19 Properties of Molecular Surfaces

learning objectives:

- color-coding of electrostatic potentials on the van der Waals surface of a molecule
- electrostatic potentials, like any other molecular surface property, can be mapped onto a 2-D plane by a Kohonen neural network
- different conformations will show different Kohonen electrostatic potential maps
- the Kohonen maps of the electrostatic potentials can be used for finding molecular similarities, e.g., in substrates binding to the same receptor
- **artificial** neural networks can be used to eludicate chemical processes in **biological** neural networks
- Kohonen networks can be used to study the shape of molecules
- Kohonen networks can help in the search and optimization of pharmaceutical lead structures

19.1 The Problems

The shape of a molecule, and properties on molecular surfaces such as hydrophobicity, hydrogen bonding potential, and the electrostatic potential are of profound influence on many physical, chemical, or biological properties. A ligand must have a certain shape in order to fit into a receptor, and the properties on the surface of a ligand must correspond to those in the pocket of a receptor.

The study of the geometry of molecular surfaces and of the distribution of electronic and other properties on molecular surfaces may, therefore, give important insights into mechanisms of interactions of molecules and their influence on the properties of compounds.

We will largely concentrate our discussion on the molecular electrostatic potential (MEP) although much of what is being said is also applicable to other molecular surface properties. After the analysis of the molecular electrostatic potential we will turn our attention to the geometric shape of molecular surfaces.

Molecular electrostatic potentials give detailed information for studies on chemical reactivity or pharmacological activity of a compound. The spatial distribution and the values of the electrostatic potential determine the attack of an electrophilic or nucleophilic agent as the primary event of a chemical reaction. By the same token, the three-dimensional distribution of the electrostatic potential is largely responsible for the binding of a substrate molecule at the active site of a biologically active receptor.

The *electrostatic potential* at a certain point near a molecule is the work involved when a unit positive charge is brought from infinity to this point. It can be calculated by quantum mechanical procedures of various degrees of sophistication, or by a simple point-charge model if the atoms of a molecule have been assigned partial charge values. Negative values of the electrostatic potential indicate attraction for a positive charge, positive values stand for repulsion.

The three-dimensional nature of the electrostatic potential makes it difficult to simultaneously visualize its **spatial distribution** and its **magnitude**.

To simplify matters, in the rest of the study we will deal with only the electrostatic potential on the van der Waals surface of a molecule. This part of the total electrostatic potential is the most important part as molecules come into contact with other molecules at the van der Waals surface. It is calculated by a point-charge model at points evenly distributed over this surface. The partial charges on the atoms are calculated by iterative partial equalization of orbital electronegativity (PEOE), a well-established empirical method for the rapid calculation of charge distributions (see References 19-1 and 19-2).

It has become customary to color-code the **magnitude** of the electrostatic potential, showing values from negative to positive as colors from red through yellow and green to blue.

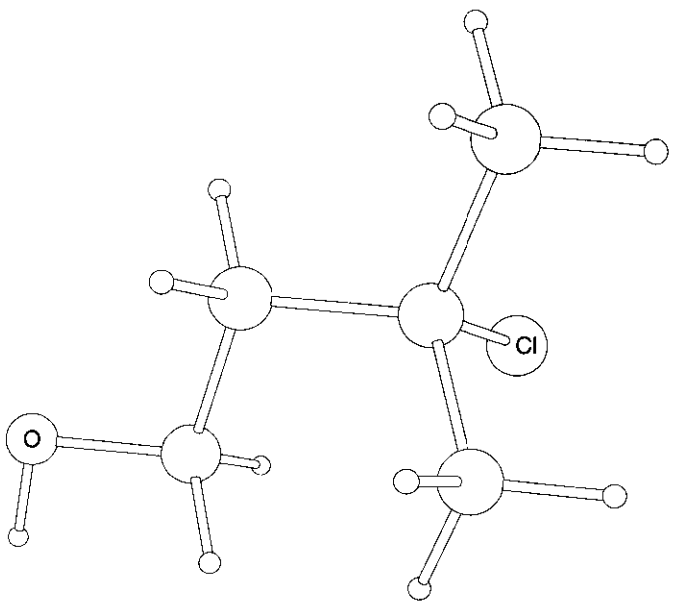

Figure 19-1: Ball and stick model of *3-chloro-3-methylbutan-1-ol.*

The next problem is how to represent the **spatial distribution** of the electrostatic potential. This is usually done by choosing an observation point away from the molecule and showing, on a graphics screen, whatever part of the van der Waals surface can be seen from this point. Basically, this means that a parallel **linear projection** of the molecular electrostatic potential is performed, indicating only that side of the molecule visible from the observation point.

Figure 19-1 shows a three-dimensional model of *3-chloro-3-methylbutan-1-ol* viewed from an observation point from which the chlorine atom is hardly visible.

The electrostatic potential for this molecule is then calculated by moving a unit positive charge across the van der Waals surface and calculating at evenly distributed points the electrostatic interactions with all the atoms that bear partial charges.

The electrostatic potential on the van der Waals surface is shown in Figure 19-2 from the same observation point as the one chosen (Figure 19-1) for looking at the ball and stick model. The positive charge-attracting (nucleophilic) site corresponding to the oxygen atom can clearly be seen by the red color (negative electrostatic potential).

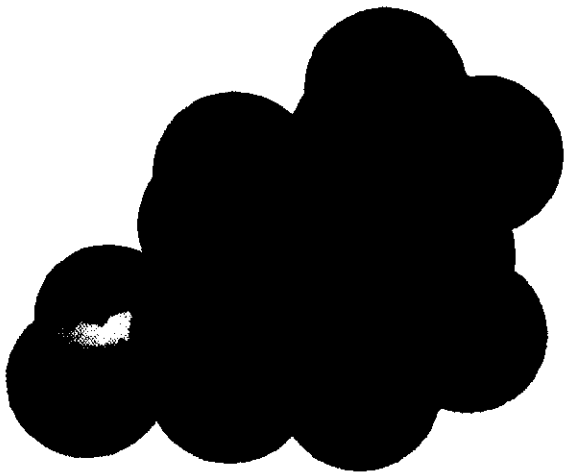

Figure 19-2: View (linear projection) of the molecular electrostatic potential of *3-chloro-3-methylbutan-1-ol* on the van der Waals surface from the same observation point as in Figure 19-1.

On the other hand, there is also a strongly negative electrostatic potential at the surface of the chlorine atom, but this cannot be seen from this viewpoint.

With a linear projection a large part of the van der Waals surface **cannot** be seen from the observation point. Thus, a **series** of observation points and associated projections of the color-coded van der Waals surface have to be chosen to get the overall picture of the electrostatic potential.

It would be quite helpful if the entire electrostatic potential of the molecule on the van der Waals surface could be shown in one picture. Clearly, this requires a **nonlinear** projection method; a Kohonen network can do this for us.

19.2 The Network Architecture and Training

In Section 6.4, we saw how a two-dimensional map of the surface of a sphere can be obtained by training a Kohonen network with the three Cartesian coordinates of a series of points taken from that surface.

Basically, the same procedure is followed in this example to generate a map of the distribution of the electrostatic potential; 20,000 points are randomly selected from the van der Waals surface and a (60 x 60) Kohonen network is trained with the three Cartesian

coordinates of each of these points. Thus, each neuron in the network has three weights.

Recall that a Kohonen network undergoes unsupervised training. Hence, the value of the electrostatic potential is not used in the training phase.

The plane of projection is the surface of a torus. For any point of the van der Waals surface that enters the learning phase, the neuron with the three weights **most similar** to the input coordinates of this particular point is selected as the winning neuron, and its weights and those of neurons in an appropriate neighborhood are corrected according to Equations (6.6).

When the Kohonen network stabilizes, all neurons are investigated to see which points from the van der Waals surface they contain.

It is found that points which are very close together on the van der Waals surface (and therefore have nearly the same potential) map to the same neuron, and adjacent neurons are excited by points with **similar** electrostatic potentials. Thus, the magnitude of the electrostatic potential of the points that map to a neuron can be used to decide its color. This leads to the color-coded map of the electrostatic potential shown in Figure 19-3.

In this map of the electrostatic potential, the effect of the oxygen and of the chlorine atom can both be seen simultaneously. The oxygen atom corresponds to the yellow and orange spot whereas the chlorine atom is indicated by the large blue-green area. Thus, the Kohonen network has achieved a nonlinear projection of the entire molecular surface and can, therefore, indicate all features of a property, such as the electrostatic potential, on a molecular surface.

As discussed in Section 6.2, the map is projected onto the surface of a torus. In fact, the learning process adjusts the form of the torus as well as possible to the three-dimensional structure of the molecule. This is shown in Figure 19-4.

The torus can be cut open and flattened into a square arrangement of neurons at any two perpendicular lines on the torus. Thus, the map can be shifted right, left, up or down. The two maps shown in Figure 19-5 are just as valid representations of the electrostatic potential as the map of Figure 19-3.

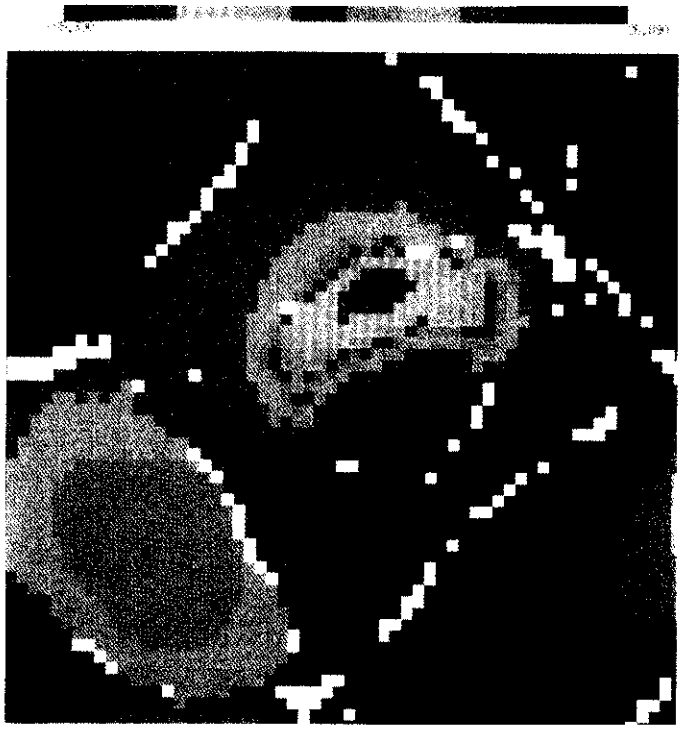

Figure 19-3: Color-coded Kohonen map of the electrostatic potential of *3-chloro-3-methylbutan-1-ol* in the conformation shown in Figure 19-1.

Figure 19-4: Adjustment of the torus used for projection of the surface of the molecule.

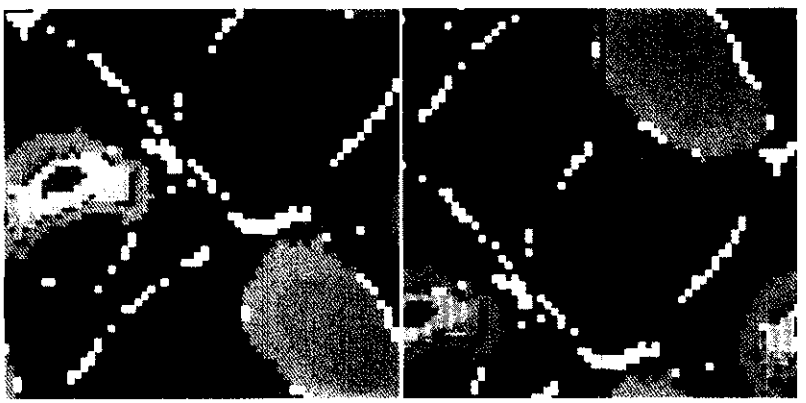

Figure 19-5: Shifted maps of the electrostatic potential of *3-chloro-3-methylbutan-1-ol*.

19.3 Tiling with Kohonen Maps; Conformational Effects

In order to show all of the features of Kohonen maps of the electrostatic potential, it is helpful to put several identical maps together like tiles (see Section 6.4). Figure 19-6 shows the composite of six identical Kohonen maps of the electrostatic potential of *3-chloro-3-methylbutan-1-ol* as shown in Figure 19-3.

The electrostatic potential of a molecule depends on the relative arrangement of the individual atoms. If a molecule has conformational flexibility, the different conformations will result in different distributions of the electrostatic potential on the van der Waals surface; consequently, different conformations give different Kohonen maps of the electrostatic potential.

The Kohonen map of Figure 19-3 was obtained from a conformation of *3-chloro-3-methylbutan-1-ol* having a torsional angle for the sequence C1–C2–C3–Cl of 90° (see Figure 19-1). Figure 19-7 shows a conformation of *3-chloro-3-methylbutan-1-ol* that has a torsional angle of 0° for the sequence C1–C2–C3–Cl and furthermore a hydrogen bridge between the OH group and the chlorine atom.

The molecular electrostatic potential is influenced by the three-dimensional arrangement of atoms and thus also by the conformation of a molecule. This should also be reflected in the Kohonen maps of the molecular electrostatic potential. Figure 19-8 shows the Kohonen map of the electrostatic potential of *3-chloro-3-methylbutan-1-ol* calculated for the conformation shown in Figure 19-7.

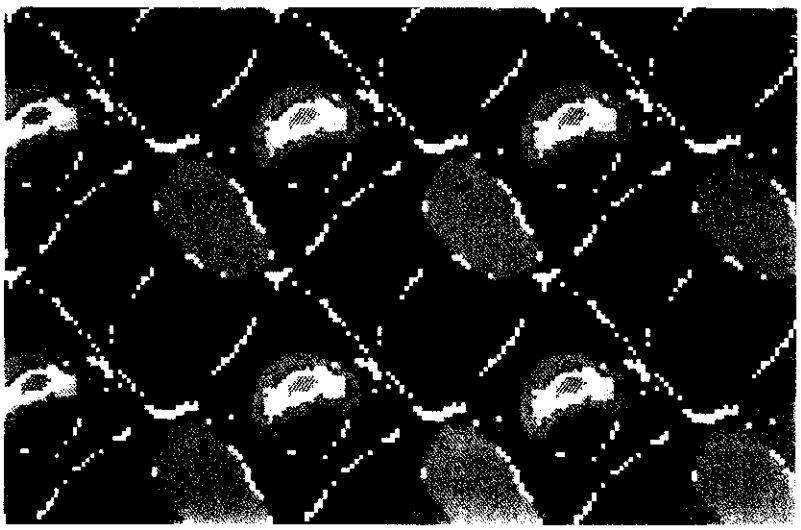

Figure 19-6: Tiling six identical Kohonen maps equivalent to Figure 19-3.

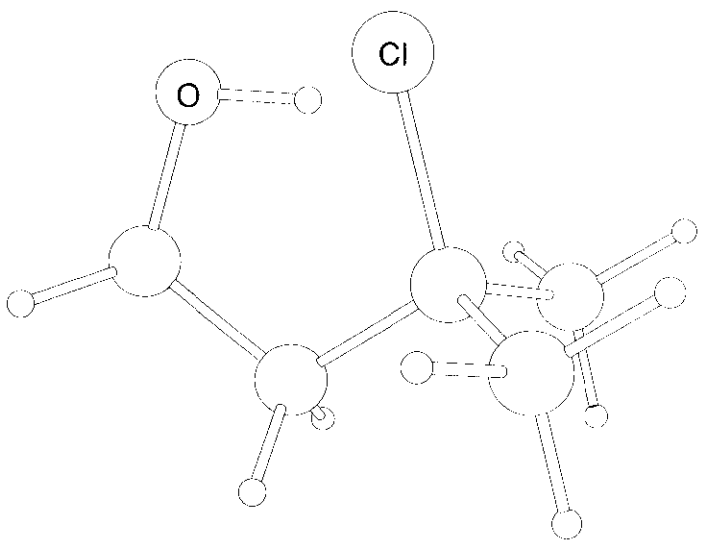

Figure 19-7: Ball and stick model of *3-chloro-3-methylbutan-1-ol* in a conformation different from the one in Figure 19-1.

It can be clearly seen that the maps shown in Figure 19-3 and 19-8 are quite different. In particular, the areas of negative electrostatic potentials resulting from the oxygen and the chlorine atom have now merged due to the close proximity of these two electronegative atoms. Thus, indeed, Kohonen maps of molecular electrostatic potentials give information on the influence of conformation on this property.

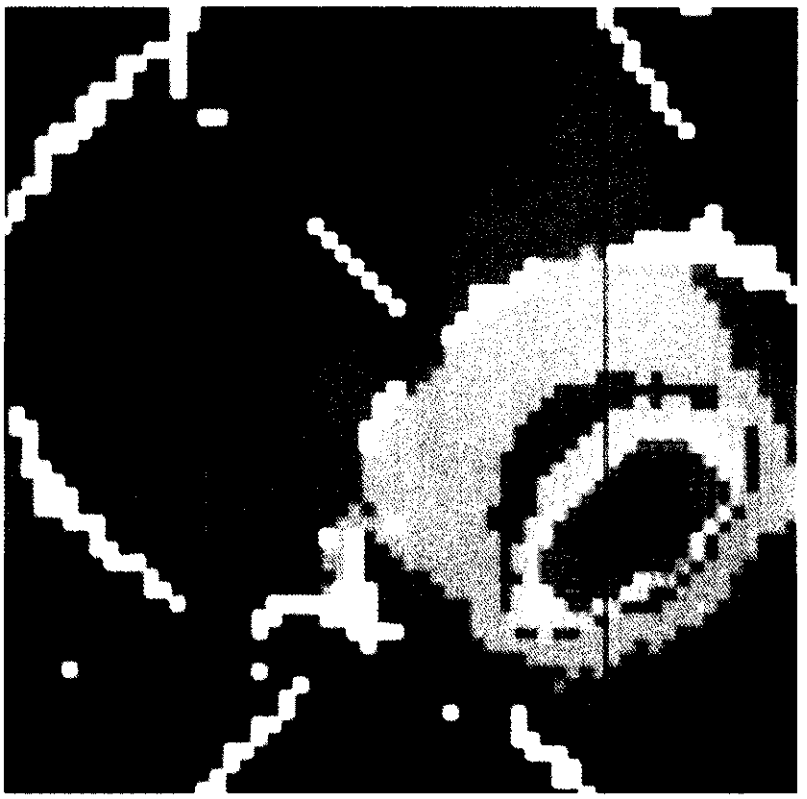

Figure 19-8: Kohonen map of the electrostatic potential of *3-chloro-3-methylbutan-1-ol* in the conformation shown in Figure 19-7.

19.4 Investigation of Receptors of Biological Neural Networks

In the following Section we will show how these Kohonen maps can be used to draw chemical inferences. Since the electrostatic potential of a biologically active molecule plays a major role in substrate binding, Kohonen maps of electrostatic potentials can be valuable for finding similarities in structures that bind to the same receptor. In fact, we will now use an **artificial** neural network (Kohonen network) to shed some light onto chemical compounds involved in **biological** neural networks.

Signals are transmitted **within** a biological neuron electrically, but **between** two different neurons – across the synaptic gap **chemically**. A compound called a *neurotransmitter* is released at the end of an axon, crosses the synaptic gap and initiates another electric signal in the dendrite of a second neuron (Figure 19-9).

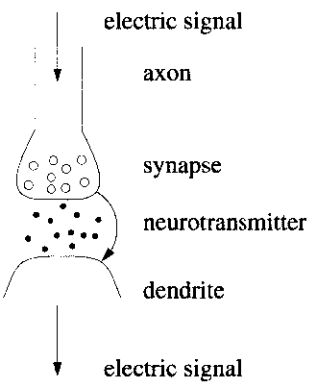

Figure 19-9: The synapse between two neurons.

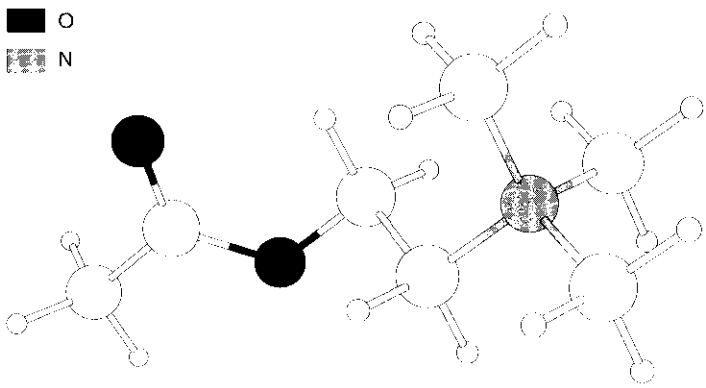

Figure 19-10: Ball and stick model of *acetylcholine*.

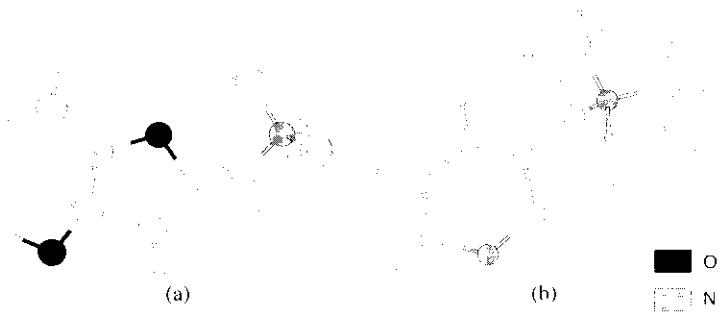

Figure 19-11: Ball and stick model of *muscarine* (a) and of *nicotine* (b).

One of the most important neurotransmitters is *acetylcholine* (Figure 19-10) which can bind to at least to two different receptors, the *muscarinic receptor* and the *nicotinic receptor*.

On the other hand, some molecules can only bind to one or the other of the two receptors: *muscarine* (Figure 19-11a) binds only to the muscarinic receptor, and *nicotine* (Figure 19-11b) binds only to the nicotinic receptor.

Apparently, both *muscarine* and *nicotine* are structurally so rigid, because of their ring systems, that they fit only in one of the receptors. *Acetylcholine*, however, is structurally more flexible; because of its open chain character, it has more conformational degrees of freedom: in one conformation it can bind to the muscarinic receptor, and in another one, to the nicotinic receptor.

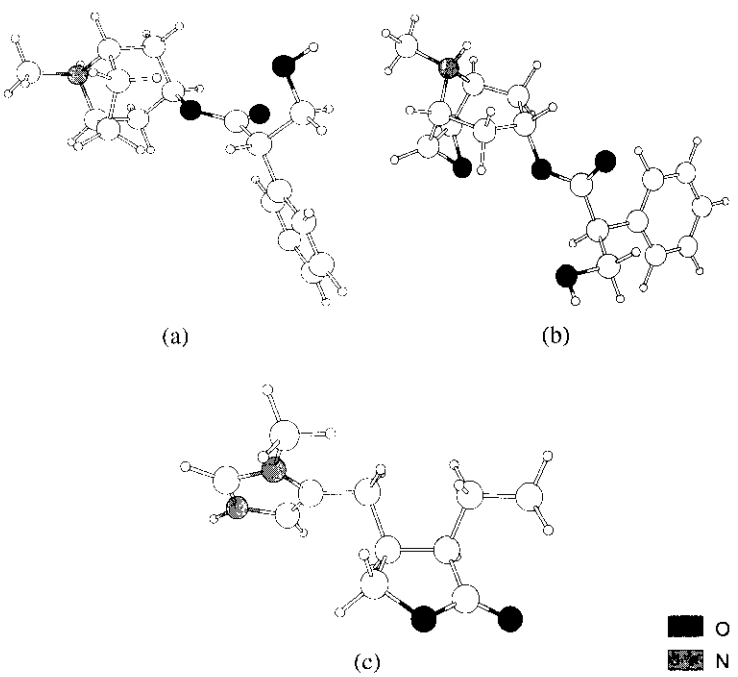

Figure 19-12: Molecules that bind to the muscarinic receptor: *atropine* (a), *scopolamine* (b), and *pilocarpine* (c).

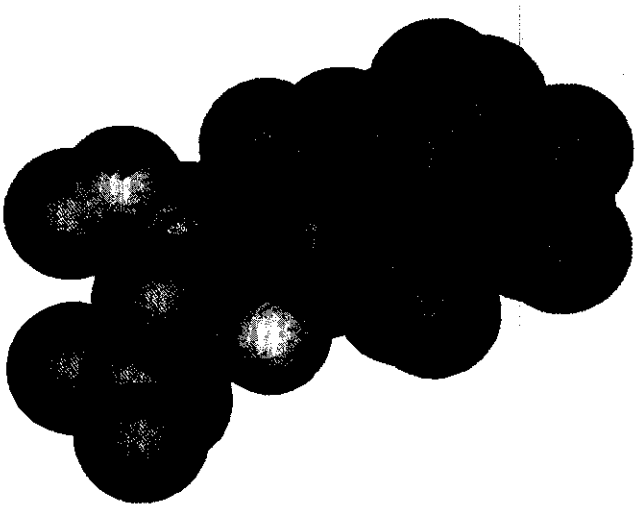

Figure 19-13: Linear projection of the van der Waals electrostatic potential of *muscarine*.

Figure 19-12 shows three more molecules that bind to the muscarinic receptor: *atropine*, *scopolamine*, and *pilocarpine*. Clearly, there is a close structural relationship between *atropine* and

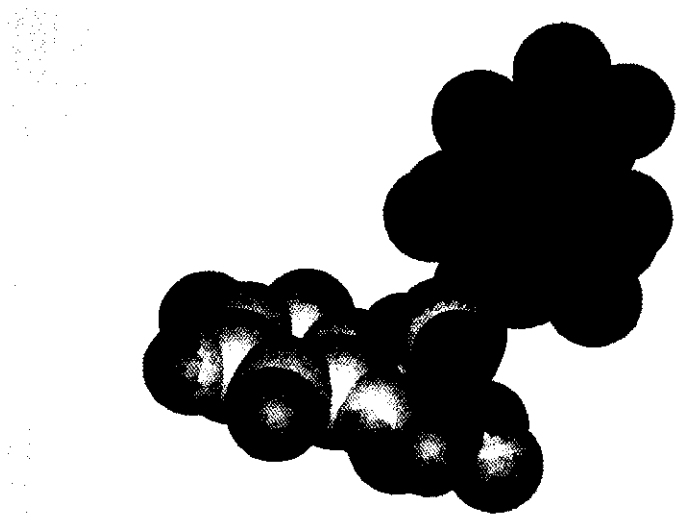

Figure 19-14: Linear projection of the van der Waals electrostatic potential of protonated *scopolamine*.

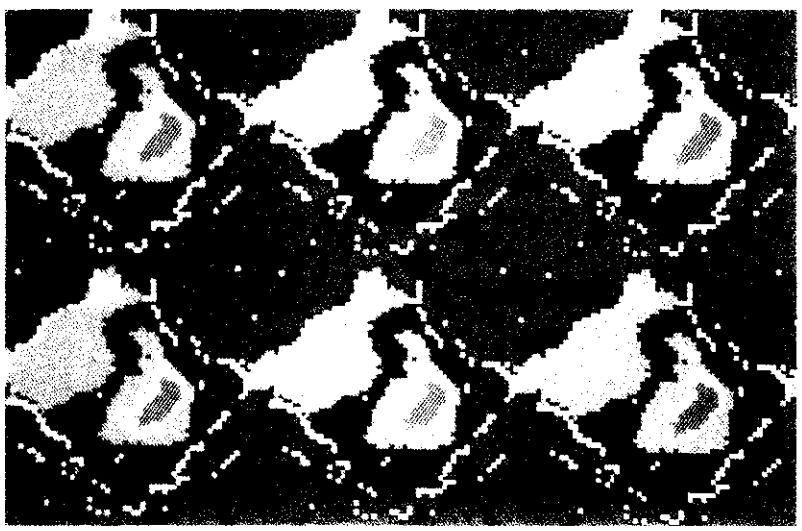

Figure 19-15: Kohonen map of the van der Waals electrostatic potential of *muscarine* (six fold).

scopolamine, but these two structures are quite different from *muscarine* and from *pilocarpine*.

Nevertheless, there must be quite an extensive similarity among these molecules, because all four of them bind to the same receptor. If it is true that the electrostatic potentials of molecules play a major role in binding, the electrostatic potentials of these four molecules should be quite similar. *Acetylcholine* and *muscarine* both have a quaternary nitrogen atom, i.e., a nitrogen atom bearing a positive charge. To enable a correct comparison, the other uncharged molecules have been protonated at the most basic nitrogen before calculating the

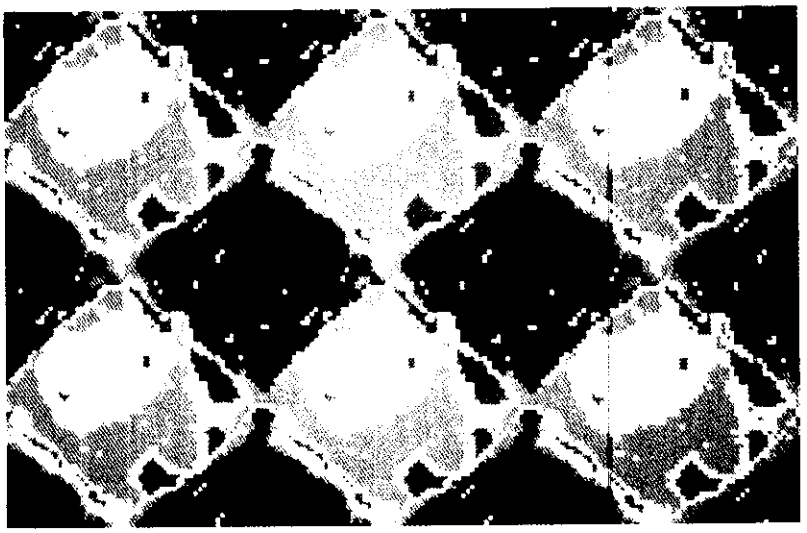

Figure 19-16: Kohonen map of the van der Waals electrostatic potential of *scopolamine* (six fold).

electrostatic potentials. This also etter reflects the situation at the receptor where the molecule surely is protonated. Figure 19-13 and Figure 19-14 show the electrostatic potentials of *muscarine* and of *scopolamine* (in its protonated form).

Clearly, a comparison of the electrostatic potentials of the two molecules suffers from the deficiencies of a linear projection, which shows only that part of the electrostatic potential that can be seen from some chosen viewpoint. In order to compare the entire electrostatic potentials of the two molecules, we have generated the Kohonen maps of both. These are shown in Figures 19-15 and 19-16.

A comparison of Figures 19-15 and 19-16 shows certain similarities in the gross features of the two Kohonen maps. These similarities will become even clearer when the Kohonen maps of the four muscarinic compounds (*muscarine*, *atropine*, *scopolamine*, and *pilocarpine*) are compared with those of the four compounds that bind at the nicotinic receptor (*nicotine*, *anatoxine*, *mecamylamine*, and *pempidine*). This is done in Figure 19-17.

A detailed discussion of the features in these eight Kohonen maps is beyond the scope of this book (for further details see References 19-5 to 19-7). However, it is hoped that the reader realizes and appreciates the similarities in the four Kohonen maps in the top and bottom row, respectively, and perceives the differences between the two groups of maps.

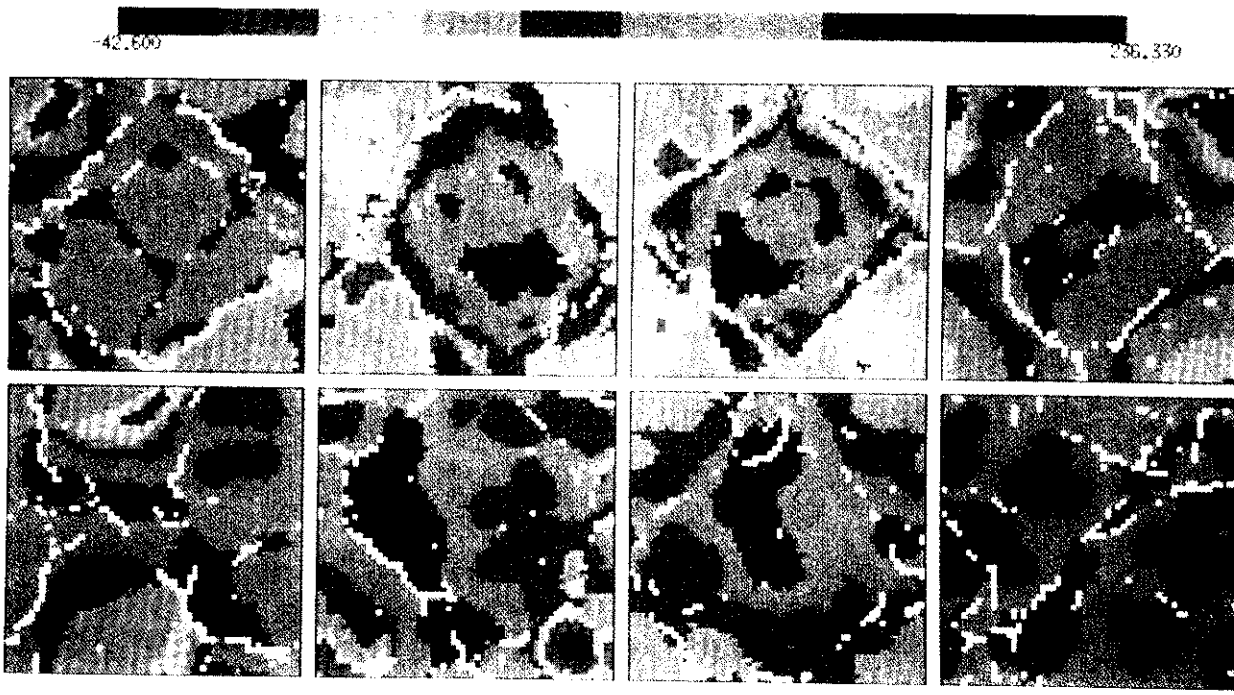

Figure 19-17: Kohonen maps of the electrostatic potentials of four muscarinic compounds (*muscarine*, and protonated *atropine*, *scopolamine* and *pilocarpine*) (top) and four protonated nicotinic compounds (*nicotine*, *anatoxine*, *mecamylamine*, and *pempidine*) (bottom).

The mapping of molecular electrostatic potentials by Kohonen networks is apparently able to extract essential features responsible for the binding of biologically active compounds to their receptors..

We hope that this application of an **artificial neural network** to a study from the field of **biological neural networks** underscores the broad applicability of neural networks. Again, applications have their limits only in the imagination of the potential user.

19.5 Comparison of Kohonen Maps

The example in the previous section has shown the benefits in comparing Kohonen maps of the molecular electrostatic potential. Particularly, the eight maps assembled in Figure 19-17 emphasize how the maps of the molecular electrostatic potential can show similarities of ligands binding to the same receptor. In order to support this comparison, the maps had to be aligned, i.e., the location of the cuts in the torus (cf. Figure 6-7) were chosen such that maximum similarity in the maps could be discerned.

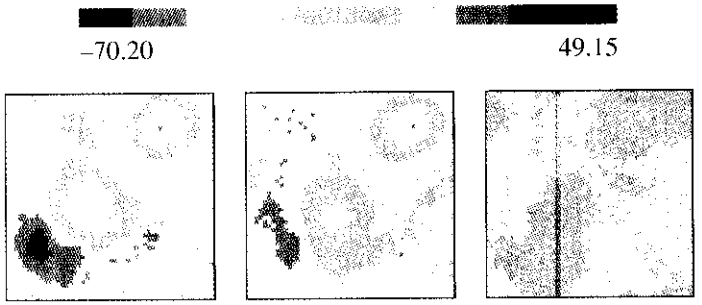

Figure 19-18: Averaged Kohonen maps of the molecular electrostatic potential of steroids binding to the CBG receptor, having (a) high, (b) intermediate, and (c) low binding affinity.

Similarity was assigned, in this case, by human inspection building on the powerful pattern recognition capabilities of the human mind. Clearly, it would be quite desirable to have an automatic procedure for defining or even quantifying similarities in such maps.

We have not yet found a universally convincing method for the calculation of similarity in these maps, as both global and local features of the maps have to be taken into account. Furthermore, the maps may have to be rotated and reflected because the orientation of the features in a map can change with different random initialization of the weights of the network. In one approach, we have analyzed the maps of 31 steroids binding to the *corticosteroid binding globulin* (CBG) receptor (cf. Sections 13.5 – 13.8 and, particularly, Table 13-2). Averaged maps were calculated for the compounds having high, intermediate, or low binding affinities; they are shown in Figure 19-18.

The pattern of the molecular electrostatic potential of the most polar area in the averaged map of the highly active compounds is the most pronounced feature (Figure 19-18a). In the three averaged maps, the distinction of the polar spaces decreases according to decreasing activity of the compounds.

The averaged map of the highly active compounds can be used to build a pharmacophore model. Thus, a comparison of the map of a steroid with the averaged maps allows one to establish whether a molecule belongs to the active or inactive CBG compounds.

The pattern of the molecular electrostatic potential in the Kohonen maps can also be converted into quantitative descriptors. Such descriptors have successfully been used in quantitative structure-activity studies.

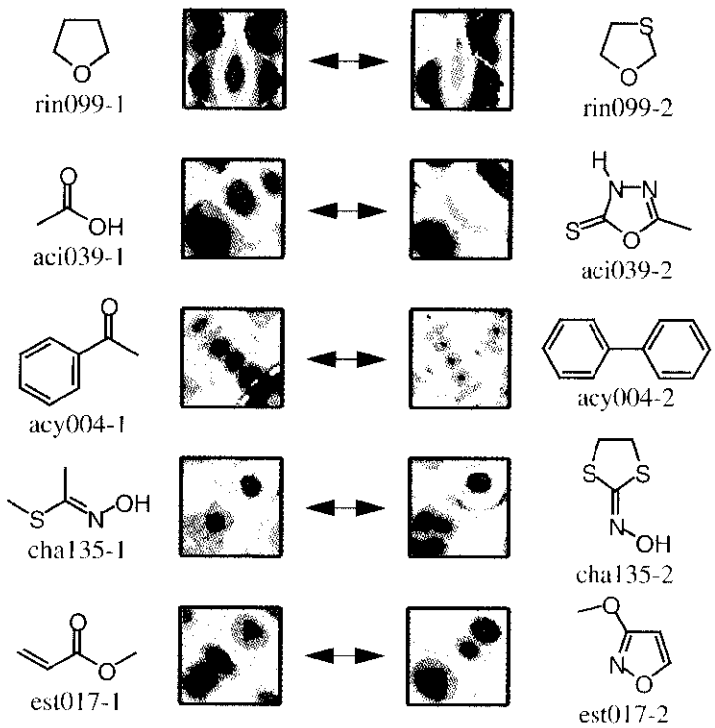

Figure 19-19: Pairs of bioisosteric groups that have quite similar Kohonen maps of their molecular electrostatic potential.

19.6 Bioisosteric Design

In drug design, a large amount of efforts has to be devoted to arrive from an initial lead structure at the final, highly active drug. Many structural variations have to be explored to increase biological activity. In this process, the concept of bioisosteric groups, of structure fragments having similar influence on biological activity has been found quite helpful.

Kohonen maps of the molecular electrostatic potential of several hundred pairs of bioisosteric groups have shown clear similarities in the maps of groups considered to be bioisosteric. Some pairs of bioisosteric groups are shown in Figure 19-19. In fact, this study allowed a more stringent definition of bioisosterism and led to the development of a new, highly active drug.

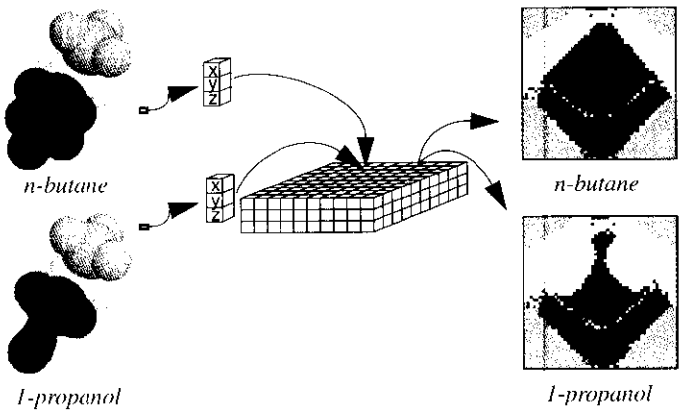

Figure 19-20: With the surface of *n-butane* a template Kohonen network is generated providing a reference map. Sending the coordinates from the surface of *propanol* through this template network provides a map with empty neurons at those places where the surface of *n-butane* differs from that of *propanol*.

19.7 Molecular Shape Analysis

The projection of molecular surface properties by a Kohonen network as outlined in Section 19.2 involves the training of the Kohonen network by the Cartesian coordinates of points from the molecular surface. Thus, the Kohonen network, in effect, stores the geometry of a molecular surface in its weights (there are three weights for each neuron!). If this is true, such a Kohonen network trained with the coordinates of a molecular surface can be used as a template for the comparison of other molecular surfaces. If the coordinates of a point from the surface of a second molecule (the compared molecule) correspond to the coordinates of a point of the first molecule, the neuron that stores the point from the first molecule (the template) will indicate a hit. Those parts of the surface of the template molecule that have no correspondence in the surface of the compared molecule will provide empty neurons.

Figure 19-20 indicates this process with the surface of *n-butane* as a template and the surface of *1-propanol* as the surface under study (the second molecule). The two areas of empty neurons in the compared Kohonen map correspond to the surface of the two hydrogen atoms that the methyl group of *n-butane* has on top of the atoms of an OH-group propanol. Note, that a proper comparison requires an alignment of the two molecules; in our case, three of the carbon atoms of *n-butane* were superimposed onto the three carbon atoms of *1-propanol*.

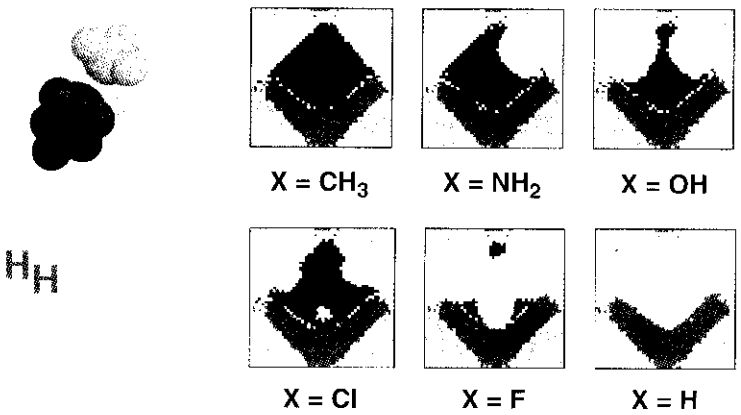

Figure 19-21: A series of 1-substituted *propane* derivatives and the corresponding compared maps obtained from the Kohonen network of *n-butane* as template.

Figure 19-21 shows the results of a more extensive comparison of the surfaces of 1-substituted propane derivatives sent through the template network of *n-butane*. The location of the empty neurons shows where the surface of the reference compound differs from that of the compared compound. An even clearer picture of the location of the differences in the shapes of the two molecular surfaces can be obtained by projecting the Kohonen map of the compared molecule back onto the 3D molecular surface of the template molecule.

Figure 19-21 indicates that the number of empty neurons might be taken as a quantitative measure of the difference in the surfaces of the reference and the compared compound. This has been verified in a number of studies. Thus, Kohonen networks can be used for the comparison of molecular shapes allowing the definition of a quantitative measure of shape difference.

19.8 References and Suggested Readings

19-1. J. Gasteiger and M. Marsili, "Iterative Partial Equalization of Orbital Electronegativity – A Rapid Access to Atomic Charges", *Tetrahedron* **36** (1980) 3219 – 3228.

19-2. J. Gasteiger and H. Saller, "Calculation of the Charge Distribution in Conjugated Systems by a Quantification of the Resonance Concept", *Angew. Chem.* **97** (1985) 699 – 701, *Angew. Chem. Int. Ed. Engl.* **24** (1985) 687 – 689.

19-3. T. Kohonen, "Self-Organized Formation of Topologically Correct Feature Maps", *Biol. Cybern.* **43** (1982) 59 – 69.

19-4. H. Ritter and T. Martinetz, K. Schulten, *Neuronale Netze, Eine Einführung in die Neuroinformatik selbstorganisierender Netzwerke*, Addison-Wesley, Bonn, FRG, 1990.

19-5. J. Gasteiger and X. Li, "Mapping the Electrostatic Potential of Muscarinic and Nicotinic Agonists with Artificial Neural Networks", *Angew. Chem.* **106** (1994) 671 – 674, *Angew. Chem. Int. Ed. Engl.* **33** (1994) 643 – 646.

19-6. X. Li, J. Gasteiger and J. Zupan, "On the Topology in Self-Organizing Feature Maps", *Biol. Cybern.* **70** (1993) 189 – 198.

19-7. J. Gasteiger, X. Li, C. Rudolph, J. Sadowski and J. Zupan, "Representation of Molecular Electrostatic Potentials by Topological Feature Maps", *J. Am. Chem. Soc.* **116** (1994) 4608 – 4620.

19-8. S. Anzali, G. Barnickel, M. Krug, J. Sadowski, M. Wagener, J. Gasteiger and J. Polanski, "The Comparison of Geometric and Electronic Properties of Molecular Surfaces by Neural Networks: Application to the Analysis of Corticosteroid Binding Globulin Activity of Steroids", *J. Comput.-Aided Mol. Design* **10** (1996) 521 – 534.

19-9. J. Polanski, J. Gasteiger, M. Wagener and J. Sadowski, "The Comparison of Molecular Surfaces by Neural Networks and Its Application to Quantitative Structure Activity Studies", *Quant. Struct.-Act. Relat.* **17** (1998) 27 – 36.

19-10. U. Holzgrabe, M. Wagener and J. Gasteiger, "Comparison of Structurally Different Allosteric Modulators of Muscarinic Receptors by Self-organizing Neural Networks", *J. Mol. Graphics* **14** (1996) 185 – 193.

19-11. J. Gasteiger, X. Li and A. Uschold, "The Beauty of Molecular Surfaces as Revealed by Self-organizing Neural Feature Maps", *J. Mol. Graphics* **12** (1994) 90 – 97.

19-12. S. Anzali, G. Barnickel, M. Krug, J. Sadowski, M. Wagener and J. Gasteiger, "Evaluation of Molecular Surface Properties Using a Kohonen Neural Network: Studies on Structure-Activity Relationships", in *Neural Networks in QSAR and Drug Design*, Ed.: J. Devillers, Academic Press, London, UK, 1996, pp. 209 – 222.