in vivo* virtually all of the amino acids making up the molecules are only the L- form: no D- amino acids should be present at all. The physiology of living bone also ensures that this asymmetry is maintained throughout life, since bone is constantly being resorbed and remodelled by cell activity, which results in the destruction of any proteins containing D-amino acids which may be formed.

At death, physiological control ceases, and the bone can be thought of as making the transition from a biochemically to a chemically (or geochemically) controlled system. Under these conditions, the equilibrium condition of the amino acids is the racemic mixture, and in theory the process of racemization begins to increase the number of D- amino acids present in the collagen. Thus the percentage of D- amino acids will increase gradually until the ratio of D/L is equal to one, at which point the system is again in dynamic chemical equilibrium.

The true situation is more complex than this, in that amino acids within the protein chain might be expected to racemize at a different rate to terminal amino acids, which in turn will be different to free amino acids (*i.e.*, amino acids existing as discrete molecules). The actual mechanism of racemization of free amino acids is thought to be via the formation of an unstable planar intermediate – the proton (hydrogen ion) attached to the chiral carbon atom at the centre of the molecule can be removed in aqueous solution, forming a *carbanion* intermediate (*i.e.*, one with a negative charge on the carbon) in which the other three

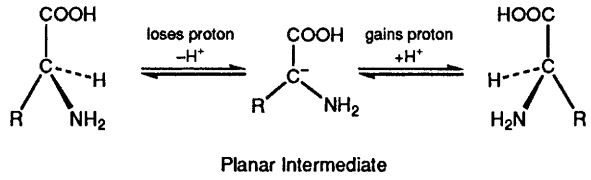


Figure 8.2 *Schematic postulated mechanism of racemization*

radicals temporarily re-orientate themselves into a triangular configuration which is planar (see Figure 8.2). This structure is unstable in water, where there are plenty of protons available to reconstitute the original molecule. If the new proton approaches from the same side as that from which the original one was lost, then the amino acid will reconfigure in the same enantiomorphic form, with the net result being that nothing has changed. If, however, a proton happens to approach from the other side, then the other enantiomer will result – effectively the molecule will have racemized (Smith and Evans, 1980). Most of the studies of racemization of amino acids have been carried out on free amino acids dissolved in water, which is not the situation which pertains in bone protein. It is well-known that one of the effects of joining amino acids together via peptide bonds is to slow down the rate that a particular amino acid will racemize compared to its free form. Moreover, research has shown that the position within the chain can be important: terminal amino acids (*i.e.*, those situated at the end of the peptide chain) racemize at a different rate to the same amino acid in an 'interior' position, although the evidence is as yet equivocal about which is faster (Smith and Evans, 1980; 265). The factor which affects the rate is largely geometrical, and is called *steric hindrance* – molecules in some chemical bonding environments are effectively more free to move than the same molecule in a different environment. This can be most easily demonstrated using molecular models. As discussed further below, the process of racemization in archaeological and geological bone is therefore complicated, and is still relatively poorly understood.

Because of these factors, it is very difficult to predict the rates of racemization of the various amino acids in a particular chemical environment. Furthermore, because racemization is a chemical reaction (unlike radioactive decay, which is a nuclear process), the rate is also

affected by environmental conditions, such as temperature, water availability and pH, chemical complexation, and the presence of other organic and inorganic species in solution. As a general rule of thumb, it is expected that the rate of a chemical reaction will double for every 10 °C increase in temperature, showing just how sensitive these processes are to environmental factors. Certain generalizations are possible which will cover most circumstances encountered when using racemization as a dating or ageing technique: for example, it has been found that aspartic acid derived from collagen racemizes roughly twice as fast as alanine, which in turn is twice as fast as leucine (Smith and Evans, 1980; 274). Bada and Shou (1980; 252) have reported that the rate of racemization of aspartic acid in 'typical uncontaminated fossil bone' is nearly twice that of free aspartic acid at 142 °C, and that in fossil bone the sequence of rate constants is:

$$\text{aspartic acid} >> \text{glutamic} \approx \text{acid phenylalanine} > \text{alanine} > \text{leucine} \approx \text{isoleucine} \approx \text{valine} > \text{proline}.$$

There is now considerable interest in the relationship between the preservation of collagen and the rate of racemization of its constituent amino acids, and further data may be expected to clarify these problems.

Because at ambient temperatures the racemization rates of all amino acids are slow, it is usually found that aspartic acid is most useful archaeologically, but over much longer geological timescales aspartic acid may become racemic (and therefore useless as a chronological indicator), and the slower rate of leucine becomes more useful.

AMINO ACID RACEMIZATION DATING OF THE CALIFORNIAN PALEOINDIANS

The phenomenon of amino acid racemization came to the attention of the archaeological world, especially in North America, with the publication of the first amino acid racemization dates on some of the Paleoindian skeletons from California in 1974 (Bada *et al.*, 1974), although this was not the first published work using AAR of relevance to archaeology. In the previous year, Schroeder and Bada (1973) had used the measured aspartic acid enantiomeric ratio in radiocarbon dated bones from Spain, Kenya, and Olduvai Gorge, Tanzania to estimate that the average temperature rise between the glacial and the post-glacial period in the Mediterranean had been 4 °C, and 5–6 °C for East Africa. In the same year, Bada and Protsch (1973) published age estimates for fossil bones from Olduvai Gorge, giving ages of between 5000 and 60–70 000 BP for

a range of bones. Their method used 10 g of bone, as compared to the 412 g required for the radiocarbon date obtained on the same bone, and, of course, yielded dates which were older than could have been measured using radiocarbon (which then had a time depth limit of about 30–40 000 years, because of the decreasing amount of ^{14}C present with time). They used an automatic amino acid analyser (*i.e.*, a liquid chromatographic method), after first hydrolysing the demineralized protein to give a solution of amino acids, and separating out the aspartic acid component by ion-exchange chromatography. The enantiomers were separated on the amino acid analyser after derivatizing with L-leucine-*N*-carboxyanhydride. Bones from the middle section of the beds at Olduvai were used to calibrate the rate of racemization for the site, which gave a radiocarbon date of 17 550 ± 1000 BP (UCLA–1695). Subsequent AAR measurements for aspartic acid were done using gas chromatography after converting the purified amino acid into the volatile *N*-trifluoroacetyl-(+)-2-butyl ester (Bada *et al.*, 1973).

This work demonstrated that AAR could give reasonable dates from smaller samples of bone than were necessary for radiocarbon, and had a time depth of at least 70 000 years, and possibly more if one of the more slowly racemizing amino acids such as alanine was used. The key paper came in 1974 (Bada *et al.*, 1974), which published dates of between ~ 6000 and 48 000 BP for various samples of human bone from the Californian coast (Table 8.1). The SDM (San Diego Museum) samples from site W-2 were from a shell midden near La Jolla excavated in 1926. Subsequently, it appears that 19 individual burials were recovered in a rescue operation from this site, known as La Jolla Shores: SDM-16755 is thought to refer to more than one individual (La Jolla Shores I and II), with a third (La Jolla Shores III) identified as SDM-16740 (Taylor *et al.*, 1985; Table 1). Site W-34 was located between Del Mar and Solano Beach, from a shell midden which had been largely destroyed by coastal erosion. Details of the other samples are given in an earlier publication (Berger *et al.*, 1971), where the relevant radiocarbon dates had already been published. The so-called Los Angeles Man was found during excavation work in 1936 in a location north of the Baldwin Hills: subsequently, the remains of a mammoth were found at the same depth, but 370 m away from the skull. This gave rise to speculation that the skull might be of a late Pleistocene date. Amino acids from the collagen extracted from 100 g of skull were radiocarbon dated at UCLA, giving a date of >23 600 BP (UCLA-1430). Despite the relatively large sample, there was insufficient carbon extracted to give a finite date. The Laguna Skull was the first Paleoindian skull to be found in California, and was discovered in 1933 during road building operations at 255 St. Ann's

Table 8.1 *Aspartic acid racemization dates from Californian Paleoindian skeletons*
(Reprinted with permission from Bada *et al.*, 1974; Tables 1 and 2. Copyright 1974 American Association for the Advancement of Science)

Sample	*Description and location*	D/L *aspartic acid*	*Predicted age (yr)*
Laguna Skull	Skull and long bones found at Laguna Beach in 1933	0.25	(used as calibration)
Los Angeles Man	Skull fragment from north of Baldwin Hills, found 1936	0.35	26 000
SDM-18402	Long bones from base of shell midden site W-2	0.16	~6000
SDM-16755	Rib and fragments in fill of site W-2, found 1926	0.36	28 000
SDM-16742	Human frontal found in white sand at W-2, 1926	0.50	44 000
SDM-16704 (femur)	Skull and mandible, long bones and scapula fragments from lower midden W-34–A, 1929	0.53	48 000

Drive, Laguna Beach. It was reported to be lying alongside some long bone fragments. A purified organic sample from 78.5 g of the skull was dated by radiocarbon to give a date of $17\,150 \pm 1470$ years BP (UCLA-1233A). A smaller sample (23 g) of the associated long bone fragments gave a minimum age of >14 800 years (UCLA-1233B), supporting the Pleistocene age of the Laguna human. The association between the skull and the long bones was confirmed by measurements of their fluorine, nitrogen, and uranium contents (the measurement of which had elsewhere finally revealed the Piltdown hoax in Britain), which were found to be closely similar. Given that these were radiocarbon dates carried out to the best standards of the day, there appeared to be no doubt on the radiocarbon evidence alone that humans were present in California during the late Pleistocene.

Given this evidence, Bada and co-workers used the date on the Laguna Skull to calibrate the rate of racemization of aspartic acid in human bone in a Californian coastal environment, giving a rate constant (k_{asp}) of 1.08×10^{-5} yr^{-1}. They then applied this rate constant to the other samples listed in Table 8.1, yielding even older dates, and pushing back the arrival of humans to around 50 000 years ago. Corroboration

was provided by comparing an AAR date of 33 000 years obtained on a dwarf mammoth bone from Santa Rosa Island (after correcting the racemization constant for the present-day temperature difference between coastal California and the Southern Californian Channel Islands) with a radiocarbon date obtained on the same charred bone of $30\,400 \pm 2500$ years (UCLA 1898). It should be noted that the radiocarbon dates are uncalibrated, and are therefore likely to be too young (see below). In principle AAR dates are absolute and do not require calibration. It is possible, however, that AAR dates should be viewed as equivalent to uncalibrated radiocarbon dates if the racemization constant is determined by reference to the radiocarbon age of a 'calibration' bone.

Further work followed on both Californian and other human material, summarized by Bada and Helfman (1975). Relevant to the Californian Paleoindian debate are the additional data listed in Table 8.2. The result on the SDM-16704 is a duplicate (by NASA) of that listed in the first set of data (Table 8.1), and is in close agreement. Sample SDM-19241 was given a date of around 6000 years using a 'southern California post-glacial' k_{asp} value of 1.5×10^{-5} yr^{-1}. This agreed with a radiocarbon date on associated shells of 6700 ± 150 yr (LJ-79). Using this 'post-glacial average value' for k_{asp} also gave Stanford Man a date of ~7000 years. The Sunnyvale skeleton was also measured by NASA, using a slightly different value for the racemization constant (~7×10^{-6} yr^{-1}). Different

Table 8.2 *AAR dates on other Californian Paleoindian skeletons* (Bada and Helfman, 1975; Table 7, by permission of Routledge Publishers)

Sample	*Description and location*	D/L *aspartic acid*	*Predicted age (yr)*
SDM-16704 (skull)	Del Mar W-34A	0.520	47 000
SDM-16704 (femur)		0.470	41 000
SDM-16740	W-12A Cliffs north of	0.458	39 000
SDM-16724	Scripps	0.347	27 000
SDM-16706	Batiquitos Lagoon	0.505	45 000
SDM-19241	W-9 Cliff N of Scripps	0.154	~6000
Sunnyvale skeleton		0.522 (ulna) 0.498 (skull)	70 000?
Stanford Man I		0.14	~7000

values of k_{asp} were argued to be necessary for very old or very young samples to account for the different average deposition temperatures experienced by these samples. All samples measured by Bada were also assayed for nitrogen, as an approximate check on their antiquity (Oakley, 1963), and the paper reasonably concludes that these data constitute definitive evidence for the presence of humans in North America before 40 000 BP.

As noted above, the sceptics were still distinctly unconvinced, dismissing this evidence as coming from an unproven dating technique, and noting the lack of archaic features on these supposedly very old human skulls. In a scholarly review of the nature of the evidence for and against the presence of pre-Clovis humans in North America, Dincauze (1984) investigated the AAR evidence closely. She noted that the radiocarbon age on the Laguna Skull used to calibrate the AAR method was not secure, on the basis that earlier studies of the alluvial fan in which it was found produced bracketing radiocarbon dates in the 9th Millennium BP from marine molluscs, with a reverse relationship between stratigraphy and date. It therefore did not have the required geological or archaeological integrity for such a key measurement. She also noted that '*the great ages claimed for the skeletons were seriously incompatible with their wholly modern physical types*' (p. 289). Already by 1984 evidence from other dating techniques was beginning to suggest that the AAR dates were insupportable. The Sunnyvale skeleton, in particular, had been re-dated using first of all uranium series dating (Bischoff and Rosenbauer, 1981) and then by radiocarbon dating (Taylor *et al.*, 1983). The uranium series work (for a description of the method, see Aitken, 1990) used two independent decay schemes ($^{238}U \rightarrow ^{230}Th$ and $^{235}U \rightarrow ^{231}Pa$) as an internal check on the consistency of the measurements, and concluded from measurements on postcranial fragments that the Sunnyvale skeleton dated to 8300 (+230, −100) years BP (absolute) using the ^{230}Th method, or 9000 (±600) using ^{231}Pa. The Del Mar tibia (SDM-16704) was similarly re-dated to either 11 000 (+500, −100) or 11 300 (+1300, −1200) years BP. These are internally consistent, and compare very badly with 70 000 BP for Sunnyvale by AAR, and 48 000 for Del Mar. Furthermore, two radiocarbon dates on molluscs thought to stratigraphically bracket the Sunnyvale burial yielded dates of 10 100 to 10 400 BP (Bischoff and Rosenbauer, 1981; Table 1). Taylor *et al.* (1983) reported four ^{14}C measurements on organic fractions taken from the postcranial bones of the Sunnyvale skeleton, including three by the (then) novel method of accelerator mass spectrometry, all of which gave relatively recent dates: 4390±150 BP (UCR-1437A) using the 'conventional' method, and 3600±600 (UCR-1437A/AA-50), 4850±400 BP

(UCR-1437B/AA52) and 4650 ± 400 BP (UCR-1437D/AA-51) by AMS ('direct') dating. They concluded that this evidence was consistent with other geological, archaeological, and anthropometric evidence – by implication, therefore, there was a serious problem with the AAR dates.

As a result of these inconsistencies, the very same amino acid extracts that had been used to produce the contentious AAR dates were independently dated by the AMS method at the Oxford Radiocarbon Accelerator Unit of Oxford University (OxA numbers: Bada *et al.*, 1984) and the NSF Accelerator Facility for Radioisotope Analysis, University of Arizona, Tucson (AA numbers). The resulting combined publication (Taylor *et al.*, 1985) was given the conclusive subtitle 'none older than 11 000 C-14 years BP'. It is a good example of how advances in another scientific discipline (in this case, the development of accelerator physics) can come along at exactly the right time to apparently resolve the questions thrown up in another area entirely. In the early days of radiocarbon, as noted above, sample requirements were so large that it was normal to combust the entire bone – mineral, collagen, plus whatever else might be present – to give what was known as a 'whole bone' date. In the 1970s, it was realized that the mineral fraction was potentially subject to post-depositional contamination by dissolved carbonates in circulating groundwaters, and that purified collagen was a much better material to date (the 'collagen' dates discussed above). With the advent of AMS, sample preparation could go even further – down to the individual amino acid fraction, especially hydroxyproline which is virtually unique to mammalian collagen – thus enabling contaminating compounds to be removed, hopefully giving even more reliable dates (Stafford *et al.*, 1991). Without doubt, AMS dating revolutionized our approaches to prehistory during the 1980s, just as conventional radiocarbon dating had done two decades earlier (Gowlett and Hedges, 1986).

The revisions published by Taylor *et al.* (1985) (summarized also by Gowlett, 1986), as well as assigning all the Californian Paleoindian remains to the Holocene, also pointed to the apparent cause of the problem – not the AAR methodology itself, but in the use of a seemingly erroneous radiocarbon date of 17 150 BP for the Laguna Skull as the calibration. Unfortunately, AAR was to receive all of the blame for the furore, and has hardly yet recovered in the eyes of most archaeologists! Nevertheless, the revised dates themselves were taken as clear support – or at least, as not offering contradictory evidence – for the 'late school' of thought about the colonization of the Americas. Table 8.3 is compiled from a number of published sources giving relevant data (Stafford *et al.*, 1984; Bada *et al.*, 1984; Taylor *et al.*, 1985, Gowlett, 1986) with an attempt made to cross-correlate the dates (and remove some apparent

Table 8.3 *Comparison of AMS and AAR dates on Paleoindian skeletons, compiled from various sources listed in the text*

Skeleton	*AAR Age*	*Revised AMS C-14 Age (yr)*		*Other Dates (yr)*
Laguna	calibration	OxA-189	5100 ± 500	>14 800–17150 (C-14)
Sunnyvale	70 000?	AA-60 UCR-1437A	3600 ± 600	8300–9 000 (U-series)
		AA-51 UCR-1437D	4650 ± 400	UCR-1437A 4390 ± 150
		AA-52 UCR-1437B	4850 ± 400	
		OxA-187	6350 ± 400	
La Jolla SDM-16755	28 000	OxA-186 LJS II	5600 ± 400	LJS I 1770 ± 790
		AA-610A LJS II	4820 ± 270	LJS I 1850 ± 200
		AA-610B LJS II	5370 ± 250	LJS I 1930 ± 200
		AA -611 LJS II	6330 ± 250	
Del Mar SDM-16704	41–48 000	OxA-188	5400 ± 120	11 000–11 300 (U-series)
		OxA-774	5270 ± 100	
Stanford Man I (UCLA-1425)	~7000	OxA-152	4850 ± 150	
		OxA-153	4950 ± 130	
			5130 ± 70	
Yuha	23 600		1650 ± 250	5800 (U-series)
			2820 ± 200	
			3850 ± 250	
Taber, Alberta		OxA-773	3390 ± 90	[22 000–60 000 on
		Chalk River	3550 ± 500	geological grounds]
Los Angeles Man, Baldwin Hills	26 000		3560	

discrepancies!) as far as possible. A more complete (and first hand) review of the story regarding the dating of the Sunnyvale and Yuha skeletons was subsequently published by Taylor (1991).

It was immediately obvious that the AAR dates were serious overestimates, if the AMS dates were correct. Bada (1985) lost no time in pointing out that, if the correct calibration date had been used, the AAR dates were nowhere near as outrageous as they seemed, and specifically observed that the calibration date of 17 150 BP had already provided evidence for the presence of humans in America in the late Pleistocene, and that the AAR dates were consistent at the time with other available radiocarbon dates. He recalculated the racemization constant for aspartic acid, this time using the AMS dates on both the Laguna and Los Angeles (Baldwin Hills) material, to arrive at a value of $k_{asp} = 6.0 \pm 2 \times 10^{-5}$ yr^{-1}, which he felt was applicable to bones with poor amino acid preservation. Applying this to the other data gave dates which were uniformly within the Holocene, and are 'compatible with their AMS radiocarbon ages' (Bada, 1985; 645). Table 8.4 shows the results of this recalculation, adapted from Bada (1985). One feature of the revised AAR dates in Table 8.4 is that the error estimates have increased considerably, in some cases by an order of magnitude (probably as a result of increased uncertainty in k_{asp}), and are now significantly poorer than the error estimates attached to the AMS dates.

There are a number of observations to be made at this stage. The first is to note the poor concordance between the older (in the sense of those measurements carried out earlier) radiocarbon dates and the AMS dates. This is largely attributable to either the use (or incorporation) of the mineral fraction in the bone, or to the fact that the bones are poorly preserved – the total organic content is much lower than one would expect with fresh bone. This is a salutary reminder to archaeologists working in regions of the world where bone is generally poorly preserved, or to those tempted to use older radiocarbon dates from the literature. The second is that if we accept Bada's assertion that the revised AAR dates are consistent with the AMS dates, then this is only because the AAR dates have an unacceptably large associated error. Nor is there particularly good agreement between the other dating evidence (principally uranium series dates) for the same material (Table 8.3) with either the AMS or the AAR dates, beyond the fact that they are all Holocene. Furthermore, as Bada *et al.* (1984) themselves pointed out, there appears to be no straightforward relationship between the AMS radiocarbon age of the skeletons and the measured D/L ratio of the aspartic acid extracts, in contrast to the general correlations found in other parts of the world. They note that in the Paleoindian material, the

Table 8.4 *Revised aspartic acid racemization dates for Californian Paleoindians*
(After Bada, 1985; Table 1, reproduced by permission of the Society for American Archaeology from *American Antiquity*, vol. 50 no. 3 1985)

Sample	*Conventional C-14 age (yr)*	*AMS age (yr)*	*Previous AAR age (yr)*	*Revised AAR age (yr)*
Laguna	17 150 ± 1470 >14 800	5100 ± 500	calibration	calibration
Los Angeles	>23 600	3560 ± 220	26 000	calibration
Del Mar	–	5400 ± 120	46 000	7500 ± 3000
La Jolla				
SDM-16724			27 000	5000 ± 2000
SDM-16740			39 000	7100 ± 3000
SDM-16742			44 000	8000 ± 3000
SDM-16755	1770 ± 800 1850 ± 200 1930 ± 200	5330 ± 245 5600 ± 400 6326 ± 250	28 000	5100 ± 2000
Batiquitos Lagoon			45 000	8000 ± 3000
San Jacinto	3020 ± 140		37 000	5100 ± 2000
Sunnyvale	4390 ± 600	3600 ± 600 4650 ± 400 4850 ± 400 6300 ± 400	70 000?	8200 ± 3000

best correlation appears to be between the degree of organic preservation in the bone (as measured by the total amino acid content of the bones) and the degree of racemization. They attribute this to faster rates of racemization of aspartic acid in N-terminal (end chain) positions in the protein, which could occur if hydrolysis was cleaving the chain to leave the aspartic residues exposed. This point has been further studied by Taylor *et al.* (1989), who also correlate this non-age-related variability to the state of preservation of the bone – specifically, the total nitrogen content and the ratio of glycine to glutamic acid. Clearly, the state of preservation is an extremely important factor in any dating technique when applied to bone.

In fact, developing this hypothesis, the story has yet one more twist. Further work by Stafford and co-workers (Stafford *et al.*, 1990; 1991) on the AMS radiocarbon ages of mammoth and human bone in various states of preservation has shown that, depending on the level of collagen surviving in the bone, significantly different ages can be obtained on different amino acid fractions from the same bone. The problem is so severe in what they term 'non-collagenous' bone (bone with less than about 10% of the original amino acids remaining, or less than 0.3% total nitrogen) that age estimates varying between 2500 and 8000 years BP can be obtained on the same bone. Because the Californian Paleoindian skeletons are generally 'non-collagenous', they conclude that the revised Holocene dates themselves are questionable, and may be a result of poor preservation. Even if the Paleoindian bones were actually late Pleistocene, they might be expected to give Holocene AMS dates because of their poor preservation. The circle is complete!

SUMMARY OF THE CURRENT POSITION RELATING TO THE FIRST AMERICANS

But what of the original question – when did humans first enter North America? The debate shows no sign of lessening – in fact there is every reason to expect further intensification. The removal of early AAR dates was a blow to all 'long chronologists', although the addendum with regard to the reliability of even AMS dates from poorly preserved bone certainly counsels caution in the interpretation of the Holocene dates. In fact, the question of interpretation is more complex than this, since it has long been accepted that all radiocarbon dates must be calibrated to give correct calendar dates, usually against dendrochronological measurements (Bowman, 1990). Uncalibrated dates are uniformly several hundred years (or even thousands of years) too young by the late

Pleistocene. This calibration process has become somewhat bewildering, since it opens up the possibility of multiple probabilistic calibrated dates for a single radiocarbon date (because of 'wiggles' in the calibration curve), and usually means that apparently precise uncalibrated dates may get 'smeared' out by the calibration process. Up until recently this was of little concern to late Pleistocene geologists or archaeologists, since the means of calibration for dates earlier than around 9000 calendar years BP was not available, but recent work comparing uranium series dated coral with the dates obtained on the same sample by AMS has changed that (Stuiver and Reimer, 1993). It is now possible, at least in a crude and preliminary fashion, to calibrate radiocarbon dates going back to nearly 20 000 years BP. Batt and Pollard (in press) have carried out a study of the dates associated with the 'Clovis horizon' in North American prehistory, and have concluded that instead of Clovis being confined to a very tight timescale of 11 200 to 10 900 radiocarbon years BP (as suggested by Haynes, 1991), the true time period represented by these Clovis sites might be closer to a range of around 1850 calendar years, between about 14 000 and 12 150 cal BP (cal BP is the accepted terminology for calibrated radiocarbon dates, equivalent to calendrical years). This has implications for the interpretation of the radiocarbon evidence at sites such as Meadowcroft in the USA and Monte Verde in Chile, claimed to be definitely pre-Clovis on the basis of the uncalibrated dates (Batt and Pollard, in press). The full impact of these calibrated dates is yet to be felt in earlier prehistory (in North America, as elsewhere). It could mean that some of the chronological problems which have been debated for decades and which are apparently resolvable by radiocarbon are not, in fact, capable of resolution using radiocarbon alone – particularly, for example, the question of the synchronicity between the arrival of the Clovis people and the demise of the American megafauna.

The debate about which sites are definitive evidence of pre-Clovis occupation continues unabated. Purportedly early sites come and go with almost Warholian frequency. Adovasio (1993) lists four which 'will not go away' – the Bluefish Caves complex in Northern Yukon, the Nenana Valley Site complex in Alaska, Monte Verde in Chile, and Meadowcroft Rockshelter in Pennsylvania. All have apparently solid evidence for human occupation before the so-called 'Clovis horizon' at 11 500 radiocarbon years BP, but still the detractors are unconvinced. The debate is extremely polarized, and one wonders exactly what evidence would be necessary to sway the contestants either way. For a brief time, it looked as if the racemization of amino acids in bone protein might provide the clinching factor, but it was not to be!

OTHER ARCHAEOLOGICAL USES OF AMINO ACID RACEMIZATION: AGE AT DEATH

One of the lasting features of the AAR episode has been a (largely undeserved) reluctance to take the study of amino acid racemization seriously on the part of most archaeological scientists. This has not been the case in the wider field of Quaternary geochronology, where amino acid chronostratigraphy has continued to be widely applied: for example, the work of Bowen *et al.* (1989) on British Pleistocene deposits, where the epimerization of L-isoleucine to D-alloisoleucine is used to create a relative stratigraphic sequence which has been related to the oxygen isotope timescale using independent dating techniques (uranium series, thermoluminescence, and electron spin resonance techniques; for details of these, see Aitken, 1990).

Of considerable interest in recent years to archaeologists and forensic scientists has been the observation that not all mammalian tissues maintain their protein amino acids in the L-form. It has been known for some time that D-aspartyl residues accumulate *in vivo* with time in the metabolically stable proteins found in tooth enamel and dentine, and in eye lens tissue, due to the racemization of L-aspartyl residues (Helfman and Bada, 1976; Masters *et al.*, 1977). Work in Cardiff and Bradford has demonstrated that the ratio of D/L aspartic acid in dental collagen from

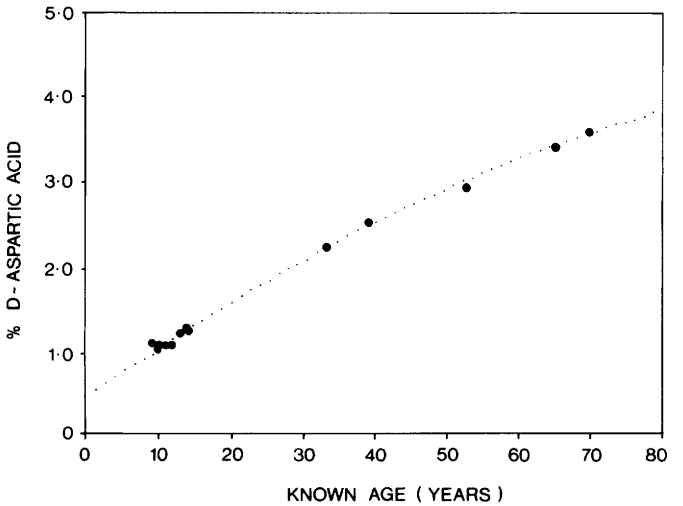


Figure 8.3 *Calibration curve for racemization of aspartic acid in modern dental collagen* (From Gillard *et al.*, 1991; Figure 1, by permission of Birkhäuser Verlag AG)

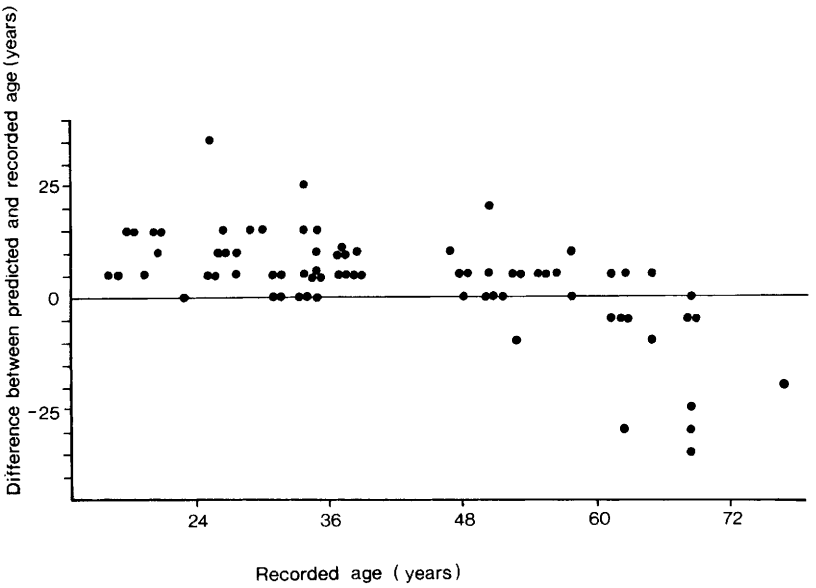


Figure 8.4 *Difference between 'real' age and age estimated from skeletal measurements plotted as a function of 'real' age in the Spitalfields population* (Redrawn from Gillard *et al.*, 1991; Figure 3, by permission of Birkhäuser Verlag AG)

human first premolars gives a very good prediction of the age of the individual, when using modern (extracted or fresh *post-mortem*) teeth (Gillard *et al.*, 1990; 1991; Child *et al.*, 1993). Figure 8.3 (from Gillard *et al.*, 1991) shows the calibration curve obtained on modern teeth, and Table 8.5 shows the result of using this calibration curve on modern teeth in a blind test, with very satisfactory results. However, when the same technique was applied to the known-age dental remains from inhumations in Spitalfields Church crypt, London (Reeve and Adams, 1993; Molleson and Cox, 1993), the results began to show some discrepancies (Figure 8.4, redrawn from Table II in Gillard *et al.*, 1991). This shows a general trend for the predicted age being an overestimate for individuals younger than about 50, and a tendency to underestimate the age of individuals over 60. Assuming that the 'known ages' are correct, the reason for this is not entirely clear. It is worth noting, however, that the amino acid method produces age estimates which are no worse than those from other techniques using skeletal parameters. In fact, even more interestingly, exactly the same pattern of predicted *versus* real ages is observed in the data obtained from combined skeletal observations in the

Table 8.5 *Blind test of aspartic acid calibration data on teeth extracted from living individuals*

Tooth	*%D-asp*	*Predicted age (yr)*	*Known age (yr)*
Lower first premolar	3.30	61.7	60.5
Lower lateral incisor	2.27	33.5	31.0
Upper canine	2.67	43.5	42.0

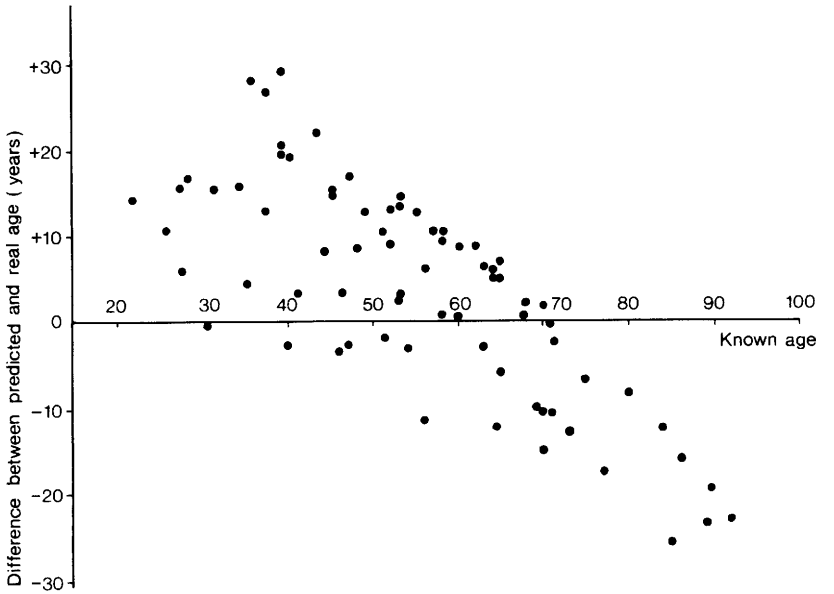


Figure 8.5 *Difference between 'real' age and age estimated from skeletal measurements plotted as a function of 'real' age in the Spitalfields population*
(Adapted from Molleson and Cox, 1993; Figure 12.4, by permission of the CBA and the author)

Spitalfields collection (Figure 8.5, adapted from Molleson and Cox, 1993; 171). These estimates were obtained by combining four well-established criteria for estimating age (based on observations of the pelvis, the femur, cranial sutures, and humerus). The Spitalfields work is almost unique in archaeology, in that it allows archaeological methods to be compared against the 'true' answer, and the observed discrepancies are therefore disturbing. From the amino acid viewpoint, it is at least encouraging to observe that the pattern and magnitude of the deviations

Table 8.6 *Reproducibility of Spitalfields aspartic acid age estimates from multiple teeth in the same dentition*

Catalogue no.	*Recorded age (yr)*	*Age estimates for different teeth (yr)*	*Range (yr)*
070	35	37.0, 39.0, 46.2	9.2
301	35	37.8, 51.0, 41.0	13.2
545	21	34.8, 34.5, 31.6	3.2
643	63	58.0, 57.0, 33.0	3.5
670	68	33.0, 37.2, 41.6	8.6
708	37	46.0, 45.6, 40.2	5.8
714	34	33.0, 34.6, 37.8	4.8
899	31	33.5, 30.2, 30.0	3.5

observed are very similar to those using other methods, suggesting the problem might have a common origin. Work currently underway suggests that some, at least, of the systematic errors common to all these ageing techniques might be due to the regression methodology generally employed, and may be reducible (Aykroyd, Lucy and Pollard, unpublished observations).

One of the more surprising results of this study is the data shown in Table 8.6, where three of the four first premolars from the same individual have been analysed independently. Obviously, it is expected that three teeth from the same dentition should give the same estimated age: it is also clear that they do not! Individual number 899, for example, gives an average age of 31.2 years (range = 3.5 years), comparing satisfactorily with the known age of 31; number 545 gave an average of 33.6 yr (range only 3.2 yr), but the true age was 21 (*i.e.*, a precise but not accurate age estimate); number 643, with a true age of 63, gave a range from 33.0 to 58.0 years (both inaccurate and imprecise). Similar variability has been observed in the use of aspartic acid age estimates in modern forensic cases. In the case of an individual burned to death in a motorway crash, the age obtained from all four first premolars was found to be 24.2 with a standard deviation of 4.5 years, compared with the subsequently-known true age of 25. In a second case, of an individual exposed for three months after death on the Brecon Beacons (South Wales) two estimates (from two first premolars) were obtained of 40.4 and 38.9 years; subsequently the true age was found to be 25 (Child *et al.*, 1993).

It is clear from these data, and forensic work carried out by other investigators (*e.g.*, Masters, 1986b), that the technique sometimes becomes unreliable when applied to potentially partially degraded dental

collagen. We have postulated elsewhere (Child *et al.*, 1993) that the problem may be caused by selective microbial attack on the collagen, possibly via collagenase activity, resulting in fragmentation of the molecule and subsequent variation in the rate of terminal amino acid racemization (Smith and Evans, 1980). Current work in Bradford is aimed at improving the reliability of the method by being more selective about the molecular integrity of collagen fraction extracted.

An obvious limitation to the use of racemization-derived age at death estimates on 'genuine' archaeological material (as opposed to the individuals from Spitalfields, where the time since death is relatively short) is the uncertainty in the contribution of *post-mortem* racemization to the measured ratio. We have calculated from thermodynamic considerations that, for aspartic acid, the post-mortem contribution in an individual aged 50 at death in 1740 AD and buried in the Spitalfields vaults should be around 0.1% D-aspartic acid (Gillard *et al.*, 1990), which is therefore negligible compared to the 3% expected *in vivo* (it should be noted that this figure includes an element of induced racemization during the measurement procedure, estimated at about 0.5%). This calculation also suggests that the observed discrepancies in the Spitalfields data were unlikely to be due to chemical factors. This would not, however, be the case for human remains of greater antiquity, nor if the average burial temperature were significantly higher. For example, a Roman person aged 50 should have accumulated approximately 2.8% D-aspartic acid *in vivo*, but could now have a post-mortem contribution of about 1% (2000 years at a burial temperature of 10 °C). The problem will become more serious with younger individuals further back in time, particularly in warm regions. It is not difficult to envisage a case where the *post-mortem* contribution outweighs the *in vivo* racemization by a factor of two or more. This problem must be resolved if the technique is to be reliable and become generally applicable archaeologically. There is still plenty of good chemical work to be done before the application of amino acid racemization measurements can be regarded as a routine method for archaeological and forensic science.

SUMMARY

The racemization of amino acids – the interconversion of one form to its mirror image – sprang to archaeological prominence as a potential dating technique, at a time when radiocarbon dating of bone required relatively large samples, and when the time depth limit of radiocarbon (around 35 000 years BP) was felt to be a serious drawback. A technique was devised which allowed a date to be obtained on very small samples

of amino acids extracted from bone collagen, but which, unfortunately, did not yield 'absolute' dates. Uncertainties in the rate constant, which may depend on the chemical environment of the amino acid, as well as the influence of changing temperature on the rate constant, meant that 'calibration' against a 'known age' (*i.e.*, radiocarbon dated) sample was deemed to be the best option. With hindsight, the sample chosen to calibrate the amino acid measurements on one of the most important archaeological problems – the first date of human occupation in the Americas – was unfortunate, in that it turned out to be in error, and the resulting controversial dates have subsequently been withdrawn. The main impetus for using amino acid racemization measurements as a dating technique has now largely gone. The advent of accelerator-based methods of radiocarbon dating has meant that good quality dates can be obtained on very small samples, although the age limit of ^{14}C dating still remains at around 30–40 000 years BP. Measurements of amino acid racemization for dating purposes are still used in geological circumstances to establish chronostratigraphic markers (*e.g.*, on protein extracted from shells) prior to the time period accessible by radiocarbon.

Apart from the fundamental interest in understanding the racemization process in mineralized tissue (still a great challenge), there is still reason to be involved in amino acid racemization research for archaeological purposes. It has been shown that measurements of the D-aspartic acid accumulated *in vivo* in certain tissues, especially dentine, are systematically related to the age of the individual, and this has profound implications, both archaeologically and, increasingly, forensically. Application of this technique to archaeological cases has so far proved problematic, although it has to be said that ages predicted from racemization measurements are no worse than those obtained by more traditional methods!

The episode of AAR dating relating to the origins of the Native American cultures is interesting as a case study for a number of reasons. Essentially, it is a good illustration of the scientific process at work in archaeology – a problem is identified, a technique devised to solve it, the result (in this case, archaeologically unpopular) is tested by other methods, and the solution is evaluated (in this case, eventually rejected). In his masterly review of the current state of archaeological science, Renfrew (1992; 290) noted that:

> *'Sometimes archaeologists and, I am afraid, archaeological scientists, rather readily take the view that conclusions offered by the application of the methods of the natural sciences carry with them more weight than do those deriving from archaeology as such.'*

and (p. 292):

'For is it not, these days, a defining character of real science that it is testable?' ... 'That archaeological science should sometimes give the wrong answers, and that these can later be shown to be indeed erroneous, must be counted one of the subjects greatest strengths.' ... 'Archaeological science has certainly now come of age, and can take such differences of opinion as these as a characteristic feature of scientific progress.'

The AAR dating episode certainly illustrates this second statement!

REFERENCES

Adovasio, J.M. (1993). The ones that will not go away. In *From Kostenki to Clovis*, eds. Soffer O. and Praslov N.D., Plenum Press, New York, pp. 199–218.

Aitken, M.J. (1990). *Science-based Dating in Archaeology*. Longman, London.

Bada, J.L. (1985). Aspartic acid racemization ages of California paleoindian skeletons. *American Antiquity* **50** 645–647.

Bada, J.L. (1991). Amino acid cosmogeochemistry. *Philosophical Transactions of the Royal Society of London B* **333** 349–358.

Bada, J.L., Gillespie, R., Gowlett, J.A.J. and Hedges, R.E.M. (1984). Accelerator mass spectrometry radiocarbon ages of amino acid extracts from Californian palaeoindian skeletons. *Nature* **312** 442–444.

Bada, J.L. and Helfman, P.M. (1975). Amino acid racemization dating of fossil bones. *World Archaeology* **7** 160–173.

Bada, J.L., Kvenvolden, K.A. and Peterson, E. (1973). Racemization of amino acids in bones. *Nature* **245** 308–310.

Bada, J.L. and Protsch, R. (1973). Racemization reaction of aspartic acid and its use in dating fossil bones. *Proceedings of the National Academy of Sciences of the USA* **70** 1331–1334.

Bada, J.L., Schroeder, R.A. and Carter, G.F. (1974). New evidence for the antiquity of man in North America deduced from aspartic acid racemization. *Science* **184** 791–793.

Bada, J.L. and Shou, M-Y. (1980). Kinetics and mechanism of amino acid racemisation in aqueous solution and in bones. In *Biogeochemistry of Amino Acids*, eds. Hare P.E., Hoering T.C. and King K. Jr., John Wiley, New York , pp. 235–255.

Batt, C.M. and Pollard, A.M. (in press). Radiocarbon calibration and the peopling of North America. Paper presented at *Archaeological Chemistry*, American Chemical Society Meeting, Anaheim, California, April 1995 (proceedings to be published).

Berger, R., Protsch, R., Reynolds, R., Rozaire, C. and Sackett J.R. (1971). New radiocarbon dates based on bone collagen of California

Paleoindians. *Contributions of the University of California Archaeological Research Facility* **12** 43–49.

Bischoff, J.L. and Rosenbauer, R.J. (1981). Uranium series dating of human skeletal remains from the Del Mar and Sunnyvale sites, California. *Science* **213** 1003–1005.

Bowen, D.Q., Hughes, S., Sykes, G.A. and Miller, G.H. (1989). Land-sea correlations in the Pleistocene based on isoleucine epimerization in non-marine molluscs. *Nature* **340** 49–51.

Bowman, S.G.E. (1990). *Radiocarbon Dating.* British Museum Press, London.

Child, A.M., Gillard, R.D., Hardman, S.M., Pollard, A.M., Sutton, P.A. and Whittaker, D.K. (1993). Preliminary microbiological investigations of some problems relating to age at death determinations in archaeological teeth. In *Archaeometry: Current Australasian Research*, eds. Fankhauser B.L. and Bird J.R., ANU Occasional Papers in Prehistory, Canberra, pp. 85–90.

Dawber, J.G. and Moore, A.T. (1980). *Chemistry for the Life Sciences.* Macmillan, London (2nd edn.).

Dillehay, T.D. and Meltzer, D.J. (eds.) (1991). *The First Americans.* CRC Press, Boca Raton.

Dincauze, D.F. (1984). An archaeo-logical evaluation of the case for pre-Clovis occupations. *Advances in World Archaeology* **3** 275–323.

Fagan, B.M. (1991). *Ancient North America.* Thames and Hudson, London.

Gillard, R.D., Hardman, S.M., Pollard, A.M., Sutton, P.A. and Whittaker, D.K. (1991). Determination of age at death in archaeological populations using the D/L ratio of aspartic acid in dental collagen. In *Archaeometry '90*, eds. Pernicka, E. and Wagner, G.A., Birkhauser, Basel, pp. 637–644.

Gillard, R.D., Pollard, A.M., Sutton, P.A. and Whittaker, D.K. (1990). An improved method for age at death determination from the measurement of D-aspartic acid in dental collagen. *Archaeometry* **32** 61–70.

Gowlett, J.A.J. (1986). Problems in dating the early human settlement of the Americas. In *Archaeological Results from Accelerator Dating*, eds. Gowlett, J.A.J. and Hedges, R.E.M., Oxford University Committee for Archaeology Monograph 11, Oxford, pp. 51–59.

Gowlett, J.A.J. and Hedges, R.E.M. (eds.) (1986). *Archaeological Results from Accelerator Dating.* Oxford University Committee for Archaeology Monograph 11, Oxford.

Haynes, C.V. Jr. (1991). Geoarchaeological and paleohydrological evidence for a Clovis-age drought in North America and its bearing on extinction. *Quaternary Research* **35** 438–450.

Haynes, C.V. Jr. (1992). Contributions of radiocarbon dating to the geochronology of the peopling of the new world. In *Radiocarbon after Four Decades*, eds. Taylor, R.E., Long, A. and Kra, R.S., Springer-Verlag, New York, pp. 355–374.

Helfman, P.M. and Bada, J.L. (1976). Aspartic acid racemization in dentine as a measure of ageing. *Nature* **262** 279–281.

Martin, P.S. (1984). Prehistoric overkill: the global model. In *Quaternary Extinctions: A Prehistoric Revolution*, eds. Martin P.S. and Klein R.G., University of Arizona Press, Tucson, pp. 354–403.

Masters, P.M. (1986a). Amino acid racemisation dating – a review. In *Dating and Age Determination of Biological Materials*, eds. Zimmerman M.R. and Angel J.L., Croom Helm, London, pp. 39–58.

Masters, P.M. (1986b). Age at death determinations for the autopsied remains based on aspartic acid racemization in tooth dentine: importance of postmortem conditions. *Forensic Science International* **32** 179–184.

Masters, P.M., Bada, J.L. and Zigler, J.S. (1977). Racemization in dentine as measure of ageing and in cataract formation. *Nature* **268** 71–73.

McConnell, D. (1973). *Apatite. Its Crystal Chemistry, Mineralogy, Utilization, and Geologic and Biologic Occurrences*. Springer-Verlag, Vienna.

McLean, F.C. and Urist, M.R. (1968). *Bone. Fundamentals of the Physiology of Skeletal Tissue*. University of Chicago Press, Chicago (3rd edn.).

Meltzer, D.J. (1989). Why don't we know when the first people came to North America? *American Antiquity* **54** 471–490.

Miller, E.J. and Gay, S. (1982). Collagen: an overview. *Methods in Enzymology* **82** 3–32.

Molleson, T. and Cox, M. (with Waldron A.H. and Whittaker, D.K.) (1993). *The Middling Sort. The Spitalfields Project Volume 2 – the Anthropology*. Council for British Archaeology Research Report 86, York.

Morrison, R.T. and Boyd, R.N. (1983). *Organic Chemistry*. Allyn and Bacon, Boston (4th edn.).

Oakley, K.P. (1963). Fluorine, uranium and nitrogen dating of bones. In *The Scientist and Archaeology*, ed. Pyddoke, E., Roy Publishers, New York, pp. 111–119.

Reeve, J. and Adams, M. (1993). *Across the Styx. The Spitalfields Project Volume 1 – the Archaeology*. Council for British Archaeology Research Report 85, York.

Renfrew, A.C. (1992). The identity and future of archaeological science. In *New Developments in Archaeological Science*, ed. Pollard, A.M., Proceedings of the British Academy 77, Oxford University Press, Oxford, pp. 285–293.

Schroeder, R.A. and Bada, J.L. (1973). Glacial-postglacial temperature difference deduced from aspartic acid racemization in fossil bones. *Nature* **182** 479–482.

Smith, G.G. and Evans, R.C. (1980). The effect of structure and conditions on the rate of racemisation of free and bound amino acids. In *Biogeochemistry of Amino Acids*, eds. Hare P.E., Hoering T.C. and King K. Jr., John Wiley, New York , pp. 257–282.

Stafford, T.W. Jr., Hare, P.E., Currie, L., Jull, A.J.T. and Donahue, D. (1990). Accuracy of North American human skeletal ages. *Quaternary Research* **34** 111–120.

Stafford, T.W. Jr., Hare, P.E., Currie, L., Jull, A.J.T. and Donahue, D.J. (1991). Accelerator radiocarbon dating at the molecular level. *Journal of Archaeological Science* **18** 35–72.

Stafford, T.W. Jr., Jull, A.J.T., Zabel, T.H., Donahue, D.J., Duhamel, R.C., Brendel, K., Haynes, C.V. Jr., Bischoff, J.L., Payen, L.A. and Taylor, R.E. (1984). Holocene age of the Yuha burial: direct radiocarbon determinations by accelerator mass spectrometry. *Nature* **308** 446–447.

Stuiver, M. and Reimer, P.J. (1993). Extended ^{14}C data base and revised Calib 3.0 ^{14}C age calibration programme. *Radiocarbon* **35** 215–230.

Taylor, R.E. (1991). Frameworks for dating the late Pleistocene peopling of the Americas. In *The First Americans*, eds. Dillehay, T.D. and Meltzer, D.J., CRC Press, Boca Raton, pp. 77–111.

Taylor, R.E., Ennis, P.J., Slota, P.J. Jr. and Payen, L.A. (1989). Non-age-related variations in aspartic acid racemization in bone from a radiocarbon-dated late Holocene archaeological site. *Radiocarbon* **31** 1048–1056.

Taylor, R.E., Payen, L.A., Gerow, B., Donahue, D.J., Zabel, T.H., Jull, A.J.T. and Damon, P.E. (1983). Middle Holocene age of the Sunnyvale human skeleton. *Science* **220** 1271–1273.

Taylor, R.E., Payen, L.A., Prior, C.A., Slota, P.J. Jr., Gillespie, R., Gowlett, J.A.J., Hedges, R.E.M., Jull, A.J.T., Zabel, T.H., Donahue, D.J., Stafford, T.W. and Berger, R. (1985). Major revisions in the Pleistocene age assignments for North American human skeletons: none older than 11 000 ^{14}C years B.P. *American Antiquity* **50** 136–140.

Weiner, S. and Traub, W. (1992). Bone structure: from ångstroms to microns. *FASEB Journal* **6** 879–885.

Williams, R.A.D. and Elliott, J.C. (1989). *Basic and Applied Dental Biochemistry*. Churchill Livingstone, Edinburgh (2nd edn.).

Chapter 9

Lead Isotope Geochemistry and the Trade in Metals

INTRODUCTION

The possibility of using some form of chemical fingerprinting to trace metal objects back to their ore source, and hence reconstruct prehistoric economic contacts, has long been one of the great goals of archaeological chemistry. Ever since the advent of instrumental methods for the chemical analysis of metals in the 1930s, large programmes of analysis of prehistoric metal objects have been undertaken. This approach is fraught with problems – far more than have been encountered with the study of other archaeological materials, with the possible exception of glass. The relationship between the trace element composition of a metalliferous ore and that of a metal object derived from it is an extremely complicated one, which is influenced by a number of factors, considered below.

Most metallic elements exist naturally as different isotopes – atoms of the same element which have the same chemical characteristics, but vary in weight (see Appendix 2). Normally, however, the relative abundance of the different isotopes of the same metal varies very little across the surface of the earth. Lead is unique in that it has a large range of natural isotopic compositions, due to the fact that three of its four stable isotopes (^{206}Pb, ^{207}Pb, and ^{208}Pb, with the fourth being ^{204}Pb) lie at the end of major radioactive decay chains. Geologists have been quick to exploit this fact – firstly, to obtain an estimate of the age of the Earth, and more recently to estimate the geological age of the various metalliferous deposits. The discovery therefore that the ratios of the stable isotopes of lead vary measurably from metal deposit to metal deposit, and are apparently unaffected by anthropogenic processes, was naturally hailed as a major breakthrough in the scientific study of archaeological metals.

Since then, it has been widely applied, particularly to Late Bronze Age metal production in Anatolia and the eastern Mediterranean. Although the results of lead isotope analysis have been widely quoted in the archaeological literature, the technique has recently entered a phase of fierce debate regarding the archaeological interpretations of the data, which involves a great deal of technological discussion of the genesis of lead isotope deposits and, ultimately, the uniqueness of a particular isotopic signature for a particular *'ore field'*. The bulk of this chapter gives the geochemical background to the technique of lead isotope analysis (which has never been fully presented to an archaeological audience), in order that the ensuing debate can be more easily interpreted. The chapter ends with a brief résumé of the current areas of debate, and a suggestion of alternative frameworks for the interpretation of lead isotope data.

THE TRACE ELEMENT APPROACH TO METAL PROVENANCE

As noted above and in Chapter 6, there are a number of theoretical difficulties associated with the archaeological interpretation of the chemical analyses of metal objects, putting aside for current purposes the practical difficulties relating to sampling precious objects, and producing representative chemical analyses of potentially inhomogeneous artefacts. The first source of confusion is the use of the term *'ore'*. In modern parlance, an ore is defined as *'a natural aggregation of minerals from which a metal or metallic compound can be recovered with profit on a large scale'* (Richards, 1909; 1). The application of this concept to prehistoric mineral processing is problematical, in that we have virtually no way of telling what value was placed upon the metal produced. It is quite possible that, at least in the early part of the Bronze Age, the symbolic value of possessing a metal object far outweighed any modern perception of value, which suggests that the current classification of 'economic mineral deposits' may not be useful in archaeometallurgy. Again, the modern term 'ore' is usually defined in terms of 'grams of metal extracted per tonne of ore processed', which is a measure of the abundance of metalliferous minerals in a particular vein, and therefore takes account of the presence of a large amount of *gangue* (vein material or parent rock, which has no significant metalliferous content). Most modern metal-bearing deposits contain mixed mineralization – either mixtures of different species of minerals of the metal required, or, more commonly, a mixture of different metalliferous minerals. Commercially exploitable ores for, say, gold, may only contain a few parts per million of the metal. In antiquity,

it is likely that miners were able to exploit extremely rich deposits of metalliferous minerals, in which case the 'ore' may well have contained very little gangue, and may even have been a relatively pure mineral species. The importance of this is that when one discusses the trace element composition of an ancient metal deposit, it is important to keep clearly in mind the possible difference between the minerals which make up the deposit and the mineralogical constitution of any 'ore' which may have been extracted from the deposit. This makes the use of modern-day data on ore deposits a hazardous business.

Further complications set in if the ore is reduced to the metal by means of some sort of furnace technology. Almost certainly the ore extracted from the mine (or removed from an open cut vein, or simply scraped from the surface) would have been enriched via some sort of *beneficiation* process (possibly involving crushing or washing), which may influence the mineralogical composition of the ore. Depending on the nature of the ore, it may have been necessary to employ some form of *roasting* prior to reduction – partly to mechanically break up the ore, and partly to convert sulfides into oxides. This step would almost certainly result in the loss of some of the more volatile components of the ore, especially elements such as arsenic. The furnace reduction process itself may have caused further volatilization, and the addition of any other material to liquefy the slag produced (if any) may have added additional trace elements to the final metal. Thermodynamic considerations of the partition of trace elements between different liquid phases in the melt may also need to be considered. In all but the earliest stages of metallurgy, the properties of the metal were modified via the addition of other metals to produce alloys – tin added to copper to produce bronze early on in the Bronze Age, with lead being added at a later stage to produce an alloy which was sufficiently fluid at moderate temperatures to produce complex castings. Any such additions will certainly complicate the interpretation of trace element data.

There has been considerable debate about the nature of these earliest alloying processes. Adherents to the 'minimalist school' of early technology believe that many of the early alloys – certainly the copper-arsenic alloys which usually predate the use of tin bronzes, and possibly even the earliest tin bronzes themselves – were the result of the fortuitous smelting of particular mixed mineral deposits. That is not to say that the production of these alloys may not have been deliberate, but that it was deliberate by geographical choice of deposit rather than careful mixing of smelted metals. Others believe that control over the composition of the alloys was exerted by subsequent additions of the appropriate quantities of alloying metal to a relatively pure smelted metal. Half way

between these two views is the possibility that control took place by careful blending of ores from different geographical locations before smelting, to give the appropriate balance of composition in the final metal. Whatever the true position, the important point from our perspective is that the metal which emerged from the primary manufacturing process almost certainly represents the smelting products of several minerals, possibly from several geographical locations. The manufacturing processes of metal goods from these primary metals can only further complicate the picture, since these would almost certainly have involved further high temperature fabrication processes (with associated volatilization), and may have required additional mixing.

Combine this with the almost inevitable practice of recycling metals in all but the most primitive metal using communities, and the further electrochemical modification of the structure and composition of buried metalwork, and it is not surprising that many archaeometallurgists have come to the view that it is extremely unlikely that anything positive can be said about the relationship between trace element composition of a metal object and its precise ore source. Clearly, some statements can be made, such as if a copper alloy object contains nickel, then it is probable that the copper comes from a nickeliferous deposit, as discussed in Chapter 6. The reverse, however, is much less certain. The reason why a copper object does not contain nickel may indeed be the result of a lack of nickel in the ore source, but it may also be influenced by the manufacturing processes. Laboratory simulation work has shown, for example, that nickel-bearing copper ores do not contribute nickel to the smelted copper unless the temperature is in excess of 950 °C (Thomas, 1990; Chapter 6). We must conclude that the measured chemical composition of a metal artefact is a complex function of the chemistry of the ore source(s) from which it is derived, the thermodynamics and kinetics of the high temperature processes employed, of anthropogenic factors such as alloying and recycling of metals, and the electrochemistry of the corrosion processes acting during burial (and possibly after burial, either as a result of conservation processes, or as the object equilibrates with a new environment). Little wonder that few archaeological scientists place much faith in the results of trace element provenancing of metals! Although these comments are primarily aimed at smelted non-ferrous metals (in practice archaeologically this means copper alloy objects), the same considerations apply to ferrous metals and to those metals which occur '*native*' (principally gold and platinum, although silver and copper are also known in the native state, and meteoritic iron can be considered a native metal). Few people have seriously attempted to provenance iron objects on the trace

element composition of the metal, although some work has been done on the composition of the slag inclusions, which is likely to pose serious problems in terms of the partitioning of elements between metal and slag in the liquid state, and the trace element contribution of the fluxes added to the furnace charge. A great deal of work has been done on the trace element and inclusion patterning in gold alloys, but there appears to have been little critical questioning of the underlying geochemical assumption that the impurity patterns should carry any geographical significance. In general it might be safe to assume that they do not.

When it was demonstrated in the late 1960s that the lead isotopic composition of metal objects might give directly an indication of the ore source, it was therefore eagerly applied by archaeological scientists with access to the necessary high precision heavy element mass spectrometers. This was even more the case when it was realized that the method applied not only to lead artefacts (relatively rare in the archaeological record), but also to the traces of lead left in silver extracted from argentiferous lead ores by the cupellation process, and to the lead impurities left in copper objects smelted from impure copper ores. It certainly appeared as if a prayer had been answered, and that some of the key questions asked of archaeometallurgists (such as the origin of the silver used in classical Athenian coinage, and the sources of the primary copper used in the Aegean Bronze Age) could now be answered with considerable confidence. Thirty years later there is a growing feeling that some of this confidence might have been misplaced, and that the time is ripe for a reassessment.

NATURAL RADIOACTIVITY AND THE STABLE ISOTOPES OF LEAD

The basic concepts of nuclear structure and isotopes are explained in Appendix 2. This section derives the mathematical equation for the rate of radioactive decay of any unstable nucleus, in terms of its *half-life.*

The rate of decay of an unstable parent nucleus at any time t is proportional to the number (N) of atoms left (Faure, 1986; 38). In other words, the rate at which the number of radioactive nuclei decline is proportional to the number left at that time. Expressed mathematically, this becomes:

$$-\frac{dN}{dt} \propto N \tag{1}$$

The symbol λ is introduced as the constant of proportionality, which is

termed the *decay constant* of the parent nucleus, and is characteristic of that nucleus (with units of inverse time):

$$-\frac{\mathrm{d}\mathcal{N}}{\mathrm{d}t} = \lambda\mathcal{N} \tag{2}$$

Rearranging terms and integrating gives the following equation:

$$-\int \frac{\mathrm{d}\mathcal{N}}{\mathcal{N}} = \lambda \int \mathrm{d}t \tag{3}$$

which integrates to give:

$$-\ln \mathcal{N} = \lambda t + C \tag{4}$$

The constant of integration C can be defined by setting the starting condition so that $\mathcal{N} = \mathcal{N}_0$ at time $t = 0$, giving $C = -\ln \mathcal{N}_0$:

$$-\ln \mathcal{N} = \lambda t - \ln \mathcal{N}_o \tag{5}$$

Rearranging this gives:

$$\begin{aligned} &\ln \mathcal{N} - \ln \mathcal{N}_o = -\lambda t \\ &\ln \frac{\mathcal{N}}{\mathcal{N}_o} = -\lambda t \\ &\frac{\mathcal{N}}{\mathcal{N}_o} = e^{-\lambda t} \\ &\mathcal{N} = \mathcal{N}_o \mathrm{e}^{-\lambda t} \end{aligned} \tag{6}$$

This is the basic equation describing the decay of all radioactive particles, and, when plotted out, gives the familiar exponential decay curve. The parameter λ is characteristic of the parent nucleus, but it is not the most readily visualized measure of the rate of radioactive decay. This is normally expressed as the *half-life* ($T_{\frac{1}{2}}$), which is defined as the time taken for half the original amount of the radioactive parent to decay. Substituting $\mathcal{N} = \mathcal{N}_0/2$ into the above equation gives:

$$\begin{aligned} &\ln\left(\frac{1}{2}\right) = -\lambda T_{\frac{1}{2}} \\ &\ln 2 = \lambda T_{\frac{1}{2}} \\ &T_{\frac{1}{2}} = \frac{\ln 2}{\lambda} = \frac{0.693}{\lambda} \end{aligned} \tag{7}$$

The half-lives of natural radioisotopes vary widely, between fractions of a second and many billions (10^9) of years, and are widely tabulated (*e.g.*, Littlefield and Thorley, 1979; Appendix C).

Nuclei which are radioactively unstable usually decay by the emission of one of three particles from the nucleus, traditionally labelled α, β and γ particles. The largest, slowest, and least penetrating of these are the α particles, which turn out to be the nucleus of the helium atom – *i.e.*, two protons and two neutrons, with an overall charge of +2. Decay by α emission is restricted to the heavier elements, and can be summarized in the following general equation:

$$^{A}_{Z}X \rightarrow ^{A-4}_{Z-2}Y + ^{4}_{2}He$$

The ejection of the α particle (labelled as a helium nucleus in the above equation) from the nucleus of element X results in the transmutation of X into Y, which has an atomic number two less than X (*i.e.*, two positions below it in the Periodic Table). The particular isotope of element Y which is formed is that with an atomic mass of four less than the original isotope of X.

The next heaviest of the particles which can be emitted during radioactive decay is the β particle, which has been identified as being the same as an electron – a much lighter particle, with a mass of approximately 1/1840 of that of either the proton or the neutron, but carrying a single negative charge. It is important to realize that this is still a particle which has been ejected from an unstable nucleus, and not to confuse it with the orbital electrons, which are (initially at least) unaffected by these nuclear transformations. The general equation for β decay is:

$$^{A}_{Z}X \rightarrow ^{A}_{Z+1}Y + ^{0}_{-1}\beta$$

Here the effect of the emission is to increase the atomic number by 1 (*i.e.*, to transmute X into the next heaviest element in the Periodic Table, Y), to leave the atomic weight unchanged (a so-called *isobaric* transmutation), and to emit the β particle, which is conventionally given a mass of 0 and a charge of −1. Although it does not actually happen like this, it is often useful to think of the β process as being the conversion of a neutron into two equal but oppositely charged particles, the proton and the electron, as follows:

$$^{1}_{0}n \rightarrow ^{1}_{1}p + ^{0}_{-1}e$$

In this notation, the upper line of superscripts refer to the mass of the particles in atomic mass units (u), and the subscripts refer to the charge. The resulting electron is then ejected from the nucleus as the β particle.

For the current discussion, the third type of particle, the γ particle, is of less interest, even though it is the most energetic and penetrating of the radioactive particles. It is in fact not a particle in the same sense as the α and β particles – it is a quantum of high energy electromagnetic radiation, which can nevertheless be thought of as a particle as a result of particle-wave duality. It effectively has zero mass, and no electrical charge, and therefore the emission of a γ particle leaves the nucleus unchanged in terms of A and Z. It is best thought of as being a mechanism for removing energy from a nucleus in an energetically excited state as a result of other processes.

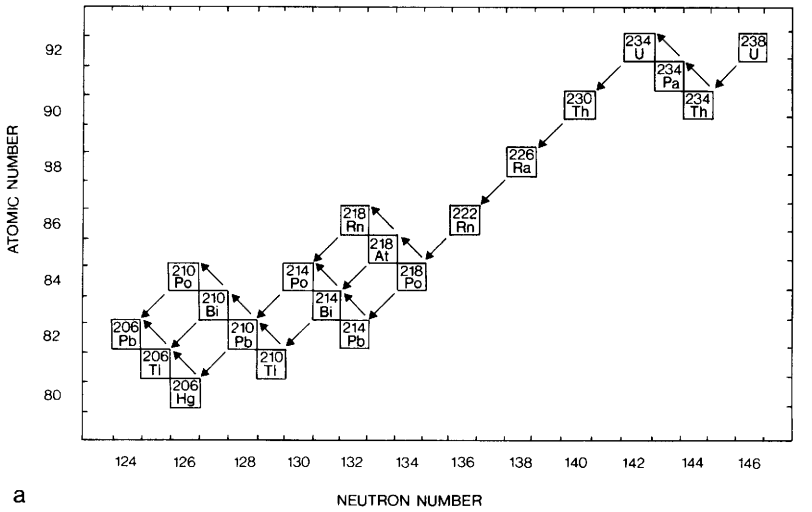
Although many nuclei are naturally radioactive, there are three main radioactive series in nature, all of which are relevant to a discussion of the isotopic composition of natural lead. These start with the elements uranium and thorium (^{238}U, ^{235}U, and ^{232}Th) and all end in one of the three stable isotopes of lead (^{206}Pb, ^{207}Pb, and ^{208}Pb respectively). Although each chain goes through a large number of intermediate unstable nuclei (Figure 9.1), the three chains can be summarized as follows:

$$^{235}_{92}\mathrm{U} \rightarrow {}^{207}_{82}\mathrm{Pb} + 7^{4}_{2}\mathrm{He} + 4\beta^{-} + \text{energy}$$

$$^{238}_{92}\mathrm{U} \rightarrow {}^{206}_{82}\mathrm{Pb} + 8^{4}_{2}\mathrm{He} + 6\beta^{-} + \text{energy}$$

$$^{232}_{90}\mathrm{Th} \rightarrow {}^{208}_{82}\mathrm{Pb} + 6^{4}_{2}\mathrm{He} + 4\beta^{-} + \text{energy}$$

Thus the chain that starts with ^{238}U goes through eight radioactive decay processes which result in the emission of an α particle, and six involving β particles, and the stable end member is ^{206}Pb, at which point the series ends. The ^{206}Pb produced as a result of these radioactive processes is termed *radiogenic*, to distinguish it from any other ^{206}Pb which may exist. As can be seen in Figure 9.1, in detail each of these decay chains involve a number of radioactive intermediates (termed *daughters*), all of which have a particular half-life. As an example, the full list of the half-lives of the elements involved in the decay chain of ^{238}U is given in Table 9.1. Inspection of this shows that the first step in the chain (^{238}U to ^{234}Th) has by far the longest half-life (approximately 4.5×10^{9} years, compared to the next longest, which is 2.5×10^{5} years). This is also true of the other two decay chains (listed in full in Russell and Farquhar, 1960; 4–5). Consideration of the behaviour of these chains under conditions of secular equilibrium (*i.e.*, when the rate of decay of the daughter isotope becomes equal to the rate of decay of the parent, and assuming a closed system), shows that it is possible to consider the decay chain simply in terms of the parent decaying directly to the stable



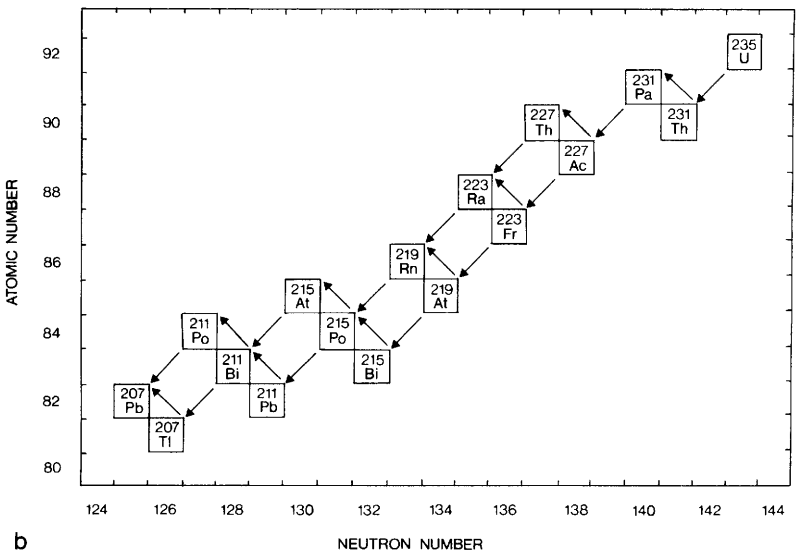


Figure 9.1 *Radioactive decay chains of ^{238}U, ^{235}U, and ^{232}Th.* (a) *decay of ^{238}U to stable ^{206}Pb;* (b) *decay of ^{235}U to stable ^{207}Pb;* (c) *decay of ^{232}Th to stable ^{208}Pb (see facing page). On these plots, a downward-pointing arrow (to the left) generally signifies decay by α emission, and an upward-pointing arrow (also to the left) decay by β emission. Note that all three chains 'branch' at several points, and that all involve isotopes of the radioactive gas radon (Rn) at some stage.* (After Faure, 1986; Figs. 18.1-3. Copyright 1986 John Wiley & Sons, Inc. Reprinted by permission of the publisher)

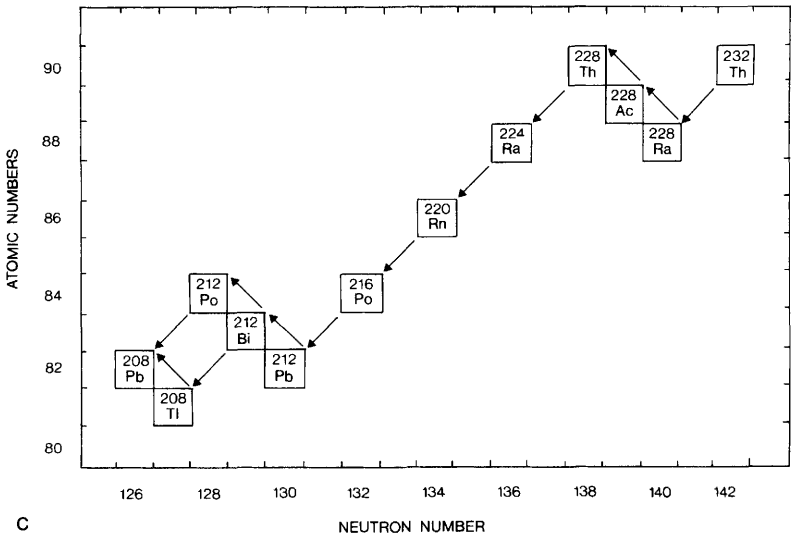


Figure 9.1 *continued*

Table 9.1 *Decay series of* ^{238}U
(From Russell and Farquhar, 1960, Table 1.2, with half-lives modified by data in Littlefield and Thorley, 1979; Appendix C)

Isotope	*Particle emitted*	*Particle energy (MeV)*	*Half-life*
^{238}U	α	4.18	4.5×10^9 yr
^{234}Th	β	0.205, 0.111	24.1 d
^{234}Pa	β	2.32, 1.50, 0.60	1.14 min
^{234}U	α	4.763	2.5×10^5 yr
^{230}Th	α	4.68, 4.61	8.0×10^4 yr
^{226}Ra	α	4.77	1620 yr
^{222}Rn	α	5.486	3.825 d
^{218}Po	α	5.998	3.05 min
^{214}Pb	β	0.65	26.8 min
^{214}Bi	α 0.04%	5.46	19.7 min
	β 99.96%	1.65, 3.17	
^{214}Po	α	7.680	160 μs
^{210}Tl	β	1.8	1.32 min
^{210}Pb	β	0.018	25 yr
^{210}Bi	β	1.17	4.8 d
^{210}Po	α	5.298	140 d
^{206}Pb			Stable

lead end point, with a half-life essentially the same as the longest half-life in the system, which happens to be that of the parent isotope (Faure, 1986; 285). Thus each of the three decay chains can be simplified to the following, which is particularly useful when considering the evolution of the isotopic compositions of terrestrial lead deposits:

$$^{238}U \rightarrow ^{206}Pb \qquad T_{\frac{1}{2}} = 4.468 \times 10^9\ yr$$

$$^{235}U \rightarrow ^{207}Pb \qquad T_{\frac{1}{2}} = 0.7038 \times 10^9\ yr$$

$$^{232}Th \rightarrow ^{208}Pb \qquad T_{\frac{1}{2}} = 14.01 \times 10^9\ yr$$

THE LEAD ISOTOPIC COMPOSITION OF METALLIFEROUS DEPOSITS

Having now established that three of the four stable isotopes of lead lie at the end of very long-lived radioactive decay chains, it is now appropriate to consider models for the development of the lead isotope composition of metallic ores. The fourth stable isotope (^{204}Pb) is not produced radiogenically, and is therefore termed *primeval* – its existence is the result of being present at the beginnings of the solar system, and therefore being incorporated into the earth as it solidified. The abundance of the three radiogenic isotopes also has a primeval component, to which has been added a radiogenic component. It is conventional to use isotopic ratios when discussing lead isotope geochemistry – geologists use the ratios $^{206}Pb/^{204}Pb$, $^{207}Pb/^{204}Pb$, and $^{208}Pb/^{204}Pb$, since ^{204}Pb is non-radiogenic, and these ratios occur in the equations for the isotopic evolution of ore bodies (see below), but there is also a practical reason. Modern mass spectrometrists use a technique called *thermal ionization mass spectrometry* (TIMS) to make these very high precision measurements, and one way of achieving the precision necessary is to measure all the isotope abundances simultaneously as ratios, since this minimizes variations due to small fluctuations in the ion beam flux (see Chapter 2). The raw data therefore consists of isotopic ratios, and it is convenient to retain these in subsequent considerations.

It is now conventional to classify lead-bearing deposits into two types – *ordinary* or *common* lead deposits, and *anomalous* deposits. Since these names were first applied in the 1960s, it has become apparent that in fact anomalous deposits are more common than ordinary deposits, but the names are still in use. *Ordinary lead* is found in the '*conformable*' mineral deposits of volcanic island arcs, where the mineralization is hosted in stratigraphic sequences of marine volcanic and sedimentary rocks. These

deposits have simple lead evolution histories with lead being derived by volcanic activity from the lower crust and mantle, without radiogenic lead contamination from the upper crust. *Anomalous lead* occurs in deposits that have had more complex evolutionary histories experiencing radiogenic lead contamination from the upper crust. In general the isotopic composition of these deposits cannot be explained by the simple models described below. The earliest model for the isotopic evolution of lead minerals is called the *Holmes–Houtermans* model. The impetus for developing such a model came from a desire to be able to calculate the age of the Earth from the isotopic composition of common lead ores (Holmes, 1946).

The Holmes–Houtermans model makes a number of assumptions which are important to enumerate (Faure, 1986; 310):

(i) originally the Earth was fluid and homogeneous, at which time the U, Th, and Pb were evenly distributed;
(ii) the isotopic composition of this *primeval lead* was the same everywhere;
(iii) on cooling, the Earth became rigid, and local variations arose in the U/Pb and Th/Pb ratios;
(iv) in any region, the U/Pb and Th/Pb ratios subsequently change only as a result of radiogenesis;
(v) at the time of formation of a *common (ordinary) lead* mineral, the Pb was separated from the U and Th, and there was no further change in its isotopic composition.

It is important to distinguish clearly in this scenario between the general solidification of the Earth's crust, which had the effect of 'freezing in' variations in the U/Pb and Th/Pb ratios, and the specific mineralization event which created the galena (lead sulfide, PbS) deposits, which removed the lead from the uranium and thorium, and effectively therefore 'froze' the isotopic composition of the lead in the galena at the values representative of the time of mineralization.

The equation for the growth of a stable daughter from a radioactive parent can be easily derived from equation (6) above, which is the familiar radioactive decay curve. We can write that:

$$D = N_0 - N$$

where D and N are the abundances of the daughter and parent after time t, with N_0 being the quantity of parent present at $t = 0$. Combining these two gives:

$$D = N_0 - N_0(e^{-\lambda t})$$

or:

$$D = N_0(1 - e^{-\lambda t}) \tag{8}$$

This equation expresses the growth of the daughter from the radioactive parent in terms of the amount of the parent originally present (N_0). It is more useful if we use equation (6) to replace N_0 with N (the number of parent nuclei remaining after time t) in this equation, since N is the quantity which is actually measurable. Thus:

$$N_0 = Ne^{\lambda t}$$

and therefore:

$$D = N(e^{\lambda t} - 1) \tag{9}$$

This is the equation for the growth of a stable radioactive daughter, assuming that no primeval D was present at time $t = 0$. If this was the case (and the amount is termed D_0), then the equation becomes:

$$D = D_0 + N(e^{\lambda t} - 1) \tag{10}$$

Using the simplification outlined in the previous section (*i.e.*, that the radiogenic production of ^{206}Pb from ^{238}U can be regarded as a single step with a half-life equal to that of the decay of ^{238}U), the equation for the growth of radiogenic ^{206}Pb can therefore be written as:

$$^{206}Pb = (^{206}Pb)_i + {}^{238}U(e^{\lambda_1 t} - 1)$$

where the subscript i denotes the initial (primeval) amount of ^{206}Pb present, and λ_1 is the effective decay constant of ^{238}U to ^{206}Pb. Conventionally, as noted above, this is expressed as a ratio to the abundance of ^{204}Pb, as follows:

$$\frac{^{206}Pb}{^{204}Pb} = \left(\frac{^{206}Pb}{^{204}Pb}\right)_i + \frac{^{238}U}{^{204}Pb}\left(e^{\lambda_1 t} - 1\right)$$

Similar equations can be written for ^{207}Pb and ^{208}Pb using their appropriate radioactive parents and decay constants. If $t = 0$ is taken to represent the time of the formation of the Earth's crust, then these three equations describe the trajectory of the isotopic composition of terrestrial lead from that time. If T is the time elapsed since the formation of the Earth, (*i.e.*, the age of the Earth), and t_m is the time before present at which the lead minerals were formed, then, using the assumptions of the Holmes–Houtermans model given above, the isotopic composition of a common lead deposit formed t_m years ago is given as follows:

$$\frac{^{206}Pb}{^{204}Pb} = \left(\frac{^{206}Pb}{^{204}Pb}\right)_i + \frac{^{238}U}{^{204}Pb}\left(e^{\lambda_1 T} - e^{\lambda_1 t_m}\right)$$

According to the model, this isotopic composition was fixed t_m years ago, on the separation of the lead minerals from the uranium- and thorium-bearing environment, is unchanged to the present day, and is therefore measurable. There are two similar equations for the other two radiogenic stable isotopes of lead. To simplify the manipulation, we can use the following notation:

$$\frac{^{206}Pb}{^{204}Pb} = a, \quad \frac{^{207}Pb}{^{204}Pb} = b, \quad \frac{^{208}Pb}{^{204}Pb} = c$$

and:

$$\frac{^{238}U}{^{204}Pb} = \mu, \quad \frac{^{232}Th}{^{204}Pb} = \omega$$

The third ratio required ($^{235}U/^{204}Pb$) can be calculated from the expression for μ, since the ratio of $^{238}U/^{235}U$ is accepted to be 137.88. Using these symbols, the three equations can be reduced to:

$$\begin{aligned} a &= a_o + \mu(e^{\lambda_1 T} - e^{\lambda_1 t_m}) \\ b &= b_o + \frac{\mu}{137.88}(e^{\lambda_2 T} - e^{\lambda_2 t_m}) \\ c &= c_o + \omega(e^{\lambda_3 T} - e^{\lambda_3 t_m}) \end{aligned} \qquad (11)$$

which form the basis for the Holmes–Houtermans model for the isotopic composition of common lead deposits. Providing we have a value for T (the age of the Earth) we can use these equations to predict the so-called *model age* (t_m) for such deposits. The age of the Earth has been relatively precisely estimated using lead isotopic measurements on meteorites. Certain meteorites contain an iron sulfide phase called *troilite* which contains appreciable lead but virtually no uranium or thorium, and which is believed to be the most primeval lead available to us on Earth. This allowed equations derived from the above model to be solved for T, and the currently accepted value is now 4.55×10^9 years (Faure, 1986; 312).

Using this value, we can solve the equations (11) above, enabling us to determine the age of common lead deposits with single stage histories (*i.e.*, formed by a single metallogenic event, with no subsequent alteration). This results in a pair of bivariate plots (conventionally $^{207}Pb/^{204}Pb$ *vs.* $^{206}Pb/^{204}Pb$, and $^{208}Pb/^{204}Pb$ *vs.* $^{206}Pb/^{204}Pb$), the first of which results from solving the pair of equations involving a and b, and the second from b and c. A graphical representation of the diagram for

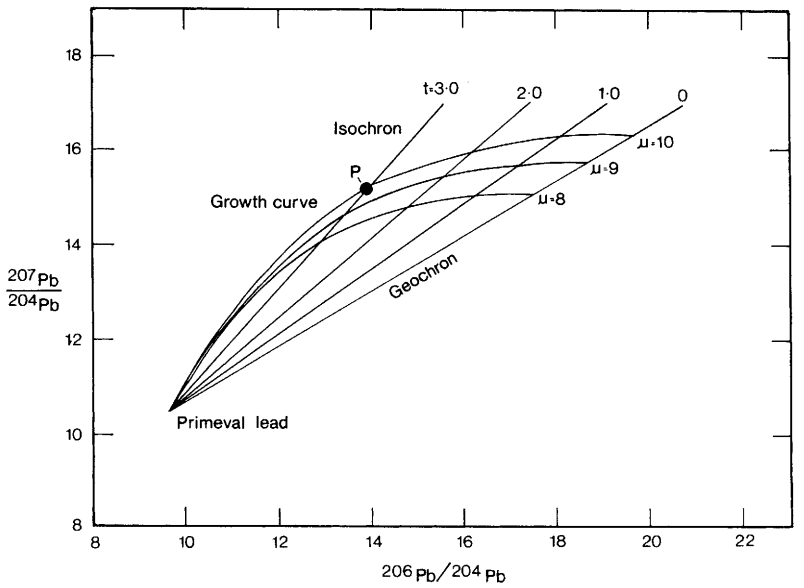


Figure 9.2 *Holmes–Houtermans model for the evolution of common lead deposits. (Growth curves for values of μ (ratio of $^{238}U/^{204}Pb$) equal to 8, 9, and 10 are shown.* Isochrons *(see text) corresponding to model ages of 1, 2, and 3 billion (10^9) years are plotted. The point P corresponds to the isotope ratios of a deposit formed 3 billion years ago with a μ value of 10. The isochron corresponding to a model age of zero gives the range of compositions expected from deposits forming now, depending on the μ value of the deposit, and is termed the* geochron
(From Faure, 1986; Figure 19.2. Copyright 1986 John Wiley & Sons, Inc. Reprinted by permission of the publisher)

$^{207}Pb/^{204}Pb$ *vs.* $^{206}Pb/^{204}Pb$ is shown in Figure 9.2. The solution of the equations is a set of growth curves, depending on the initial value of the parameter μ (the ratio $^{238}U/^{204}Pb$) at the time of the formation of the deposit (which is usually assumed to be the same as the present day overall value in the geographical region of interest, since ^{238}U decays only very slowly, and ^{204}Pb is not radiogenic). All of these trajectories start at a point which represents the isotopic composition of primeval lead (corresponding to $t_m = T$), and end at a point equivalent to the isotopic composition of common lead minerals formed today (when $t_m = 0$). The exact location of any deposit along any of these trajectories depends on the time of formation of the deposit, but all deposits formed at the same time lie along straight lines starting at the primeval point, termed *isochrons*. The equation of these isochrons can be calculated by

eliminating μ from the first two of equations (11), and the isochron gradient (m) is then given by:

$$m = \frac{1}{137.88}\left(\frac{e^{\lambda_2 T} - e^{\lambda_2 t_m}}{e^{\lambda_1 T} - e^{\lambda_1 t_m}}\right)$$

which shows that the gradient of the isochron is purely a function of t_m. It follows therefore, that all common lead deposits formed at the same time should lie on the same isochron, cutting through the primeval composition. The isochron which defines a model age of zero is termed the *geochron*, and corresponds to the predicted isotopic composition of lead minerals forming now. Distance from the primeval isotopic value along the isochron is therefore simply dependent on the value of μ in the ore-forming region. A value of $\mu = 0$ (an unlikely occurrence on Earth!) would give a mineral whose isotopic composition was the same as primeval lead. In theory therefore, no isotopic value of lead should plot in the region to the right and below the geochron, since this, according to the model, corresponds to a deposit which will form some time in the future! In fact, this does occur, but the explanation is somewhat more prosaic. In principle, however, plotting the isotopic ratios of a single stage common lead deposit onto this diagram should allow the age of formation and the value of μ to be predicted.

Unfortunately, it turns out that very few terrestrial deposits actually conform to this model. Faure suggests (1986; 316) that only about ten deposits (mostly in Australia) were so classified in 1968, and these all fitted the $\mu = 8$ growth line. He noted that these deposits all occurred in volcanic and sedimentary stratigraphic sequences of marine origin in volcanic island arcs. The lead in these deposits is thought to represent metal which was extruded through the crust and emplaced in an environment which virtually eliminated the mixing with crustal rocks (Russell and Farquhar, 1960; 53).

These so-called *conformable* deposits therefore represent the isotopic composition of lead in the upper mantle at the time of formation, which was uniform on a large scale, and are uncontaminated with radiogenic leads from the crust. This explains why they are common leads, and also why they have been observed to have a very narrow range of isotopic compositions. The majority of terrestrial deposits were not formed in this way. All deposits that are younger than the rocks hosting them, for example mineral veins, cannot be explained using the single stage model so far described. These deposits have complex lead evolution histories which involve mixing with upper crustal radiogenic lead.

The Holmes–Houtermans model does, however, form a basis for our

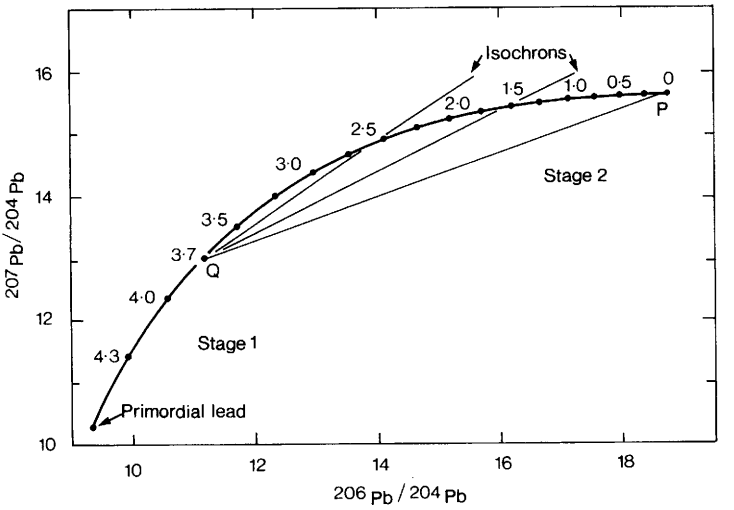


Figure 9.3 *Stacey and Kramers' two-stage evolutionary model for lead isotopes Point Q (at 3.7 billion years ago) indicates geochemical differentiation in the reservoir, and development subsequently occurs with a μ value of 9.735, compared to μ = 7.192 between P and Q. Isochrons grow from point Q – values for 0, 1.5 and 2.5 billion years are shown*
(From Faure, 1986; Figure 19.6. Copyright 1986 John Wiley & Sons, Inc. Reprinted by permission of the publisher)

further understanding of the majority of terrestrial deposits. Stacey and Kramers (1975) suggested a two stage model for the isotopic evolution of common leads, which in general gives dates of mineralization which are closer to the correct age, as judged from the geological age of the surrounding rocks. The evolution of lead deposits started as primeval lead some 4.5×10^9 years ago, as envisaged in the Holmes–Houtermans model, but at a more recent time (calculated to be 3.7×10^9 years ago) the U/Pb and Th/Pb ratios of the reservoir changed, and subsequent isotopic evolution can be defined in terms of isochrons focusing not on the primeval 'origin' but on this secondary event (see Figure 9.3). The date of the formation of a common ore deposit can therefore be calculated from a knowledge of the U/Pb and Th/Pb ratios in the appropriate reservoir, as in the Holmes–Houtermans model. The cause of the change in the U/Pb and Th/Pb ratios of the reservoir is not certain, but it is thought to represent a switch in the source of ore fluid from the mantle to a mixed mantle/subducted continental crust from which some lead had already been extracted.

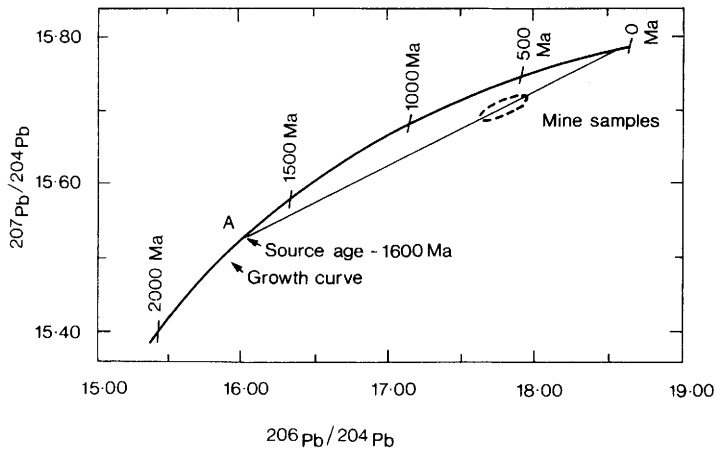


Figure 9.4 *Mixing model for the galena deposit at Bingham, Utah. The observed isotope ratios suggest that the deposit was formed from the mixing of a source of age 1.6 billion years with material of a much more recent age*
(Gulson, 1986; Figure 8.4, by permission of Elsevier Science and the author)

More general models for the isotopic evolution of lead ores have been devised by Cumming and Richards (1975), amongst others (see Gulson, 1986; Chapter 8 for a more complete summary). These are termed *continuous evolution models*, in which it is recognized that the U/Pb and Th/Pb ratios have been constantly increasing in the source reservoir with time (*i.e.*, the value of μ changes systematically, unlike the previous models which assume a fixed value). The most sophisticated of these models is that due to Amov (1983), which can be simplified to give the other models in certain circumstances. It is based on the assumption that the ratios $^{238}U/^{204}Pb$ and $^{232}Th/^{204}Pb$ have changed continually during the evolution of the Earth, and that temperature is the controlling factor determining chemical fractionation between the crust and mantle. During the early stage of the Earth's history, the crust was at a relatively high temperature and the mobility of uranium and thorium was higher than that of lead under these conditions, and so the ratios listed above increased. On cooling, the replenishment of these elements from the mantle diminished, and the ratios decreased. The equations for this model are complex, but can be reduced to either the single stage (Holmes–Houtermans) model if it is assumed that crustal formation and the beginnings of lead ore formation are contemporary, or the two-stage model if the date provided by Stacey and Kramers for the secondary

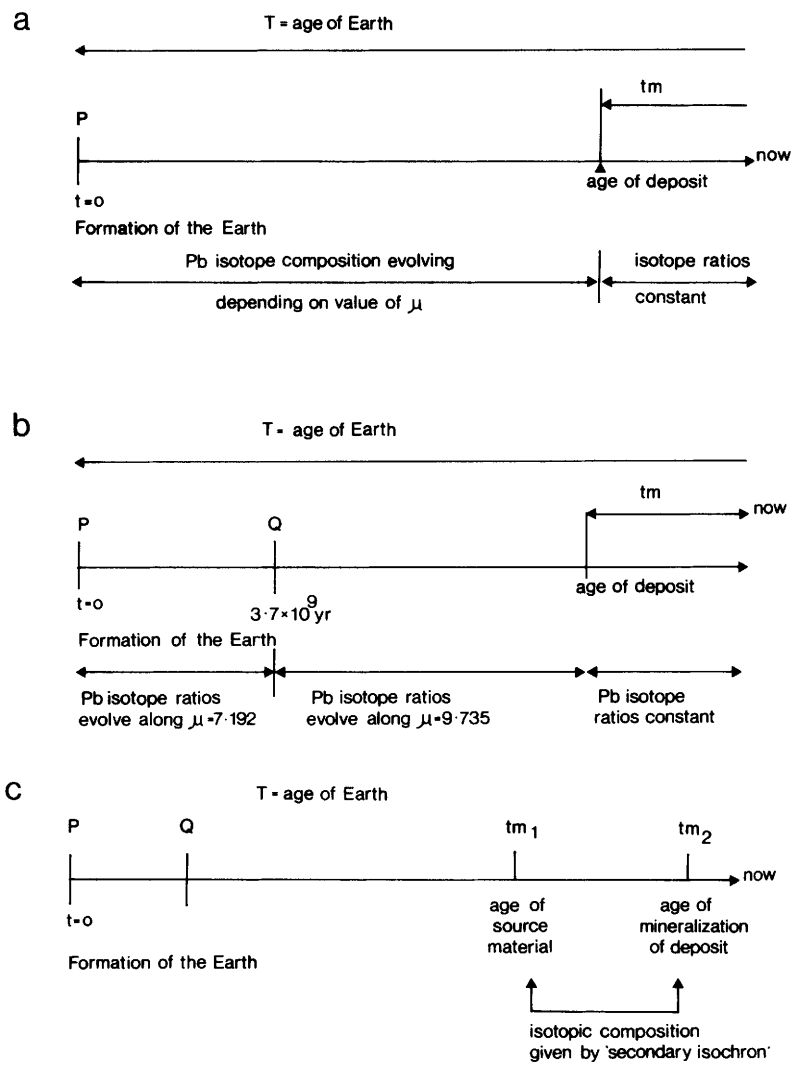


Figure 9.5 *Schematic diagrams to illustrate some of the models for the evolution of the lead isotope ratios in lead deposits: (a) Holmes–Houtermans, (b) Stacey–Kramers, (c) mixing model*

differentiation of the ore reservoir is accepted. The power of the Amov model is that other scenarios can also be considered.

These improved models still principally apply only to common leads, in which mixing with crustal rocks during emplacement was negligible. Most lead and base metal deposits are of *hydrothermal* origin, implying that

the metals were transported to the site of ore deposition by warm aqueous solutions (50–400 °C) containing high quantities of dissolved solids (up to 30 weight %). During the process of ore genesis hydrothermal fluids circulate in the surrounding country rocks leaching metals which are subsequently deposited to form the mineralization. The leaching of metals from a variety of sources, with various U/Pb and Th/Pb ratios, gives the lead an anomalous nature. Mineral deposits with complex lead evolution histories may be characterized by highly variable isotopic compositions and/or provide model ages that lie in the future due to a high radiogenic lead content, such as the *Mississippi Valley-type* deposits (Gulson, 1986; 154). These can be studied using multi-stage models which are discussed in detail by Gale and Mussett (1973), and reviewed by Gulson (1986; section 8.4). Several processes can be responsible for these deposits, but the simplest to visualize is the mixing of lead from different sources. Russell and Farquhar (1960) noticed that the isotope ratios of anomalous deposits from a particular ore body or mining region tended to lie along straight lines when plotted on an isotope ratio diagram similar to Figure 9.2. They attributed these 'anomalous lead lines' to the mixing of ordinary (single stage) lead with radiogenic lead derived from crustal minerals with different levels of uranium and thorium. Although this model can allow the age of formation of the single stage deposit to be predicted, together with the origin and age of formation of the contaminating radiogenic lead (Faure, 1986; 319), it is still limited to lead deposits which are primarily single stage in origin, and is ultimately therefore of restricted value. This approach is illustrated in Figure 9.4 (from Gulson, 1986). This shows that the isotopic composition of lead from Bingham Mine, Utah falls on a *secondary isochron* attributed to the mixing of leads of two ages, approximately 1.6 billion years old and approximately modern. It must be stated that not all anomalous deposits can be interpreted as simply as this, and not all linear distributions of isotopic data have such chronological significance.

Figure 9.5 shows a simplified schematic interpretation of some of these models. It is now generally agreed that no model is particularly appropriate at the detailed level, but, whether it can be modelled or not, it is still true to say that the measured ratios are related strongly to the age of the ore body and the geological processes such as mixing which have taken place. Therefore the isotopic 'signature' of a particular ore body may or may not be unique – ore bodies formed at the same geological time in the same manner and in the same country rock would be expected to have very similar isotopic ratios. Furthermore, the values of the isotopic ratios are constrained, in that they should conform to theoreti-

cally predictable growth curves when plotted appropriately. These constraints may well mean that isotopic data cannot be treated as if they were randomly-distributed variables. There is also good reason to suppose that the isotopic homogeneity of a deposit will be strongly influenced by its geological history. All of these factors are particularly important when the measurement of lead isotope ratios is applied to archaeology, as discussed in the next section.

LEAD ISOTOPES IN ARCHAEOLOGY

The use of lead isotopes in geology has become extremely well-established – originally because of the potential of the uranium and thorium decay series to provide an estimate of the age of the Earth, but subsequently as a technique for dating the emplacement of metalliferous deposits, and ultimately as a geochemical prospection technique for locating suitable deposits for commercial exploitation (Gulson, 1986). Given some of the problems identified in the section on the trace element approach to metal provenance, it is not surprising that the pioneering work of Brill and Wampler (1967) was subsequently eagerly seized on as a potential way out of the growing impasse that characterized trace element provenancing of metals. For reasons not explained, in this work the isotope ratios were plotted as $^{206}Pb/^{204}Pb$ *vs.* $^{206}Pb/^{207}Pb$ and $^{208}Pb/^{207}Pb$ *vs.* $^{206}Pb/^{207}Pb$ instead of the usual parameters used by isotope geologists. Nevertheless, it was clear that even with the analytical resolution available at the time, it was possible to differentiate between lead from various sources – on the basis of ore analyses, they separated samples from Laurion (Greece), England, and Spain into distinct groups, although they noted, for example, that an ore sample from northeastern Turkey fell into the same space as that occupied by three ores from England.

Brill was particularly interested in lead in glass rather than the metal itself, and this early paper also showed that the new technique was applicable to materials other than metallic lead. The two great advances in archaeological terms were the realization that lead isotope studies were applicable far beyond the study of metallic lead artefacts, which are relatively rare in the archaeological record, and the demonstration that anthropogenic high temperature processing did not, apparently, affect the isotopic ratio of the lead. The first metal to be studied in detail was silver, since the sources of silver are of great interest archaeologically (especially for coinage), and also because silver is normally extracted by cupellation from argentiferous lead sulfide deposits (*e.g.*, Tylecote, 1976; 38). Argentiferous galena may only contain around 0.1% Ag, but the

ancient cupellation process, in which the molten lead is selectively oxidized to litharge and removed, was quite capable of producing relatively high purity silver (95% Ag; Tylecote, 1976; 50) from such ores. As might be expected, lead is one of the major impurities in such silver (often of the order of a few percent), and this lead is sufficient to produce a lead isotope 'fingerprint' characteristic of the source of the argentiferous galena.

Despite the intrinsic value of silver items, they are also relatively rare in the archaeological record. The most common alloy system is normally that of copper, usually alloyed with either tin (bronze) or zinc (brass). Iron is also common in the later periods, but is usually subject to severe corrosion problems, and its survival rate is subsequently reduced (it also contains very little lead). Fortunately, copper can also be characterized from its lead isotope signature, since the primary ore of copper is chalcopyrite ($CuFeS_2$), which often occurs with galena (PbS) and sphalerite (ZnS). The copper smelted from such a deposit (even if the ore used is a secondary mineral formed by the oxidation of the primary deposit) would normally be expected to contain trace amounts of lead, which can therefore be compared with the isotope signature of the appropriate ore deposit. Apart from the potential complications discussed below, this assumes two things: (i) that the only source of lead in the copper alloy is that which is derived from the primary copper ore (*i.e.*, deliberately leaded copper alloys, such as those common from the Late Bronze Age onwards, cannot be expected to give information about the source of the copper), and (ii) that mixing of copper from different sources (or the addition of recycled metal during processing) is negligible. The second of these assumptions is currently the subject of some debate (Budd *et al.*, 1995b), and is discussed further below. Despite these possible complications, the method of lead isotope provenancing has been enthusiastically applied to copper alloy artefacts, especially those from the Late Bronze Age of the Aegean (*e.g.*, Gale and Stos-Gale, 1992, and references therein).

Although the technique of lead isotope analysis for archaeological provenancing has been in use for over 25 years, it is only now that some of the fundamental assumptions are being seriously reconsidered. The major areas under discussion at the moment can be classified under three headings:

(i) The assumption that anthropogenic processing such as roasting the ore, extraction of the metal, or cupellation produce no isotopic fractionation, and therefore that ore and artefact can always be compared directly.

(ii) The statistical, archaeological, and geological definition of the extent of a 'lead isotope field'.
(iii) The interpretation of lead isotope data, bearing in mind point (ii) and the possibility of significant pooling and recycling of metals. This is discussed below in the context of current debates relating to the reconstruction of the Bronze Age trade in metals in the Mediterranean.

The first point is probably the simplest to address. The reference most often cited in defence of the assumption that fractionation is unimportant is Barnes *et al.* (1978). In this pioneering work they compared the lead isotope ratios of galena with that of the lead smelted from it, with litharge (PbO) prepared from the smelted lead, and of a K_2O–PbO–SiO_2 glass, and a pigment ($Pb_2Sb_2O_7$) prepared from the lead. As reported in that paper, all the yields were high for each stage – 98.6% for lead recovery from galena, 95% for the production of litharge, and presumably virtually 100% for the other stages. The authors reported no measurable difference in any of the measured ratios, at least within the precision then obtainable by the measurement techniques of the day. This latter point is an important one, since Gale and Stos-Gale (1992) have observed that all of Brill's analyses published prior to 1974 'are of an accuracy too low to be of any use' (Gale and Stos-Gale, 1992; 70). (This is a comment purely on the precision of the measurements possible at the time, and applies to all of the earlier compilations of lead isotope data). In this case, the publication by Barnes *et al.* post-dates that cut-off, and although the reference in their text to the method employed is dated 1973 (Barnes *et al.*, 1973), it has to be assumed that this work conforms to the modern standards of measurement acceptability. Given that this is the case, with hindsight it is possible to show that the experiments they carried out were not those which might be expected to give significant fractionation, and the evidence cannot therefore necessarily be taken as convincing proof of the lack of fractionation in all processes.

In a pioneering piece of work, Scaife (1993) demonstrated that the theoretical considerations of the thermodynamics of *non-equilibrium evaporation* from a liquid first published by Mulliken and Harkins (1922) should have a significant effect on the isotopic ratio of lead, providing the non-equilibrium losses themselves are sufficiently large. Kinetic theory predicts that the lightest isotope will preferentially enter the vapour phase on evaporation, leaving the liquid enriched in the heavier isotope and the vapour phase in the lighter. If evaporation goes to completion (*i.e.*, all the liquid is evaporated) or the system is left to equilibrate, then there is no change in the isotopic ratio of the liquid or

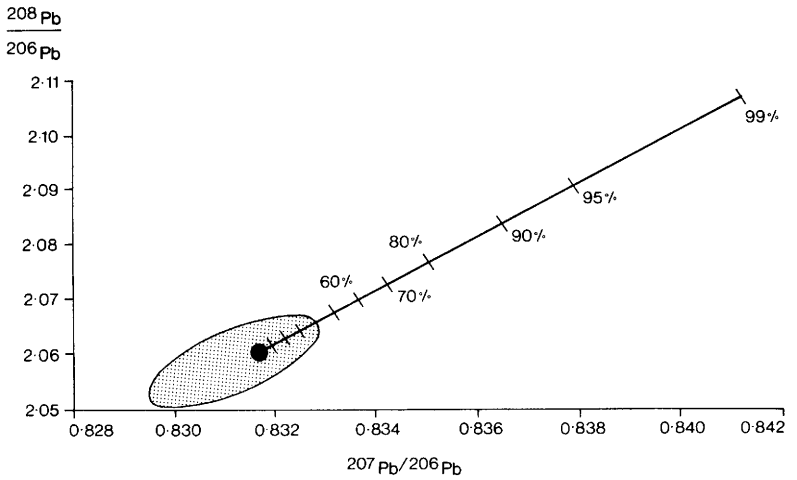


Figure 9.6 *The theoretical effect of non-equilibrium evaporation of lead on the isotopic ratio of a typical sample from the Laurion field (from Scaife, 1993). The shaded area shows the extent of the Laurion field as commonly defined, and the diagonal line shows the effect of increasing non-equilibrium losses on the isotopic ratio of a sample from the centre of the field*

condensate. Non-equilibrium evaporation is the situation in which the vapour phase is removed (by condensation, as in the early experiments, or as a result of a gas flow). Under these conditions, the lighter vapour phase will be removed and the liquid will be gradually enriched in the heavier isotope. The calculation can be extended to predict the change in isotopic ratios as a function of percentage non-equilibrium loss if the liquid contains more than two isotopes. The result of this theoretical calculation for lead is best summarized diagramatically, as shown in Figure 9.6. Here it has been assumed that a sample from close to the centroid of the published lead isotope field for Laurion (as defined by Gale *et al.*, 1980 and Stos-Gale *et al.*, 1986) has been subjected to increasing weight loss under non-equilibrium evaporation conditions, and the straight line shows the trajectory of the isotopic ratio of that sample as a function of the percentage of lead removed. It is clear in this case that 40% non-equilibrium losses would be sufficient to bring the sample to the edge of the field, and above around 60% would effectively remove it completely from the field as normally defined.

Although this is theoretically sound, the key question to be answered is whether any of the (largely unknown) ancient processes for the production of lead and silver might be expected to give such a large non-

equilibrium loss. This problem has been discussed at length in a recent paper (Budd *et al.*, 1995a). The most likely contender for a process with such high losses seems to be cupellation, in which a melt of around 99% Pb with less than 1% silver is converted via preferential oxidation to an alloy with greater than 95% silver – a total loss in excess of 99% of lead. Experiments conducted by Pernicka and Bachmann in 1983 on the cupellation of silver from Laurion galenas concluded that the process did not induce fractionation. The Ancient Metallurgy Research Group in Bradford has conducted a series of simulated laboratory metallurgical processes on galenas and lead, and, despite occasionally achieving significant lead losses, isotopic measurements have as yet shown no significant change in the isotopic ratios (Budd *et al.*, 1995c). At present, therefore, we conclude that although the model is theoretically correct, the circumstances under which non-equilibrium evaporation of lead may have taken place during most ancient processing techniques are unlikely to be significant. There does appear, however, to be another and perhaps ultimately more important conclusion to be drawn from this work – that is, that it might be unprofitable to attempt to increase the precision of the isotopic measurement techniques any further than the current levels. This is because small non-equilibrium losses may have occurred during the processing which could place an ultimate limit on the accuracy with which fabricated artefacts may be assigned to ore sources and so limit the resolution at which it is meaningful to define such fields. Subsequent consideration of non-equilibrium evaporation applied to other metallic systems (specifically tin) has given rise to a potentially fruitful method of quantifying the degree of recycling of copper alloys, as outlined below (Budd *et al.*, 1995c).

The second area of recent debate – the definition of the extent of an ore field – is a relatively complex area, which involves a discussion of the isotope geology of the ore field, the sampling strategy for the samples used to define the field, and the nature of the statistical procedures used to define the boundaries. The first question to be addressed is the extent of natural variation of the isotopic ratios in an ore deposit. The modern technique of measurement (thermal ionization mass spectrometry, TIMS) is capable of making measurements accurate to an absolute 95% error of 0.05% in the $^{207}Pb/^{206}Pb$ ratio and 0.1% for $^{208}Pb/^{206}Pb$ and $^{206}Pb/^{204}Pb$ (Gale and Stos-Gale, 1993). The isotopic homogeneity of a conformable massive sulfide deposit is often quoted to be about ± 0.1 to $\pm 0.3\%$ in the $^{206}Pb/^{204}Pb$ ratio (Gulson, 1986; 30), although it is well-known that most deposits show much greater variation, such as the Pb–Zn sulfides in the Upper Mississippi Valley, which show a large gradation across the deposit (nearly 9%, with $^{206}Pb/^{204}Pb$ varying from 21.88 to

23.96; Gulson, 1986; Figure1.4). Moreover, work on the isotopic variation across large single crystals of galena (*e.g.*, Hart *et al.*, 1981) has shown variations across a 13 cm crystal from Buick Mine in Southeast Missouri which are as large as the variations within the ore body itself (approximately 5% in $^{208}Pb/^{206}Pb$ and 4% in $^{207}Pb/^{206}Pb$). Although these are likely to be exceptional, it does highlight the necessity for good quality geological sampling of a deposit.

It is clearly important to have a good knowledge of the ore geology of the area before isotopic provenancing is undertaken. Unfortunately, ancient mining is often in areas which are no longer considered economically viable, and the data required by isotope archaeologists are not usually available in the published geological literature. This then requires that the archaeological programme should include a fieldwork strategy for sampling the relevant deposits as well as simply measuring the isotopic ratios in archaeological artefacts. This involves a number of considerations, ranging from the necessity for the mineralogical characterization of the deposit to the collection of an adequate number of samples to isotopically characterize the deposit. It cannot be assumed that if the isotopic signature of the lead minerals from a deposit have been measured, then the traces of lead in, say, a copper mineral will be the same. Pernicka (1993) has noted that an increasing number of copper minerals can be shown to have a different lead isotopic signature from the co-existing lead minerals because of incorporation of uranium and/or thorium at the time of formation of the deposit. It is therefore more important than ever only to compare metal artefacts with the kind of ore from which they may have come, such as silver objects with argentiferous galenas, copper alloys with copper minerals, *etc.* This principle has been repeatedly stated in the archaeological literature (*e.g.*, Gale and Stos-Gale, 1992; 73), but has occasionally been sacrificed in the interest of increasing the number of measurements which can be used to define a 'field'. A second, and related, principle which has been repeatedly stated is that only those ores available to ancient miners need be considered when defining a field. From one point of view this is a perfectly valid position – clearly modern isotope data from deposits so deep that no ancient process could possibly have exploited them, or from ores which are so finely dispersed that only modern extractive metallurgy can win metal, should not be included in any ore field constructed purely for archaeological purposes. In fact, one might wish to go further – only those ore deposits for which there is archaeological evidence of exploitation at the period in question should be considered as potential sources. (This raises a number of difficult side issues, such as the lack of any direct means of

dating ancient mining activity, and the validity of the assumption that we actually know – or will ever know – *all* of the sources of metal exploited in antiquity, which is very unlikely). On the other hand, this dictum can be seen as being in direct contradiction with the statement made above – that a comprehensive geological understanding of an ore deposit is required in order to characterize it fully. Clearly any research strategy must balance up both of these requirements.

One of the most controversial issues of all in the current lead isotope debate is how the source fields are constructed from collected geological samples. There is wide agreement that a statistically valid number of samples must be measured. Gale and Stos-Gale (1992; 76) stress the need for 'at least 20 geologically well-selected ore samples' per deposit, which seems to define an agreeable minimum level. There is also wide agreement that the data consists of measurements of three isotopic ratios, and that all three need to be taken into account when attempting to discriminate between deposits. Suggestions that the data should be converted into the four isotope abundance measurements (*e.g.*, Reedy and Reedy, 1992), although perhaps soundly-based statistically, are unlikely to be helpful in the light of the method used to produce the data (*i.e.*, ratio measurements). The main debate centres around how these three ratios are manipulated, and, as a side issue, the nature of 'outliers' in such data (Budd *et al.*, 1993; Scaife *et al.*, in press).

Some workers have approached the problem in exactly the same way as they would if they were dealing with analytical data from archaeological ceramics (*e.g.*, Sayre *et al.*, 1992). This approach leads to the definition of isotopic field boundaries as 95% 'confidence limits' around groups of points, calculated on the assumption of multivariate normality. Other workers have favoured the use of linear discriminant function analysis to distinguish between pre-defined groups (*e.g.*, Gale and Stos-Gale, 1992; 73ff). Scaife *et al.* (in press) have challenged both these approaches, not on the grounds of statistical propriety, but as a result of considering the nature of the data themselves. The multivariate normality of lead isotope data has to be questionable, on the grounds that the data are not a random sample drawn from a multivariately normal parent population. The parent population (the isotopic data characteristic of the ore deposit) is a constrained set – constrained by the equations of isotopic evolution discussed above. This is a fundamental difference in the nature of the data, and Scaife *et al.* have argued that the correct procedure for constructing ore fields is one which involves modelling the isotopic data, rather than treating it statistically as if it were unconstrained trace element data. Clearly this is an important point, and one which may have a strong influence over the separability of the different

isotopic fields, and therefore the degree of certainty with which archaeological samples can be identified to source. The magnitude of the differences (in terms of field definition) produced by these different models has yet to be established, and there is as yet no agreement about the correct procedure.

Another important factor in the definition of the relevant isotopic fields is the way in which 'outliers' are considered. Scaife *et al.* (in press) have argued that if an accurate and precise measurement is made on a sample which is geologically securely identified to a specific ore deposit, then it cannot be dismissed as an outlier, no matter how different it is from other samples relating to that deposit. Rather, it should be taken as an indication that the relevant 'isotopic field' has previously been defined too tightly. This may be somewhat utopian, in view of the acknowledged difficulties associated with the measurement of lead isotopic ratios in metalliferous ore samples which are very low in lead, but it must be accepted as a starting point for the definition of isotopic fields. Some concern has already been expressed about the understandable but dangerous practice of re-analysing those samples (and only those samples) which fall outside pre-defined field boundaries (Begemann *et al.*, 1995), and it would appear that a more scientific re-sampling procedure might be necessary in such cases.

LEAD ISOTOPES AND THE BRONZE AGE MEDITERRANEAN

Whilst these matters of detail form a lively debate in the literature and elsewhere, the important question to consider is how reliably can lead isotope data be interpreted archaeologically, and how does this impact on our knowledge of the ancient trade in metals. The vast majority of the archaeological effort in lead isotope analysis has been devoted to studying the trade in metals in the Mediterranean in the Bronze Age and later. This has been the subject of several reviews – Gale and Stos-Gale (1992) have summarized their lead isotope work carried out in Oxford as part of a larger British Academy project, which was an ambitious attempt to synthesize a range of scientific data with more traditional archaeological approaches to achieve an understanding of trade in the Bronze Age Mediterranean. A final report summarizing the achievements of this laudable effort in relation to Cyprus has recently been published (Knapp and Cherry, 1994). Sayre *et al.* (1992) have summarized the work carried out largely under the auspices of the Smithsonian Institution, Washington DC and the Brookhaven National Laboratory relating to the sources of metalliferous ores in ancient Anatolia, which

clearly has an important relationship to the work in the eastern Aegean. A group based in Heidelberg and Mainz have also been working for many years in the eastern Mediterranean and Anatolia (*e.g.*, Wagner *et al.*, 1986). Figure 9.7 shows a map of some of the more important prehistoric copper sources in the eastern Mediterranean.

Unfortunately, much of the basic data remains unpublished, and it is sometimes difficult to evaluate the competing interpretations put forward by these three groups. Nevertheless, it is possible to summarize essentially two competing interpretations from the lead isotope data. On the one hand, Pernicka and co-workers contend that: '... *There is, in fact, hardly any need for multivariate statistical methods at all. Searching for subtle differences in various projections of the data is likely to lead to overinterpretation.*' (Pernicka, 1993; 259). They claim that *'no convincing case has yet been presented where multivariate statistical methods revealed more information than the two diagrams* $^{208}Pb/^{206}Pb$ *vs.* $^{207}Pb/^{206}Pb$ *and* $^{204}Pb/^{206}Pb$ *vs.* $^{207}Pb/^{206}Pb$', and: *'it is simply unrealistic to pretend that a unique characterization of ore deposits will eventually be possible, if one would only use more sophisticated methods of data analysis'* (Pernicka, 1993; 259). The outcome of this approach is the view that it is only possible (or desirable) to define ore fields which cover a large geographical area, and which provide little in the way of highly detailed provenance information. As an example, Figure 9.8 shows the 'Aegean field' defined by Pernicka *et al.* (1984) (as published by Gale and Stos-Gale, 1992; Figure 9.4), which includes samples from the Troas, in western Anatolia, and many sites within the Aegean. This parsimonious approach is in direct contrast to the more ambitious attempts of the other groups to separate individual ore deposits within this area. An example of the level of discrimination claimed by the Oxford group is shown in Figure 9.9 (also from Gale and Stos-Gale, 1992; Figure 13), where some of the copper ores from Laurion near Athens, the Aegean and Anatolia are identified on one of the pair of lead isotope ratio plots. Apparent overlaps are resolved by discriminant analysis on selected groups, such as the resolution of the two Troad groups from the Kythnos samples, shown in Figure 9.10 (Gale and Stos-Gale, 1992; Figure. 14). A similar level of separation for other sources is claimed by the Brookhaven/Washington group, as exemplified by Figure 9.11 (from Sayre *et al.*, 1992; Figure 2).

These figures have been selected for no other reason than to illustrate the varying level of divisions in the data which the original authors themselves feel justifiable. It is noteworthy, however, that there is little agreement in terms of field definition between of the two teams who wish to see Pernicka *et al.*'s 'Aegean field' broken down into smaller units. For example, Figure 9.9 (from Oxford) shows a single field for the lead

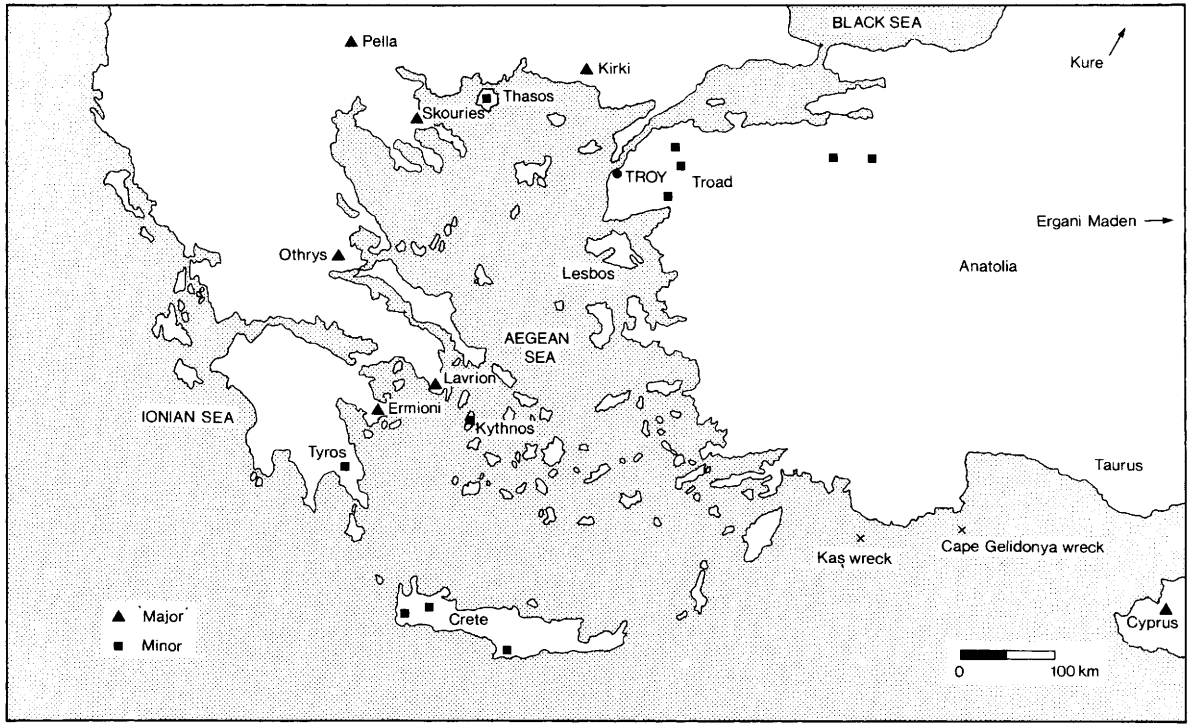


Figure 9.7 *Map of some of the more important prehistoric copper sources in the eastern Mediterranean*
(Adapted from Stos-Gale and Gale, 1990; Figure 1, in *Thera and the Aegean World III*, published by the Thera Foundation, London, with permission)

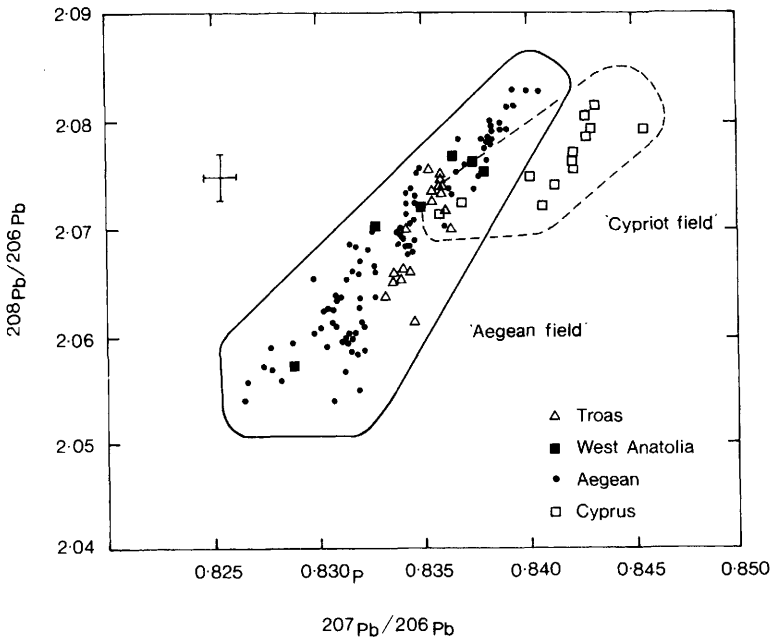


Figure 9.8 *The generalized 'Aegean field' defined by Pernicka* et al. *(1984), as given by Gale and Stos-Gale, 1992, Figure 4. (Note: the horizontal axis was labelled as $^{208}Pb/^{206}Pb$ by Gale and Stos-Gale, and has been corrected here). The cross in the upper left corner indicates the estimated precision of a single measurement* (Reproduced by permission from *Proceedings of the British Academy*, vol. 77, *New Developments in Archaeological Science.* © The British Academy 1992)

isotope ratios in Cypriot copper ores, whereas Figure 9.11 shows a similar field, but divided into two separate groups. It has been known for many years that the application of some multivariate methods of analysis to archaeological data may have a tendency to over-divide the data into apparently (statistically) valid sub-groups, but which may have no archaeological significance (Pollard, 1982), and this may be an example of that problem.

As might be anticipated, with so little agreement about the definitions of source ore fields between the major practitioners, there is a great deal of debate about the archaeological interpretations to be placed on such data. That debate is typified by the problems over the source attributions of the famous *oxhide ingots* recovered in large numbers from the Bronze Age Cape Gelidonya (13th Century BC) and Ulu Burun (Kaş; late 14th Century BC) shipwrecks (Bass, 1967, 1986; Bass *et al.*, 1989). These are

large ingots (typically between 20 and 30 kg) of relatively pure copper, whose name refers to the similarity between their shape and that of the stretched-out hide of an ox. It has been assumed that they represent trade in copper as a raw material, with an assumed origin of Cyprus, largely because of the known richness of the island in terms of copper ore, and because of the extensive evidence for prehistoric working of these deposits (Gale and Stos-Gale, 1992; 87). The earliest oxhide ingots so far known (principally from Hagia Triadha on Crete, and dated to the 16th Century BC) were analysed isotopically for lead, and appeared to have different isotope signatures from the copper sources on Cyprus, and to represent two sources with model ages calculated via the Cumming and Richards method to approximately 375 and 640 million years respectively (compared to the model age of around 100 million

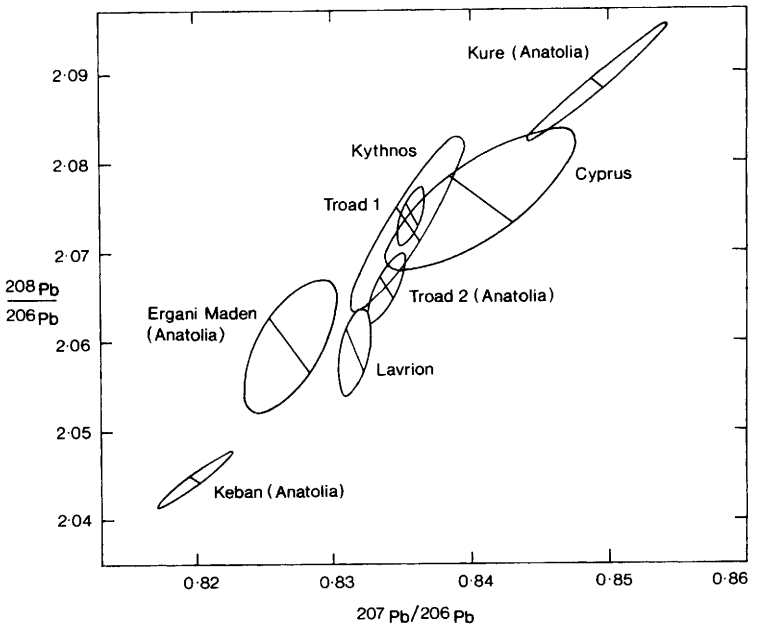


Figure 9.9 *Bivariate lead isotope ratio diagram for copper ores from some Aegean and Anatolian deposits, as defined by the Oxford group (Gale and Stos-Gale, 1992; Figure 13)*
(Reproduced by permission from *Proceedings of the British Academy*, vol. 77, *New Developments in Archaeological Science*. © The British Academy 1992)

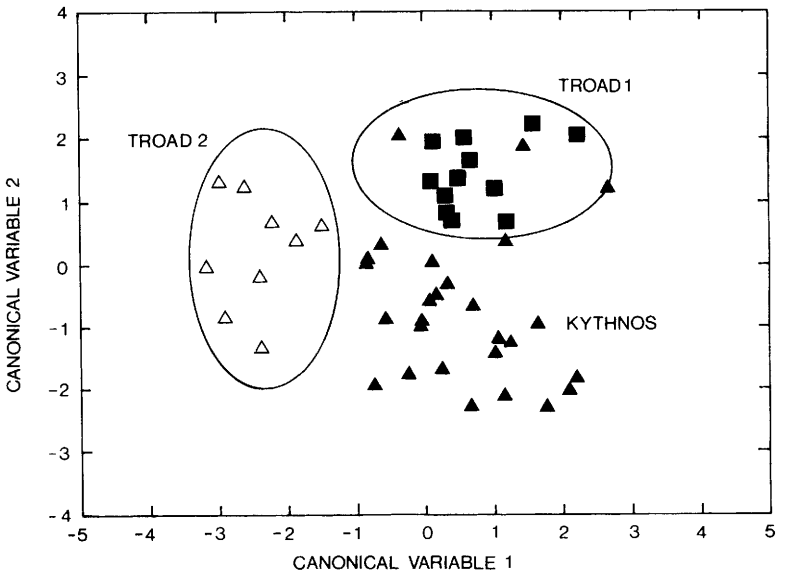


Figure 9.10 *Discriminant analysis of three overlapping groups shown in Figure 9.9, showing apparent separation of Troad 1 and 2 (in Anatolia) and Kythnian copper and slags (Gale and Stos-Gale, 1992; Figure 14)*
(Reproduced by permission from *Proceedings of the British Academy*, vol. 77, *New Developments in Archaeological Science*. © The British Academy 1992)

years for Cyprus). From a knowledge of the ore geology of the Mediterranean and Near East, Gale and Stos-Gale (1992; 91) concluded that the most likely sources were to be found in Iran or Afghanistan – a possibility not totally at odds with archaeological thinking, in the light of the known presence of lapis lazuli (presumed to be from Afghanistan) in Minoan and Mycenean contexts. Further work demonstrated that the 13th Century BC oxhide ingots found on Cyprus were isotopically consistent with the Cyprus copper ores, suggesting that by this time the vast reserves of copper on Cyprus were indeed being exploited.

More significant was the data on the ingots from the two shipwrecks found off the south coast of Turkey, mentioned above. Fifteen of the Cape Gelidonya ingots were analysed, and found to be isotopically consistent with the Cypriot field, as were the four samples taken from the 84 oxhide ingots on the Ulu Burun wreck. The latter ship was also carrying some much smaller copper ingots of a different shape, termed 'bun' ingots, some of which were also analysed, and not all were found to be consistent with Cyprus. This evidence was felt to be sufficiently strong to persuade

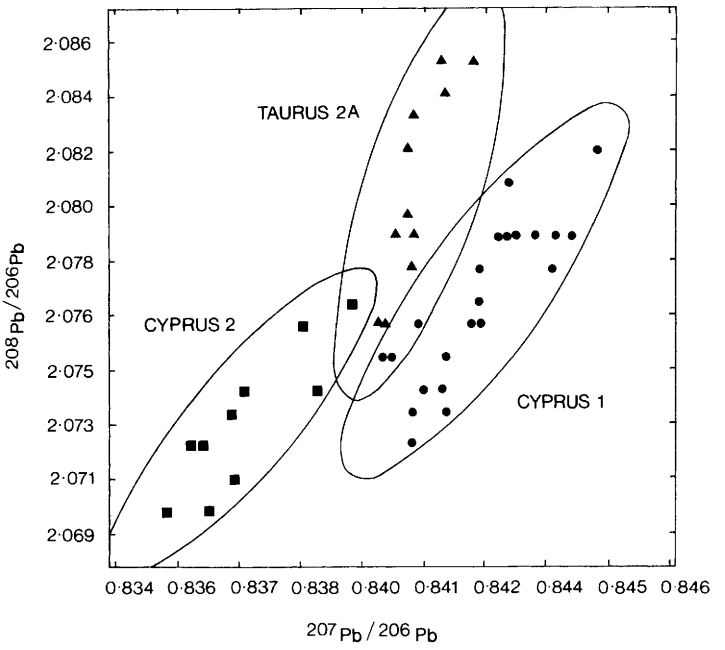


Figure 9.11 *Bivariate isotopic ratio plot of Cypriot and Anatolian ore deposits* (According to Sayre *et al.*, 1992; Figure 2, with permission)

Gale and Stos-Gale (1992; 94) to state that *'We can at least be sure that some Cypriot copper was being carried through the Mediterranean in the Late Bronze Age'.* Somewhat more controversial was the conclusion that the large number of oxhide ingot fragments found on Sardinia (some 2000 km distant in the western Mediterranean) were also isotopically identifiable as Cypriot products. Sardinia, whilst not being as rich in copper resources as Cyprus, still has extensive copper deposits. The proposal that copper was being shipped half-way across the Mediterranean to an island with its own abundant supply of copper was greeted with some scepticism in certain quarters (*e.g.*, Knapp *et al.*, 1988; Muhly, 1991).

Recalling that one of the three major groups involved in this type of work in the Mediterranean do not accept the divisions of the isotopic ore fields of the kind proposed in these interpretations, it seemed appropriate to re-evaluate the data – a proposal considerably hampered by the fact that much of the basic data remains unpublished and unavailable. The results of this reconsideration have been published recently (Budd *et al.*, 1995b), which sides very heavily with the long-held beliefs of Pernicka *et*

al. – *i.e.*, that it is essentially impossible to subdivide the various ore deposits in the eastern Mediterranean into individual separate 'fields'. There is some evidence, for example, that there is a substantial overlap between the Sardinian and Cypriot isotopic signatures, although this matter is still hotly disputed. An alternative model has subsequently been put forward for the apparent isotopic homogeneity of the lead isotopes in some of the oxhide ingots (Budd *et al.*, in press), based on the possibility of extensive mixing of copper sources, which remains to be tested on real data. Simple modelling of the recycling and mixing of metals from different sources suggests that the observed homogeneity of the isotopic signatures of objects such as the oxhide ingots may actually be a natural consequence of recycling, contrary to the popular view that it is a reliable indicator of metal from a single source (*e.g.*, *'we do not observe the smeared out lead isotope compositions for artefacts which would characterize mixing'* – Gale and Stos-Gale, 1992; 68). A novel idea has recently been put forward by Budd *et al.* (1995c) based on the possible non-equilibrium fractionation of tin isotopes in recycled bronzes which may allow the degree of recycling to be established for the first time.

SUMMARY

The use of lead isotopes to answer questions relating to the provenance and trade of metal resources has recently entered a new phase of research, in which some of the long-standing fundamental assumptions have been challenged. It is widely accepted that lead isotopes provide a unique fingerprinting technique, which is largely unaffected by anthropogenic processes, and which is considerably better than any possible use of trace elements to answer the same questions. Debate largely centres on the way in which an ore source is characterized, and the degree to which statistical methods of discrimination between closely-similar sources are applicable. The validity of the previously-held assumption that mixing and recycling of copper resources were not significant in the Bronze Age also requires further investigation. We confidently expect a fruitful period of research in this fascinating and challenging area, which brings together isotope geology, metallurgy, chemistry, and, of course, archaeology.

REFERENCES

Amov, G.A. (1983). Evolution of uranogenic and thorogenic lead, 1. A dynamic model of continuous isotopic evolution. *Earth and Planetary Science Letters* **65** 61–74.

Barnes, I.L., Gramlich, J.W., Diaz, M.G. and Brill, R.H. (1978). The possible change of lead isotope ratios in the manufacture of pigments: a fractionation experiment. In *Archaeological Chemistry II*, ed. Carter, G.F., American Chemical Society Advances in Chemistry Series 171, Washington DC, pp. 273–277.

Barnes, I.L., Murphy, T.J., Gramlich, J.W. and Shields, W.R. (1973). Lead separation by anodic deposition and isotopic ratio mass spectrometry of microgram and smaller quantities. *Analytical Chemistry* **45** 1881–1884.

Bass, G.F. (1967). Cape Gelidonya: a Bronze Age shipwreck. *Transactions of the American Philosophical Society* **78** 1–177.

Bass, G.F. (1986). A Bronze Age shipwreck at Ulu Burun (Kaş): 1984 campaign. *American Journal of Archaeology* **90** 269–296.

Bass, G.F., Pulak, C., Collon, D. and Weinstein, J. (1989). The Bronze Age shipwreck at Ulu Burun: 1986 campaign. *American Journal of Archaeology* **93** 1–29.

Begemann, F., Pernicka, E. and Schmitt-Strecker, S. (1995). Thermi on Lesbos: a case study of changing trade patterns. *Oxford Journal of Archaeology* **14** 123–136.

Brill, R.H. and Wampler, J.M. (1967). Isotope studies of ancient lead. *American Journal of Archaeology* **71** 63–77.

Budd, P., Gale, D., Pollard, A.M., Thomas, R.G. and Williams, P.A. (1993). Evaluating lead isotope data: further observations. *Archaeometry* **35** 241–263.

Budd, P., Haggerty, R., Pollard, A.M., Scaife, B. and Thomas, R.G. (1995c). New heavy isotope studies in archaeology. *Israel Journal of Chemistry* **35** 125–130.

Budd, P., Haggerty, R. Pollard, A.M. Scaife, B. and Thomas, R.G. (in press). Rethinking the quest for provenance. *Antiquity*.

Budd, P., Pollard, A.M., Scaife, B. and Thomas, R.G. (1995a). The possible fractionation of lead isotopes in ancient metallurgical processes. *Archaeometry* **37** 143–150.

Budd, P., Pollard, A.M., Scaife, B. and Thomas, R.G. (1995b). Oxhide ingots, recycling and the Mediterranean metals trade. *Journal of Mediterranean Archaeology* **8** 1–32.

Cumming, G.L. and Richards, J.R. (1975). Ore lead isotope ratios in a continuously changing Earth. *Earth and Planetary Science Letters* **28** 155–171.

Emsley, J. (1991). *The Elements*. Clarendon Press, Oxford (2nd edn.).

Faure, G. (1986). *Principles of Isotope Geology*. John Wiley, New York (2nd edn.).

Gale, N.H., Gentner, W. and Wagner, G.A. (1980). Mineralogical and

geographical sources of Archaic Greek coinage. In *Metallurgy in Numismatics Vol. 1*, ed. Metcalf, D.M., Royal Numismatic Society Special Publication 13, London, pp. 3–49.

Gale, N.H. and Mussett, A.E. (1973). Episodic uranium-lead models and the interpretation of variations in the isotopic composition of lead in rocks. *Reviews of Geophysics and Space Physics* **11** 37–86.

Gale, N.H. and Stos-Gale, Z.A. (1992). Lead isotope studies in the Aegean (The British Academy Project). In *New Developments in Archaeological Science*, ed. Pollard, A.M., Proceedings of the British Academy 77, Oxford University Press, Oxford, pp. 63–108.

Gale, N.H. and Stos-Gale, Z.A. (1993). Comments on P. Budd, D. Gale, A.M. Pollard, R.G. Thomas and P.A. Williams 'Evaluating lead isotope data: further observations', *Archaeometry*, **35 (2)** (1993), and reply. Comments...II. *Archaeometry* **35** 252–259.

Gulson, B.L. (1986). *Lead Isotopes in Mineral Exploration.* Elsevier, Amsterdam.

Hart, S.R., Shimizu, N. and Sverjensky, D.A. (1981). Lead isotope zoning in galena: an ion microprobe study of a galena crystal from the Buick Mine, Southeast Missouri. *Economic Geology* **76** 1873–1878.

Holmes, A. (1946). An estimate of the age of the Earth. *Nature* **157** 680–684.

Knapp, A.B., Muhly, J.D. and Muhly, P.M. (1988). To hoard is human: the metal deposits of LC IIC – LC III. *Report of the Department of Antiquities of Cyprus*, 233–262.

Knapp, A.B. and Cherry, J.F. (1994). *Provenance Studies and Bronze Age Cyprus: Production, Exchange and Politico-Economic change.* Monographs in World Archaeology 21, Prehistory Press, Madison.

Littlefield, T.A. and Thorley, N. (1979). *Atomic and Nuclear Physics.* Van Nostrand Rheinhold, New York (3rd edn.).

Muhly, J.D. (1991). The development of copper metallurgy in Late Bronze Age Cyprus. In *Bronze Age Trade in the Mediterranean*, ed. Gale, N.H., Studies in Mediterranean Archaeology 90, Paul Åström's Förlag, Jønsered, pp. 180–196.

Mulliken, R.S. and Harkins, W.D. (1922). The separation of isotopes. Theory of resolution of isotopic mixtures by diffusion and similar processes. Experimental separation of mercury by evaporation in a vacuum. *Journal of the American Chemical Society* **44** 37–65.

Pernicka, E. (1993). Comments on P. Budd, D. Gale, A.M. Pollard, R.G. Thomas and P.A. Williams 'Evaluating lead isotope data: further observations', *Archaeometry*, **35 (2)** (1993), and reply. Comments ...III. *Archaeometry* **35** 259–262.

Pernicka, E and Bachmann, H.G. (1983). Archäometallurgische unter-

suchungen zur antiken silbergewinnung in Laurion III. Das verhalten einigerv spurenelemente beim abtreiben des bleis. *Erzmetall* **36** 592–597.

Pernicka, E., Seeliger, T.C., Wagner, G.A., Begemann, F., Schmitt-Strecker, S., Eibner, C., Öztunali, Ö. and Baranyi, I. (1984). Archäometallurgische untersuchungen in Nordwestanatolien. *Jahrbuch des Römisch-Germanisches Zentralmuseums Mainz* **31** 533–599.

Pollard, A.M. (1982). A critical study of multivariate methods as applied to provenance data. In *Proceedings of the 22nd Symposium on Archaeometry, University of Bradford, 30th March – 3rd April 1982*, eds. Aspinall, A. and Warren, S.E., University of Bradford Press, Bradford, pp. 56–66.

Reedy, T.J. and Reedy, C.L. (1992). Evaluating lead isotope data: comments on E.V. Sayre, K.A. Yener, E.C. Joel and I.L. Barnes, 'Statistical evaluation of the presently accumulated lead isotope data from Anatolia and surrounding regions'. *Archaeometry*, **34 (1)** (1992), 73–105, and reply. Comments...IV. *Archaeometry* **34** 327–329.

Richards, R.H. (1909). *A Textbook of Ore Dressing.* McGraw-Hill, New York.

Russell, R.D. and Farquhar, R.M. (1960). *Lead Isotopes in Geology.* Interscience Publishers, New York.

Sayre, E.V., Yener, K.A., Joel, E.C. and Barnes, I.L. (1992). Statistical evaluation of the presently accumulated lead isotope data from Anatolia and surrounding regions. *Archaeometry* **34** 73–105.

Scaife, B. (1993). *Lead Isotope Analysis and Archaeological Provenancing.* Unpublished BSc dissertation, Department of Archaeological Sciences, University of Bradford.

Scaife, B., Budd, P., McDonnell, J.G., Pollard, A.M. and Thomas, R.G. (in press). A reappraisal of statistical techniques used in lead isotope analysis. Paper presented at *Archaeometry '94*, Ankara, Turkey, 9th-14th May 1994 (Proceedings to be published).

Stacey, J. S. and Kramers, J.D. (1975). Approximation of terrestrial lead isotope evolution by a two-stage model. *Earth and Planetary Science Letters* **26** 207–221.

Stos-Gale, Z.A. and Gale, N.H. (1990). The role of Thera in the Bronze Age trade in metals. In *Thera and the Aegean World III. Volume 1 Archaeology*, ed. Hardy, D.A., Thera Foundation, London, pp. 72–92.

Stos-Gale, Z.A., Gale, N.H. and Zwicker, U. (1986). The copper trade in the South-east Mediterranean region. Preliminary scientific evidence. *Report to the Department of Antiquities of Cyprus 1986*, pp. 122–144.

Thomas, R.G. (1990). *Studies of Archaeological Copper Corrosion Phenomena.* Unpublished Ph.D. thesis, School of Chemistry and Applied Chemistry, University of Wales College of Cardiff.

Tylecote, R.F. (1976). *A History of Metallurgy*. Metals Society, London.

Wagner, G.A., Pernicka, E., Seeliger, T.C., Lorenz, I.B., Begemann, F., Schmitt-Strecker, S., Eibner, C. and Öztunali, Ö. (1986). Geochemische und isotopisch Charakteristika fruher Rohstoffquellen für Kupfer, Blei, Silber und Gold in der Turkei. *Jahrbuch des Römisch-Germanischen Zentralmuseums Mainz* **33** 723–752.

Chapter 10

Summary – Whither Archaeological Chemistry?

HISTORICAL SUMMARY

The application of analytical chemistry to archaeological artefacts and sites has a long and distinguished history, going back well into the 18th Century. The pace of this activity has increased rapidly since about 1950, reflecting a growth in the availability of instrumental methods of analysis, starting with optical emission spectroscopy and neutron activation analysis, and culminating in the application of the latest techniques for sub-nanogram analysis of biochemical remains. The range of materials studied has varied enormously, ranging from the more obvious such as obsidian, flint, glass, and pottery, to skin, hair, and faint traces of putative blood on stone surfaces. Only a few have been considered in any detail here.

The analysis of archaeological material has, in general, been regarded as a specialist pursuit. Several universities and museums have gradually established well-equipped analytical laboratories to deal with some, but inevitably not all, of the materials which are encountered archaeologically. These laboratories have been able to dedicate many of the smaller pieces of analytical equipment (such as AAS, XRF, up to SEM and, occasionally, mass spectrometers) solely to the study of archaeological material. Larger pieces of equipment have tended to be shared, or used on a commercial basis, and there is every indication that this is becoming increasingly common. The justification for this level of instrumental dedication has traditionally been the special nature of archaeological material – often small, occasionally valuable, but mostly precious in the sense that it has to be returned to a museum or reference collection essentially undamaged, since it may be required for other types of study. In the early days of instrumental analysis, this was certainly a severe

restriction – sample requirements were usually measured in terms of grams, and instrumental modifications were often necessary to accommodate smaller samples. Great value was placed on techniques such as XRF, which could be used in a virtually non-destructive mode, but at the cost of restricting the analysis to surface layers.

The past five to ten years have seen a great change in analytical capabilities. Most modern instruments, whatever the technique involved, routinely deal with samples in the milligram range or smaller, and are designed to run under computer control automatically, theoretically giving astonishing throughput figures. Additionally, other scientific disciplines, such as forensic science, or biochemical palaeontology, have developed an interest in handling samples with characteristics very similar to those of archaeological material. In fact, many disciplines have always had similar requirements to archaeology, but only belatedly has this come to be recognized by either side. It is probably now true to say that archaeological chemistry is no longer a unique area of endeavour, requiring specialist equipment, although the problems to be solved often remain specific to archaeology.

One overriding concern in archaeological chemistry is the degraded or altered nature of the sample. Post-depositional alteration cannot be discounted in the discussion of any material, even the most apparently inert, such as stone or gold. The increasing focus on material of a biological or organic nature increases this concern. It simply cannot be assumed that any such material will have survived unaltered over archaeological time. The general rule has to be that archaeological samples are altered in some way from their original state. This gives rise to the conviction which has governed the compilation of this book, that a full understanding of the archaeological material can only be obtained from a thorough knowledge of the structure and behaviour of the material concerned. In this respect, for most materials, we still have a very long way to go.

THE ARCHAEOLOGICAL RELEVANCE OF CHEMICAL APPLICATIONS

An obvious and reasonable question to ask at this stage is 'what has been the value of all this effort?' In terms of the study of archaeological materials, a great deal has been learnt about subjects such as the history of technology, the exploitation of raw materials in antiquity, and the long term stability of such materials. Without doubt, our knowledge of ancient materials has been vastly enhanced by this cumulative effort. But what of the archaeological benefits? If one asks the question 'what has

been the greatest impact of the natural sciences on archaeology over the past fifty years?', without doubt the predominant answer would be 'the development of scientific techniques of dating', of which by far the most important has been radiocarbon. Chemical studies, probably the second most important contribution, have traditionally concentrated largely on the field of 'provenance studies', which has been justified in terms of reconstructing ancient trade and exchange patterns. The extent to which these have contributed to the wider archaeological narrative is difficult to assess, but in some geographical areas (such as the Mediterranean) the impact appears to have been considerable. The recognition of archaeological science in general, and archaeological chemistry in particular, as a valuable component of archaeological research is a tribute to the success of the pioneers in this field.

Nevertheless, there has been a constant debate within the discipline of archaeology about the value of the scientific approach to understanding material culture. Undoubtedly, this debate is clouded by issues of cultural identity (in the sense of C.P. Snow's 'two cultures') and financial considerations, but it is clearly a matter of great concern when a leading archaeologist can use the phrase '*Why archaeologists don't care about archaeometry*' in the title of a book review (Dunnell, 1993). It has to be said that in many cases traditional scientific applications to archaeology have failed to deliver answers to the questions which are of interest to mainstream archaeologists. The reasons for these failings are many – perhaps the original question was framed in terms having little archaeological relevance, or, quite commonly, the scientific interpretations have failed to take into account other evidence, or, more generally, have not been integrated within an appropriate theoretical framework. In this respect archaeological chemistry differs significantly from almost all other applications of chemistry – the results can only be fully utilized if they harmonize with other studies of the social and cultural context of the archaeological problem. In short, human behaviour is almost always involved somewhere along the line.

On the other hand, some archaeologists can be criticized for (understandably, perhaps) focusing on the archaeological value of scientific studies of archaeological material. Many of the greatest benefits of archaeological science have been felt outside the strictly archaeological sphere. A classic example might be radiocarbon dating itself, which has revolutionized late Quaternary geology, but has also provided a wealth of data on the changes in solar flux over the past 10 000 years, which has been of great value to solar scientists. There is a strong view in some quarters that archaeology (or at least certain aspects of it) should recognize this commonality of interest and take its place amongst the 'his-

torical sciences', such as Quaternary geology, palaeobiology, and historical environmental sciences (Pollard, 1995).

More specifically, there is a growing feeling that archaeological chemistry has unconsciously limited its usefulness to archaeology by focusing excessively on questions of provenance (Budd *et al.*, in press). Mainstream archaeology re-evaluated and largely discarded the simplistic notions of diffusionism as an explanation for cultural development during the 1960s and 70s, and chemical provenance studies played a significant role in this process. Despite contributing to the fall of diffusionism, archaeological chemistry has largely failed to move into the more complex theoretical frameworks which have grown to replace it. In many ways, the concept of provenance in its simplest sense is largely irrelevant – a knowledge of where some raw material or traded object came from is of limited value, unless questions about the social and economic structure of the supply and exchange of such materials can be addressed using these data. At the very least, it becomes a minor component of the discussion. We are beginning to see a reconciliation between scientific data and past social and economic activity (*e.g.*, Beck and Shennan, 1991; Bradley and Edmonds, 1993; Budd *et al.*, 1995). Such developments, and they are now increasing in number, will go some way towards countering the negative view of archaeological chemistry in some quarters.

WHITHER ARCHAEOLOGICAL CHEMISTRY?

Archaeological chemistry is now undoubtedly capable of utilizing many of the most sophisticated analytical techniques currently available. Sampling restrictions no longer need cause undue worries on the part of museum curators (although small scale sampling poses questions of analytical homogeneity, *etc.*). Although expensive in archaeological terms, good quality analytical facilities are widely available. The real restrictions, therefore, to good archaeological chemistry are now more in terms of ideas rather than practicalities. Careful construction of relevant archaeological questions, and intelligent interpretation of results within a sound theoretical framework, should lead to the better integration of chemical studies within archaeology. In terms of materials conservation, chemistry has a tremendous role to play in understanding the mechanisms of corrosion for a vast range of materials, and devising strategies for the control of deterioration so that the objects (or buildings!) can be stabilized for display, study, and storage. In many ways, the archaeological demand for qualified chemists or archaeologists with considerable chemical knowledge has never been greater.

It is, however, no longer acceptable (if indeed it ever was) to consider the chemistry of archaeological remains in isolation from other sources of information. A good example of the way which archaeology, chemistry, and geology must interact to answer 'real' questions is provided by the long-standing debate about the origin and transport of the bluestones at Stonehenge (Williams-Thorpe and Thorpe, 1992). It has been established beyond reasonable doubt by petrology and geochemistry that the geological origin of these stones lies in the Preseli Hills of West Wales. What is at issue is how the stones were transported to Stonehenge on Salisbury Plain. The traditional view is that human action some 4 000 years ago was responsible for the movement of the stones, although archaeological evidence for long-distance transport of the megaliths used in other stone circles is insubstantial (Thorpe and Williams-Thorpe, 1991). Recent studies, involving petrology, geochemistry, and evaluation of glacial geomorphology has suggested that glacial action around 400 000 years ago may have been the agent by which large quantities of rock was transported to southern England. The archaeological implications of these two competing models are immense, in terms of the way we interpret the significance of this major prehistoric monument. What is abundantly clear is that no one approach can possibly hope to provide a definitive answer.

There is, however, an even greater challenge for archaeological chemistry. In most countries, archaeological research over the past few years has moved away from the policy of excavation for sites threatened by development, erosion, or sheer population pressure. Many countries have adopted a legal framework of 'preservation by burial'. Planning regulations require mitigation strategies such that development can occur in such a way that known archaeological deposits are not damaged by building work. It is assumed that the safest way of preserving archaeological remains is to leave them in the ground, since they have survived, in some cases, for several millennia in that state. It is recognized, however, that development may change the local burial conditions in many ways – perhaps significantly altering the water table, or the flow of water through the deposits, or increasing the compaction of the soil and thus altering the E_h balance. The question which is now being asked within the legal framework of the planning regulations is 'what effect will these changes have on the state of preservation of the buried material?'. In some case, we can make educated guesses. Considerable effort has, for instance, been put into the study of the preservation environment of waterlogged wood, so that we have some idea of the effect of drying on such objects. More generally, however, we know very little about the detailed effects of variations in burial conditions on a wide range of

materials, such as bone, metalwork, *etc.* This requires a detailed knowledge of the deterioration mechanisms of the materials themselves, but also an ability to predict the changes arising from variations in soil conditions. This requires an understanding of the soil (strictly, burial medium)/groundwater/archaeological object interaction, which involves a very wide range of chemical and physical understanding. Such studies are likely to move rapidly to the top of the political and scientific agenda in the next few years, and form, we believe, a very great challenge and opportunity for archaeological chemistry.

REFERENCES

Beck, C.W. and Shennan, S. (1991). *Amber in Prehistoric Britain.* Oxbow, Oxford.

Bradley, R. and Edmonds, M. (1993). *Interpreting the Axe Trade: Production and Exchange in Neolithic Britain.* Cambridge University Press, Cambridge.

Budd, P., Pollard, A.M., Scaife, B. and Thomas, R.G. (1995). Oxhide ingots, recycling and the Mediterranean metals trade. *Journal of Mediterranean Archaeology* **8** 1–32.

Budd, P., Haggerty, R. Pollard, A.M. Scaife, B. and Thomas, R.G. (in press). Rethinking the quest for provenance. *Antiquity.*

Dunnell, R.C. (1993). Why archaeologists don't care about archaeometry. *Archeomaterials* **7** 161–165.

Pollard, A.M. (1995). Why teach Heisenberg to archaeologists? *Antiquity* **69** 242–247.

Thorpe, R.S. and Williams-Thorpe, O. (1991). The myth of long-distance megalith transport. *Antiquity* **65** 64–73.

Williams-Thorpe, O. and Thorpe, R.S. (1992). Geochemistry, sources and transport of the Stonehenge bluestones. In *New Developments in Archaeological Science,* ed. Pollard, A.M., Proceedings of the British Academy 77, Oxford University Press, Oxford, pp. 133–161.

Appendix 1

The Structure of the Atom, and the Electromagnetic Spectrum

In the early part of this century, a simple model of atomic structure became accepted, now known as the Bohr model of the atom. This stated that most of the mass of the atom is concentrated in the *nucleus*, which consists of *protons* (positively charged particles) and *neutrons* (electrically neutral particles, of approximately the same mass). The number of protons in the nucleus is called the *atomic number*, and essentially defines the element. The number of protons plus neutrons in the nucleus is called the *atomic mass* (see Appendix 2 for more discussion of nuclear structure and isotopes). Electrical neutrality in the atom is maintained by a number of negatively-charged electrons (equal numerically to the number of protons, but with only a fraction of the mass) circling the nucleus like a miniature solar system. This gives a picture of the atom as a very small dense nucleus at the centre containing most of the mass, surrounded by a nebulous 'cloud' of electrons some distance away (on the atomic scale of things). Rutherford demonstrated that this was a plausible model by firing relatively heavy α particles (the nuclei of the helium atom, with two protons and two neutrons) at thin metal foils. Most passed straight through, indicating that the majority of the atom was empty space. A few were deflected on transit through the foil, but some were reflected backwards, hinting that somewhere within the atom was a small but very heavy core, since it would require a relatively massive object to reflect a fast-moving α particle.

This experiment established the nuclear model of the atom. A key point which is derived from this model is that the electrons circling the nucleus are in fact in fixed stable orbits, just like the planets around the sun. Furthermore, each 'orbital' or 'shell' contains a fixed number of electrons – additional electrons are added to the next stable orbital above that which is full. This stable orbital model is a departure from

classical electromagnetic theory (which predicts unstable orbitals, in which the electrons spiral into the nucleus and are destroyed), and can only be explained by quantum theory. The fixed numbers for each orbital were determined to be two in the first level, eight in the second level, eight in the third level (but extendible to eighteen), and so on. Using this simple model, chemists derived the systematic structure of the Periodic Table, and began to understand factors such as atomic sizes, the shapes of molecules, and the underlying reason behind the observed periodicity of chemical properties. Table A1.1 shows how the electronic configuration (the number and distribution of the orbital electrons) of most of the elements can be understood by simply following the above rules for filling up the successive electron orbitals to keep pace with the increasing number of protons in the nucleus. The orbitals are labelled 1, 2, 3, *etc.*, but as can be seen from Table A1.1, sub-divisions occur increasingly with increasing number. For example, the second orbital is split into two sub-shells, labelled *s* and *p* – the 2*s* shell is full when it contains only two electrons, but the 2*p* shell can accommodate six (making eight in total for the second orbital). This splitting is a result of quantum mechanical considerations, and each sub-shell has a characteristic shape in space (*e.g.*, the *s*-orbitals are spherical, whereas the *p*-orbitals are elliptical), which is important in descriptions of chemical bonding and the shapes of molecules. A full discussion of this fascinating area can be found in any basic inorganic chemistry textbook (*e.g.*, Cotton *et al.*, 1995).

It is important to realize that, in any atom, although the orbital electrons may fill only, say, the first two or three orbitals (the exact number of electrons depending on the number of protons in the nucleus), the other unfilled energy levels still exist, and under certain circumstances an electron from a lower state can be promoted up to one of the unfilled energy levels. For example, the element sodium (Na), with 11 protons in the nucleus, and therefore 11 orbital electrons, has the following electronic configuration: $1s^2 2s^2 2p^6 3s^1$, which means that is has two 1*s*-electrons, two 2*s*-electrons, six 2*p*-electrons, and one 3*s*-electron. This is the lowest energy configuration the neutral sodium atom can have, since all of the possible first and second orbitals are full, and the final spare electron has gone into the 3*s* level, which is the next available orbital. This configuration is termed the *ground state*, and is the configuration listed for each element in Table A1.1. In the case of sodium, it is possible to promote the outer (3*s*) electron up to one of the unfilled higher orbitals, such as the 4*p*, by supplying some energy to this outer electron, although not all 'promotions' are 'allowed' – certain *selection rules* apply, as discussed below. With the outer electron temporarily up in

this higher state, the atom is said to be *excited*, and it is to be expected that the atom will return to its ground state as soon as possible, since this is energetically the most stable condition. It is possible, of course, to remove the electron from the atom completely if enough energy is supplied to do so, when the atom is then said to be *ionized*. Because it is now one electron deficient, it has one unbalanced nuclear positive charge, and hence the resulting ion carries a single positive charge – in this case, it has become the sodium *ion*, symbolized by Na^+.

Although the exact spacing between the *energy levels* (as these electronic orbitals are called) for each atom is different, it is possible to represent schematically the relative sequence of electronic orbitals, since this is the same for all atoms, as shown in Figure A1.1. It can be seen here that complications set in with the fourth set of orbitals, because the 3*d* level has an energy slightly lower than the 4*p* level. This results in a whole series of elements with full 1*s*, 2*s*, 2*p*, 3*s*, 3*p*, and 4*s*-orbitals, but instead of filling the 4*p* level, they 'pause' to fill up the empty 3*d* levels, which can accommodate up to 10 electrons. These are the so-called *d-block* or *transition* metals (Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, and Zn), whose chemical properties are intimately related to the behaviour of these *d* electrons. Some of these properties, such as being highly coloured in the ionic state, are important in archaeological contexts, since these are the natural pigments and colorants (of glass and glazes) in antiquity. Some of these properties are further explained in Chapter 5.

As noted above, not all possible transitions between energy levels are theoretically allowed. Each energy level is uniquely characterized by a set of *quantum numbers*. The integer used to define the energy level in the above discussion (1, 2, 3, *etc.*) is called the *principal quantum number*, *n*. The sub-levels described by the letters (*s, p, d, f, etc.*) are associated with the *second quantum number*, given the symbol *l*, with $l = 1$ synonymous with *s*, 2 = *p*, *etc.* The multiplicity of levels associated with each sub-level (*i.e.*, the number of horizontal lines for each orbital in Figure A1.1) is defined by a third quantum number m_l, which has values $0, \pm 1 \ldots \pm l$. Thus, *s*-orbitals only have one sub-level, *p* orbitals have three (with m_l values 0 and ± 1), *d* has five, *etc.* The selection rules can be most simply stated as follows:

$$\Delta n \geq 1$$
$$\Delta l = \pm 1$$

Thus, transitions such as $2p \rightarrow 1s$, $3p \rightarrow 1s$, $3d \rightarrow 2p$ are allowed, whereas $2s \rightarrow 1s$, $3d \rightarrow 1s$, $3d \rightarrow 3d$ are strictly forbidden. As discussed in Chapter 5, however, transitions within the d-orbitals (so-called *d–d* band transitions) do occur, and are important in the consideration of the colour developed

Table A1.1 *Electronic configuration of the elements. Elements in square brackets (e.g., [He]) may imply that the electronic configuration of the inner orbitals are identical to those of the element in brackets. Thus silver (Ag, atomic number 47) has a configuration of* $[Kr]4d^{10}5s^1$*, which if written out in full would be* $1s^2 2s^2 2p^6 3s^2 3p^6\ 3d^{10} 4s^2 4p^6 4d^{10} 5s^1$*, giving 47 electrons in all. For the heavier elements (atomic number above 55), the alternative notation K,L,M is used to denote the inner shells corresponding to orbitals 1, 2, and 3 respectively. This notation is common in X-ray spectroscopy (see p. 36)*
(Adapted from Lide, 1990)

Element	1*s*	2*s*	2*p*	3*s*	3*p*	3*d*	4*s*	4*p*	4*d*	4*f*	5*s*	5*p*	5*d*	5*f*	5*g*
1 H	1														
2 He	2														
3 Li	[He]	1													
4 Be		2													
5 B		2	1												
6 C		2	2												
7 N		2	3												
8 O		2	4												
9 F		2	5												
10 Ne		2	6												
11 Na			[Ne]	1											
12 Mg				2											
13 Al				2	1										
14 Si				2	2										
15 P				2	3										
16 S				2	4										
17 Cl				2	5										
18 Ar				2	6										
19 K					[Ar]		1								
20 Ca							2								
21 Sc						1	2								
22 Ti						2	2								
23 V						3	2								
24 Cr						5	1								
25 Mn						5	2								
26 Fe						6	2								
27 Co						7	2								
28 Ni						8	2								
29 Cu						10	1								
30 Zn						10	2								
31 Ga						10	2	1							
32 Ge						10	2	2							
33 As						10	2	3							
34 Se						10	2	4							
35 Br						10	2	5							
36 Kr						10	2	6							
37 Rb								[Kr]			1				
38 Sr											2				
39 Y									1		2				
40 Zr									2		2				
41 Nb									4		1				
42 Mo									5		1				
43 Tc									5		2				
44 Ru									7		1				
45 Rh									8		1				
46 Pd									10						
47 Ag									10		1				
48 Cd									10		2				
49 In									10		2	1			
50 Sn									10		2	2			
51 Sb									10		2	3			
52 Te									10		2	4			
53 I									10		2	5			
54 Xe									10		2	6			

Element	K	L	M	4*s*	4*p*	4*d*	4*f*	5*s*	5*p*	5*d*	5*f*	5*g*	6*s*	6*p*	6*d*	6*f*	6*g*	6*h*	7*s*
54 Xe	2	8	18	2	6	10		2	6										
55 Cs									[Xe]				1						
56 Ba													2						
57 La										1			2						
58 Ce							1			1			2						
59 Pr							3						2						
60 Nd							4						2						
61 Pm							5						2						
62 Sm							6						2						
63 Eu							7						2						
64 Gd							7			1			2						
65 Tb							9						2						
66 Dy							10						2						
67 Ho							11						2						
68 Er							12						2						
69 Tm							13						2						
70 Yb							14						2						
71 Lu							14			1			2						
72 Hf							14			2			2						
73 Ta							14			3			2						
74 W							14			4			2						
75 Re							14			5			2						
76 Os							14			6			2						
77 Ir							14			7			2						
78 Pt							14			9			1						
79 Au							14			10			1						
80 Hg							14			10			2						
81 Tl							14			10			2	1					
82 Pb							14			10			2	2					
83 Bi							14			10			2	3					
84 Po							14			10			2	4					
85 At							14			10			2	5					
86 Rn							14			10			2	6					
87 Fr														[Rn]					1
88 Ra																			2
89 Ac															1				2
90 Th															2				2
91 Pa											2				1				2
92 U											3				1				2
93 Np											4				1				2
94 Pu											6								2
95 Am											7								2
96 Cm											7				1				2
97 Bk											9								2
98 Cf											10								2
99 Es											11								2
100 Fm											12								2
101 Md											13								2
102 No											14								2
103 Lr											14				1				2

[Xe] = $1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 5s^2 5p^6$

[Rn] = $1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 4f^{14} 5s^2 5p^6 5d^{10} 6s^2 6p^6$

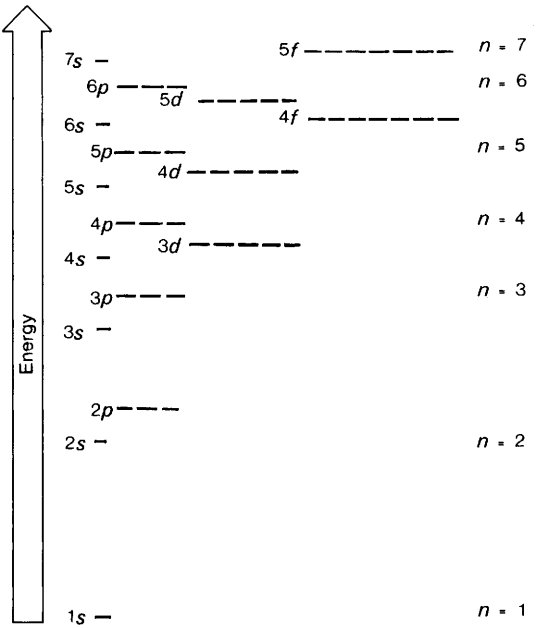


Figure A1.1 *Approximate energy level diagram for electronic orbitals in a multielectron atom. Each horizontal line can accommodate two electrons (paired as so-called spin-up and spin-down electrons), giving the rules for filling the orbitals – two in the s-levels, 6 in the p-levels, 10 in the d-levels. Note that the 3d orbital energy is lower than the 4p, giving rise to the d-block or transition elements* (From Brady, 1990; Figure 7.10)

by transition metals in solution. Other 'forbidden' transitions do also occur.

A second concept of fundamental importance to instrumental chemical analysis is that of the electromagnetic spectrum, and the relationship between particle energy and wavelength. One of the great unifications of Nineteenth Century science was the realization that everything from radio waves down through visible light to *X*-rays were essentially manifestations of the same thing – *electromagnetic radiation*, simply differentiated by having a different wavelength. Thus it was realized that visible light, for example, was simply composed of oscillating electric and magnetic fields, capable of travelling through a vacuum at a particular speed (the speed of light), with a wavelength of between 400 and 700 nm. Figure A1.2 shows the correspondence between wavelength and type of electromagnetic wave. Radio waves have a very long wavelength, with a correspondingly low frequency, which is given by the reciprocal relationship:

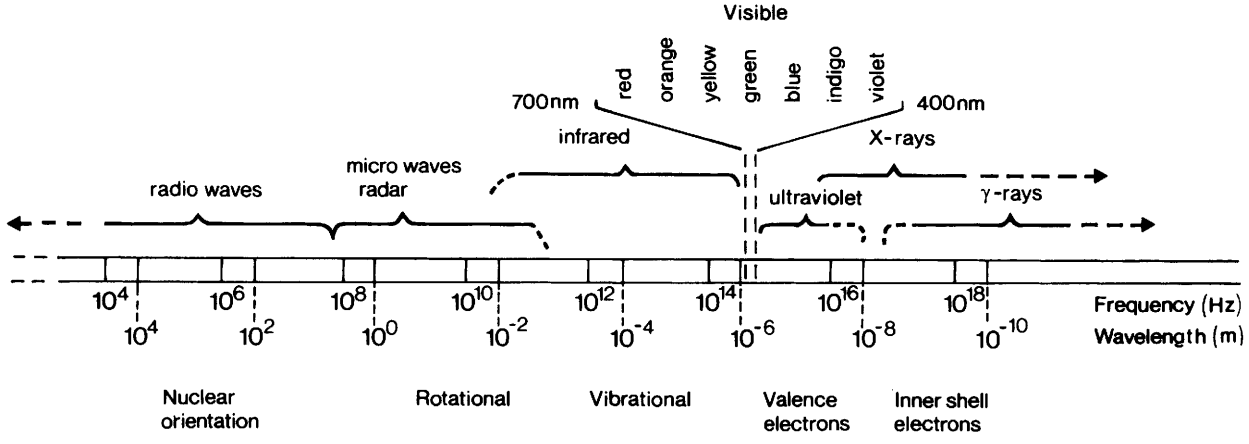


Figure A1.2 *The electromagnetic spectrum, showing the reciprocal relationship between wavelength (in metres) and frequency (in Hertz, or cycles per second). The visible band is a very narrow region, with wavelengths between 400 and 700 nm*
(Adapted from Physical Science Study Committee, 1960; Figure 31-27, reprinted by permission of D.C. Heath and Company)

$$\lambda = c/\upsilon$$

where λ is the wavelength, c is the speed of light (normally taken as 3×10^8 ms^{-1}) and υ (Greek nu) is the frequency. *X*-rays and γ rays have very short wavelengths and very high frequencies. As can be seen, visible light in fact only occupies a very small region of the total electromagnetic spectrum, but is obviously important to us because it is the wavelength range to which our eyes are sensitive.

A second discovery of the early Twentieth Century was less intuitively obvious – that of *particle-wave duality*. Simply put, this states that electromagnetic radiation can be thought of as either radiation characterized by its wavelength (as described above), or as a stream of particles (called *quanta*) with a fixed energy. The energy of the quantum (E) is in units of electron volts (eV), and is proportional to the frequency (or wavelength) of the wave; the relationship involves Planck's constant, h ($= 6.626 \times 10^{-34}$ J s):

$$E = hc/\lambda = h\upsilon$$

This is an elegant explanation for one of the early conundrums posed by quantum theory, which noted that in some experiments visible light behaves as if it were a wave (*e.g.*, Young's slits, which demonstrates diffraction of a wave front; see any classical text on optics, such as Jenkins and White, 1976), but sometimes it exhibits behaviour which can only be explained by assuming that it is manifested as a stream of particles (*e.g.*, the photoelectric effect – see, for example, Caro *et al.*, 1978). It would appear, therefore, that light (or any electromagnetic radiation) can behave as either a particle of energy E or a wave of wavelength λ, depending on the experiment that we carry out to observe it. Particle-wave duality is one of the central tenets of quantum theory.

REFERENCES

Brady, J.E. (1990). *General Chemistry*. John Wiley, New York (5th edn.).

Caro, D.E., McDonnell, J.A. and Spicer, B.M. (1978). *Modern Physics*. Edward Arnold, London (3rd edn.).

Cotton, F.A., Wilkinson, G. and Gaus, P.l. (1995). *Basic Inorganic Chemistry*. John Wiley, New York (3rd edn.).

Jenkins, F.A. and White, H.E. (1976). *Fundamentals of Optics*. McGraw-Hill, New York (4th edn.).

Lide, D.R. (ed.) (1990). *CRC Handbook of Chemistry and Physics*. CRC Press, Boca Raton, Florida (71st edn.).

Physical Science Study Committee (1960). *Physics*. D.C. Heath, Boston.

Appendix 2

Isotopes

Atoms are made up of a positively charged nucleus surrounded by a 'cloud' of orbital electrons carrying an equal negative charge – classically this is envisaged as being similar to the solar system, with the sun occupying the position of the atomic nucleus, and the planets the orbiting electrons. The nucleus contains both positively charged particles (protons) and electrically neutral particles (neutrons), which are roughly the same weight as the protons. This weight is given the value one on the *atomic mass unit* (u) scale, also termed a *Dalton* after the man who put forward the modern atomic theory in the 19th Century. On this classical model, it is impossible to explain the structure of the nucleus itself (this cannot be explained without resorting to quantum mechanics and sub-atomic physics), but it is conventionally stated that the neutrons act as some form of 'glue' which prevents the protons in the nucleus flying apart as a result of electrical repulsion between the positively charged protons.

The number of protons in the nucleus is called the *atomic number*, and is given the symbol Z. It is this number which gives all the elements their different chemical characteristics, and distinguishes one element from another. Z varies continuously from 1, which is the lightest of all elements, hydrogen, up to 103, which is currently the heaviest known element (lawrencium).* Since Z gives the number of units of positive charge in the nucleus, it also dictates the number of electrons orbiting the nucleus. In a neutral atom, the number of electrons is identical to the number of protons in the nucleus, since the charge on the proton and electron is identical but opposite. The familiar chemical symbols (*e.g.*, H, C, O, Pb, *etc.*) are effectively a shorthand code for the proton

* Heavier elements of atomic number 104–110 have also been observed. As this book goes to press, names for elements 104–109 have only been assigned on a provisional basis.

number – thus the symbol Pb stands for 'that element with 82 protons in its nucleus'.

All elements except the simplest form of hydrogen have some neutrons in their nucleus. For the lightest elements, the number of neutrons is the same (or roughly the same) as the number of protons (*e.g.*, helium – He – has two protons and two neutrons), but as the number of protons increases, there appears to be a need for an excess of neutrons to hold the nucleus together. Thus gold (Au, atomic number 79) has 118 neutrons in its nucleus, giving it a total weight on the atomic scale of 197 Daltons. The number of neutrons is given the symbol $\mathcal{N}$, and the combined number of protons plus neutrons is given the symbol A, and is referred to as the *atomic mass number.* Protons and neutrons are collectively termed nucleons, and the following simple relationship holds:

$$A = \mathcal{N} + \mathcal{Z}$$

As stated above, all elements have a unique proton number, but some also have a unique number of neutrons (at least in naturally-occurring forms), and therefore a unique atomic weight – examples are gold ($\mathcal{Z} = 79$, $\mathcal{N} = 118$, giving $A = 197$), bismuth (Bi; $\mathcal{Z} = 83$, $\mathcal{N} = 126$, $A = 209$) and at the lighter end of the scale, fluorine (F; $\mathcal{Z} = 9$, $\mathcal{N} = 10$, $A = 19$) and sodium (Na; $\mathcal{Z} = 11$, $\mathcal{N} = 12$, $A = 23$). Such behaviour is, however, rare in the Periodic Table, where the vast majority of natural stable elements can exist with two or more different neutron numbers in their nucleus. These are termed *isotopes.* Isotopes of the same element have the same number of protons in their nucleus (and hence orbital electrons, and hence chemical properties) but different numbers of neutrons, and hence different atomic weights. That is one reason why the quoted atomic weights of many elements are not whole numbers – they are averages of two or more different isotopes, with the exact value of the average depending on the relative abundance of the various isotopes. For example, copper has a quoted atomic mass of 63.54 Daltons, as a result of having two isotopes – one of mass 63 with a relative abundance of 69%, and one of 65, with an abundance of 31%. It is conventional to denote the various isotopes by using the atomic weight as a preceding superscript – thus these two isotopes of copper are ^{63}Cu and ^{65}Cu. Strictly speaking, the atomic number should also be given as a preceding subscript, but for most purposes that is unnecessary since the symbol 'Cu' is synonymous with the element whose atomic number is 29. It should be noted in passing that even elements which have only one isotope do not have an atomic weight which is exactly a whole number, because the actual mass of a nucleus is never exactly equal to the sum of the masses of its constituents, due

to the relativistic conversion of a small proportion of the mass into energy, known as the *binding energy*.

For the lightest elements, there are usually only two isotopes which are stable, and the lighter of the two is often the most common. Thus carbon has two stable isotopes, ^{12}C (abundance 99%) and ^{13}C (abundance 1%). A third isotope, ^{14}C, is extremely rare (fractional abundance 10^{-12}), radioactively unstable, but, of course, vitally important in archaeology! Similarly, nitrogen has two stable isotopes (^{14}N and ^{15}N, with abundances 99.6% and 0.4% respectively). Oxygen has three stable isotopes (^{16}O, ^{17}O, and ^{18}O, with abundances 99.76%, 0.04%, and 0.20% respectively), and is typical of the slightly heavier (but still 'light') elements. As stated above, the isotopes of the same element chemically all behave identically, but, because they have different weights, processes involving diffusion, or transport across membranes, can all give rise to *fractionation* – a systematic change in the abundance ratio of two isotopes. For these light isotopes, fractionation is particularly marked in biological processes, since the relative weight differences between the individual isotopes is large. Fractionation in light isotopes is now widely used as a natural isotopic marker system in biogeochemical cycles (*e.g.*, Lajtha and Michener, 1994), and has found extensive application in archaeology as a marker for dietary reconstruction (*e.g.*, van der Merwe, 1992). By the time the transition metals, and the heavier elements in general, are reached, the situation becomes more complex, in that some metals have a large number of stable isotopes. Lead has four (^{204}Pb, ^{206}Pb, ^{207}Pb, and ^{208}Pb, with average natural abundances of 1.3%, 26.3%, 20.8%, and 51.5% respectively). The most prolific in this respect is tin (Sn), with ten stable isotopes (112, 114, 115, 116, 117, 118, 119, 120, 122, and 124), with abundances varying from 0.4% up to 33%.

In addition to these stable isotopes, many elements have one or more radioactively unstable isotopes which are produced either as a result of specific nuclear processes (such as ^{14}C, produced by the interaction of cosmic radiation with ^{14}N) or as daughter nuclides during the radioactive decay of heavier unstable elements. Lead, for example, in addition to the four stable isotopes, has at least a further 17 unstable isotopes, ranging from ^{194}Pb up to ^{214}Pb, and with half-lives (see Chapter 9) which vary from 800 ms to 3×10^5 yr (Emsley, 1991). In fact, of the total number of nuclides known (well in excess of 1700), only 260 are stable, suggesting that radioactive instability is the rule rather than the exception. (The term 'stable' perhaps needs some clarification: it is conventional to refer to a number of isotopes as 'stable', when in fact they are known to be radioactively unstable, but with an extremely long half-life. An example is ^{204}Pb, normally thought of as stable, but actually estimated to have a

half-life of 1.4×10^{17} yr, and thus justifiably termed 'stable' on the geological time-scale.) Observations suggest that in order to maintain stability as the atomic number increases, the ratio of neutrons to protons has to increase from 1:1 for the very light elements up to about 3:1 for the heaviest. Too few or too many neutrons leads to nuclear instability. The reasons for this radioactive instability of some nuclei and not others are rather unclear, but it is has been observed that certain combinations of nucleon numbers are more stable than others, giving rise to the concept of *'magic numbers'* in nuclear stability (Faure, 1986; 15). Over half of the stable nuclides known have even numbers of both Z and N. Odd/even and even/odd combinations contribute just over another 100, but only four nuclei with an odd/odd configuration are stable. Furthermore, it has been noted that nuclei which can be imagined as being multiples of the helium nucleus (^{4}He, $A = 2$, $N = 2$), such as ^{12}C, ^{16}O, ^{32}S, and ^{40}Ca, are all particularly stable, which gives some indication of the robustness of the helium nucleus in nuclear physics (see Chapter 9). Standard tables of nuclides exist, for example Littlefield and Thorley (1979), Appendix C, which lists the nuclear con- figuration, natural abundances, decay process, and half-life of all known nuclides, although new information is constantly being added.

REFERENCES

Emsley, J. (1991). *The Elements*. Clarendon Press, Oxford (2nd edn.).

Faure, G. (1986). *Principles of Isotope Geology*. John Wiley, New York (2nd edn.).

Lajtha, K. and Michener, R. (eds.) (1994). *Stable Isotopes in Ecology and Environmental Science*. Blackwell Scientific Publications, Oxford.

Littlefield, T.A. and Thorley, N. (1979). *Atomic and Nuclear Physics*. Van Nostrand Rheinhold, New York (3rd edn.).

van der Merwe, N.J. (1992). Light stable isotopes and the reconstruction of prehistoric diets. In *New Developments in Archaeological Science*, ed. Pollard, A.M., Proceedings of the British Academy 77, Oxford University Press, Oxford, pp. 247–264.

Appendix 3

Fundamental Constants

Description	*Symbol*	*Value*
Avogadro constant	N_A	$6.022169 \times 10^{23}\ \text{mol}^{-1}$
Faraday constant	F	$96486.70\ \text{C mol}^{-1}$
Charge of electron	e	$1.6021917 \times 10^{-19}\ \text{C}$
Mass of electron	m_e	$9.109558 \times 10^{-31}\ \text{kg}$
Mass of proton	m_p	$1.672614 \times 10^{-27}\ \text{kg}$
Mass of neutron	m_n	$1.674920 \times 10^{-27}\ \text{kg}$
Planck constant	h	$6.626196 \times 10^{-34}\ \text{J s}$
Speed of light in vacuum	c	$2.9979250 \times 10^{8}\ \text{m s}^{-1}$
Rydberg constant	R_∞	$1.09737312 \times 10^{7}\ \text{m}^{-1}$
	R_H	$1.09677578 \times 10^{7}\ \text{m}^{-1}$
Bohr magneton	μ_B	$9.274096 \times 10^{-24}\ \text{A m}^2$
Gas constant	R	$8.31434\ \text{J K}^{-1}\ \text{mol}^{-1}$
Boltzmann constant	k	$1.380622 \times 10^{-23}\ \text{J K}^{-1}$

Appendix 4

Atomic Number and Approximate Atomic Weights (based on $^{12}C = 12.000$) of the Elements

Name	*Symbol*	*Atomic Number*	*Atomic Weight*
Actinium	Ac	89	(227)
Aluminium	Al	13	26.98
Americium	Am	95	(243)
Antimony	Sb	51	121.7
Argon	Ar	18	39.94
Arsenic	As	33	74.92
Astatine	At	85	(210)
Barium	Ba	56	137.3
Berkelium	Bk	97	(247)
Beryllium	Be	4	9.012
Bismuth	Bi	83	209.0
Boron	B	5	10.81
Bromine	Br	35	79.90
Cadmium	Cd	48	112.4
Calcium	Ca	20	40.08
Californium	Cf	98	(251)
Carbon	C	6	12.01
Cerium	Ce	58	140.1
Cesium	Cs	55	132.9
Chlorine	Cl	17	35.45
Chromium	Cr	24	52.00
Cobalt	Co	27	58.93
Copper	Cu	29	63.54
Curium	Cm	96	(247)
Dysprosium	Dy	66	162.5
Einsteinium	Es	99	(252)
Erbium	Er	68	167.2
Europium	Eu	63	152.0

Name	*Symbol*	*Atomic Number*	*Atomic Weight*
Fermium	Fm	100	(257)
Fluorine	F	9	19.00
Francium	Fr	87	(223)
Gadolinium	Gd	64	157.2
Gallium	Ga	31	69.72
Germanium	Ge	32	72.59
Gold	Au	79	197.0
Hafnium	Hf	72	178.4
Helium	He	2	4.003
Holmium	Ho	67	164.9
Hydrogen	H	1	1.008
Indium	In	49	114.8
Iodine	I	53	126.9
Iridium	Ir	77	192.2
Iron	Fe	26	55.84
Krypton	Kr	36	83.80
Lanthanum	La	57	138.9
Lawrencium	Lr	103	(260)
Lead	Pb	82	207.2
Lithium	Li	3	6.941
Lutetium	Lu	71	175.0
Magnesium	Mg	12	24.31
Manganese	Mn	25	54.94
Mendelevium	Md	101	(258)
Mercury	Hg	80	200.5
Molybdenum	Mo	42	95.94
Neodymium	Nd	60	144.2
Neon	Ne	10	20.17
Neptunium	Np	93	(237)
Nickel	Ni	28	58.70
Niobium	Nb	41	92.91
Nitrogen	N	7	14.01
Nobelium	No	102	(259)
Osmium	Os	76	190.2
Oxygen	O	8	16.00
Palladium	Pd	46	106.4
Phosphorus	P	15	30.97
Platinum	Pt	78	195.0
Plutonium	Pu	94	(244)
Polonium	Po	84	(209)
Potassium	K	19	39.09
Praeseodymium	Pr	59	140.9
Promethium	Pm	61	(147)
Protactinium	Pa	91	231.0
Radium	Ra	88	226.0
Radon	Rn	86	(222)
Rhenium	Re	75	186.2

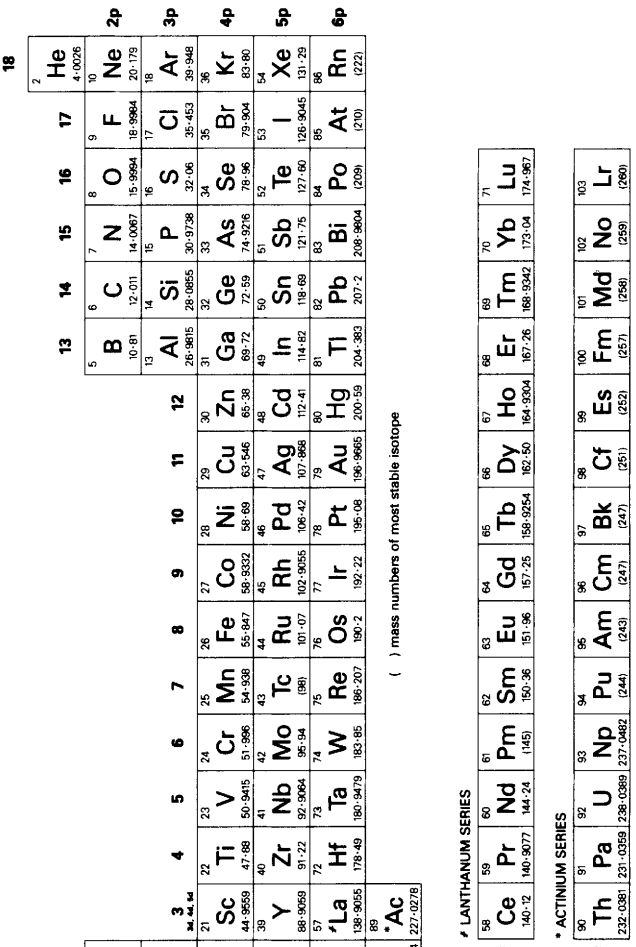
Name	*Symbol*	*Atomic Number*	*Atomic Weight*
Rhodium	Rh	45	102.9
Rubidium	Rb	37	85.47
Ruthenium	Ru	44	101.0
Samarium	Sm	62	150.4
Scandium	Sc	21	44.96
Selenium	Se	34	78.96
Silicon	Si	14	28.08
Silver	Ag	47	107.9
Sodium	Na	11	22.99
Strontium	Sr	38	87.62
Sulfur	S	16	32.06
Tantalum	Ta	73	180.9
Technetium	Tc	43	(98)
Tellurium	Te	52	127.6
Terbium	Tb	65	158.9
Thallium	Tl	81	204.3
Thorium	Th	90	232.0
Thulium	Tm	69	168.9
Tin	Sn	50	118.6
Titanium	Ti	22	47.90
Tungsten	W	74	183.8
Uranium	U	92	238.0
Vanadium	V	23	50.94
Xenon	Xe	54	131.3
Ytterbium	Yb	70	173.0
Yttrium	Y	39	88.91
Zinc	Zn	30	65.38
Zirconium	Zr	40	91.22

Atomic weights in brackets are approximate.

Appendix 5

Periodic Table of the Elements

Reproduced with the kind permission of Glaxo Wellcome plc.



Subject Index

The application of chemistry within archaeology is an important and fascinating area. It allows the archaeologist to answer such questions as "what is this artefact made of?", "where did it come from?", and "how has it been changed through burial in the ground?", providing pointers to the earliest history of mankind.

Archaeological Chemistry begins with a brief description of the goals and history of archaeological science, and the place of chemistry within it. It sets out the most widely used analytical techniques in archaeology and compares them in the light of relevant applications. The book includes an analysis of several specific archaeological investigations in which chemistry has been employed in tracing the origins of or in preserving artefacts. The choice of these investigations conforms to themes based on analytical techniques, and includes chapters on obsidian, ceramics, glass, metals and resins. Finally, the book suggests a future role for chemical and biochemical applications in archaeology in the 1990s.

Archaeological Chemistry enables scientists to tackle the fundamental issues of chemical change in archaeological materials, in order to advance the study of the past. It will prove an essential companion to students in archaeological science and chemistry, field and museum archaeologists, and all those involved in conserving human artefacts.

Mark Pollard is Professor and Head of the Department of Archaeological Sciences at the University of Bradford. He obtained a BA and DPhil in Physics from the University of York, concentrating his doctoral thesis on the chemical analysis of Medieval glass. He spent six years at the University of Oxford's Research Laboratory for Archaeology and the History of Art as Analytical Research Officer. He was appointed "New Blood" Lecturer in Inorganic Chemistry (Archaeology) at the University of Wales, Cardiff, where he spent a further six years applying chemical techniques to archaeological problems before moving to Bradford in 1990. He is a Fellow of the Society of Antiquaries of London, and a Member of The Royal Society of Chemistry.

Carl Heron is Lecturer in Archaeological Sciences at the University of Bradford. After graduating in Archaeological Sciences, he gained research experience in the Department of Chemistry at the University of Wales, Cardiff, and in the Department of Biochemistry at the University of Liverpool. He has extensive experience in archaeological fieldwork. His specialist interests include organic analysis, particularly the applications of gas chromatography and mass spectrometry to biological materials, as well as chemical and geophysical prospection. He is also interested in the historical relationship between archaeology and the 'hard' sciences.

ISBN 0-85404-523-6

THE ROYAL SOCIETY OF CHEMISTRY
Information Services