BUTTERWORTH HEINEMANN

Self-organizing maps and molecular similarity

Thomas W. Barlow

Physical Chemistry Laboratory, Oxford, U.K.

*Self-organizing maps generated by Kohonen neural networks provide a method for transforming multidimensional problems into lower dimensional problems. Here, a Kohonen network is used to generate two-dimensional representations of the electrostatic potential about the ring structures of histamine H2 agonists. Previous work by J. Gasteiger and X. Li (*Angew. Chem. Int. Ed. Engl. *1994, 33, 643) has shown the usefulness of such a method for classifying molecules as muscarinic or nicotinic agonists. Here, the method is extended to rank histamine H2 agonists in order of biological activity.*

Keywords: Kohonen neural network, self-organizing feature map, similarity

INTRODUCTION

Drug potencies are known to depend in some situations on molecular electrostatic potential. It is relatively easy to calculate the electrostatic potential surrounding a molecule. However, it can be difficult to compare the electrostatic potentials of different molecules. Gasteiger and Li[1] have demonstrated how Kohonen neural networks can be used to produce two-dimensional, topology-preserving maps of small molecules. They used these maps to compare molecules and decide whether they belonged to one of two classes. In this article we introduce a degree of quantification by dealing only with one class of molecules and ranking them according to their similarity.

THE KOHONEN NEURAL NETWORK

The neural network discussed in this article was originally developed by Kohonen[2–4] to model the laying down of sensory information in the brain. The network is described in detail elsewhere.[5–7] Its use for solving chemical problems has been pioneered by Gasteiger, Zupan, and others.[1,8–10] What follows is a brief summary of the algorithms used for this article.

The self-organizing map

The Kohonen network is used to map a three-dimensional molecular surface onto a two-dimensional map. The molecular surface is represented by s three-dimensional vectors, $\mathbf{x}_k = x_{k1} + x_{k2} + x_{k3}$ ($k = 1, 2, \ldots, s$). These are randomly chosen from a van der Waals surface. The two-dimensional map is represented by an array of units, where each unit has the same number of neighbors (eight in this case). This is achieved by wrapping each side of the array onto its opposite side, as shown in Figure 1. Associated with each unit of the array is a weight vector, $\mathbf{w}_j = w_{j1} + w_{j2} + w_{j3}$ ($j = 1, 2, \ldots, n$ labels the position of the unit in the array, which contains a total of n units). Initially the weight vectors are randomly chosen with smaller magnitude than the input vectors. They are then modified by an unsupervised learning procedure to approximate the set of input vectors. This procedure is described below in Eqs. (1)–(6).

The learning procedure is a competitive one. An input vector is selected at random and compared with each of the weight vectors by calculation of the Euclidean distance,

$$d = [(x_{k1} - w_{j1})^2 + (x_{k2} - w_{j2})^2 + (x_{k3} - w_{j3})^2]^{1/2} \quad (1)$$

The unit for which d is a minimum is selected as the "winning unit," c. Hence unit c is the unit for which the weight vector is closest to the input vector. That is,

$$|\mathbf{w}_c - \mathbf{x}_k| \leq |\mathbf{w}_j - \mathbf{x}_k| \qquad \text{for all } j \quad (2)$$

Once the "winning unit" has been selected, the weight vectors are modified so that they are closer to the input vector (see Figure 2). The components of the weight vectors are adjusted according to the following learning rule:

$$\Delta w_{ji} = \eta(k)a(k,j,c)(x_{ki} - w_{ji}) \quad (3)$$

where $i = 1, 2, 3$. $\eta(k)$ controls the learning rate. It is a linear, decreasing function of the input number:

$$\eta(k) = 0.3[1 - (k/s)] \quad (4)$$

Color plates for this article are on pp. 53–55.

Address reprint requests to Dr. Barlow, Physical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, U.K.
Received 10 June 1994; accepted 28 June 1994.

655 Avenue of the Americas, New York, NY 10010

0263-7855/95/$10.00
SSDI 0263-7855(94)00007-F

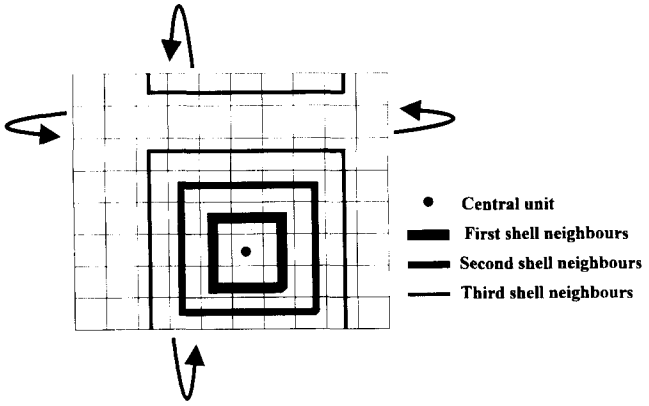

Figure 1. A two-dimensional array of units arranged as in the Kohonen network. Each unit has eight nearest neighbors. Units on the edge of the array are considered to neighbor units on the edge of the opposite side.

The neighborhood function $a(k,j,c)$ decreases with distance $|r_j - r_c|$ from the winning unit in the two-dimensional array. ($|r_j - r_c|$ is equivalent to the shell number in Figure 1.) This means that weight vectors are adjusted proportionally to their distance from the winning unit.

$$a(k, j, c) = \begin{cases} 1 & \text{for } |r_j - r_c| = 0,1 \\ 1 - (|r_j - r_c| - 1)/\sigma(k) & \text{otherwise} \end{cases} \quad (5)$$

$\sigma(k)$ defines the extent of the neighborhood function. Like $\eta(k)$, it is a linear, decreasing function of the input number. However, $\sigma(k)$ reaches its final value at $k = p$, $p < s$.

$$\sigma(k) = \begin{cases} \sigma(0) - [\sigma(0) - \sigma(p)]k/p & \text{for } k = 1, 2, \ldots p \\ \sigma(p) & \text{for } k = p, p + 1, \ldots s \end{cases} \quad (6)$$

Initially, when $\sigma(k)$ is large, the learning rule is extensive, leading to long-range ordering of the map. During this period, there are large fluctuations in the weight vectors. As $\sigma(k)$ and $\eta(k)$ decrease, with the basic ordering of the map already established, the weight vectors are fine tuned to best approximate the input vector space. The result is the formation of a map that reflects three-dimensional topological relationships in its two-dimensional array of weights.

The networks used in this paper are arrays of 24 × 24 units. $\sigma(0) = 12$ and $\sigma(p) = 1$. The total number of inputs $s = 10\,000$; $p = 8\,000$.

The feature map

The self-organizing map is an array of units, organized so that their position in the two-dimensional array reflects the topological position of their weight vectors in the input vector space. A feature, such as electrostatic potential on the surface of the molecule in three-dimensional space, can now be mapped onto the units of the array. This is summarized in Figure 3.

The electrostatic potential is calculated for each of the input vectors. The weights of the self-organized map are again compared with the input vectors as in Eqs. (1) and (2). This time, however, the weights are not adjusted. The electrostatic potential associated with the input vector is now allocated to the winning unit. This is repeated for several inputs, until a maximum number of units have been assigned a feature. If the same unit has a weight vector closest to more than one input vector, that unit adopts an electrostatic potential value equal to the average of all values assigned to it.

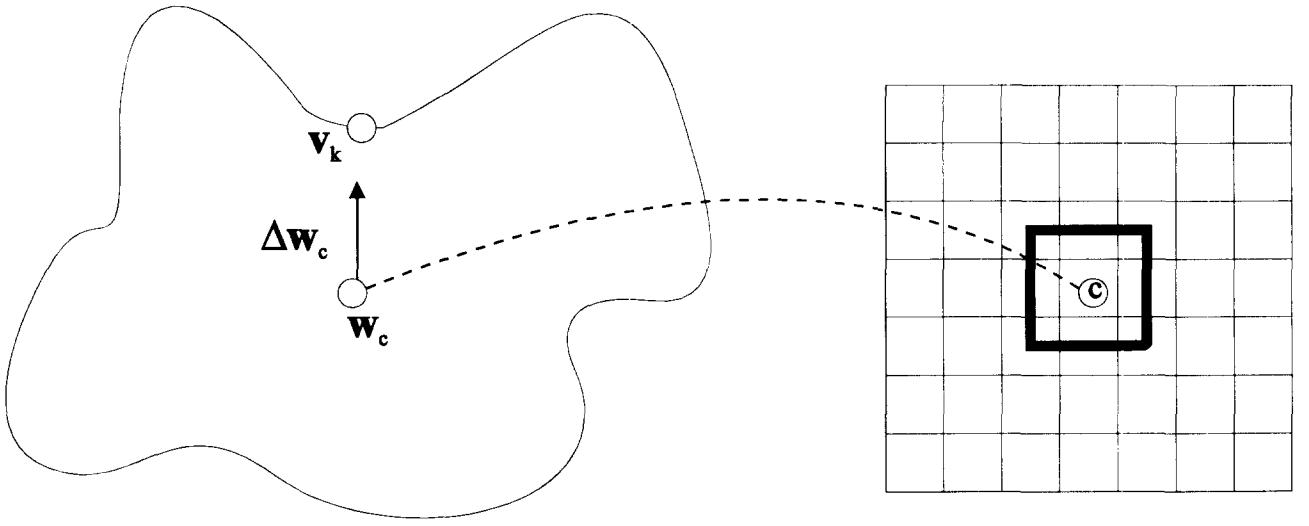

Figure 2. The winning unit, c, is the unit with weight vector, $\mathbf{w}_c$, closest to the input vector, $\mathbf{v}_k$. It is adjusted by an amount $\Delta\mathbf{w}_c$ so that it moves closer to the input vector in the higher dimensional vector space. The first shell of neighboring units will have their weights adjusted in a similar way. Other units will be modified proportionally to their distance from unit c in the two-dimensional array.

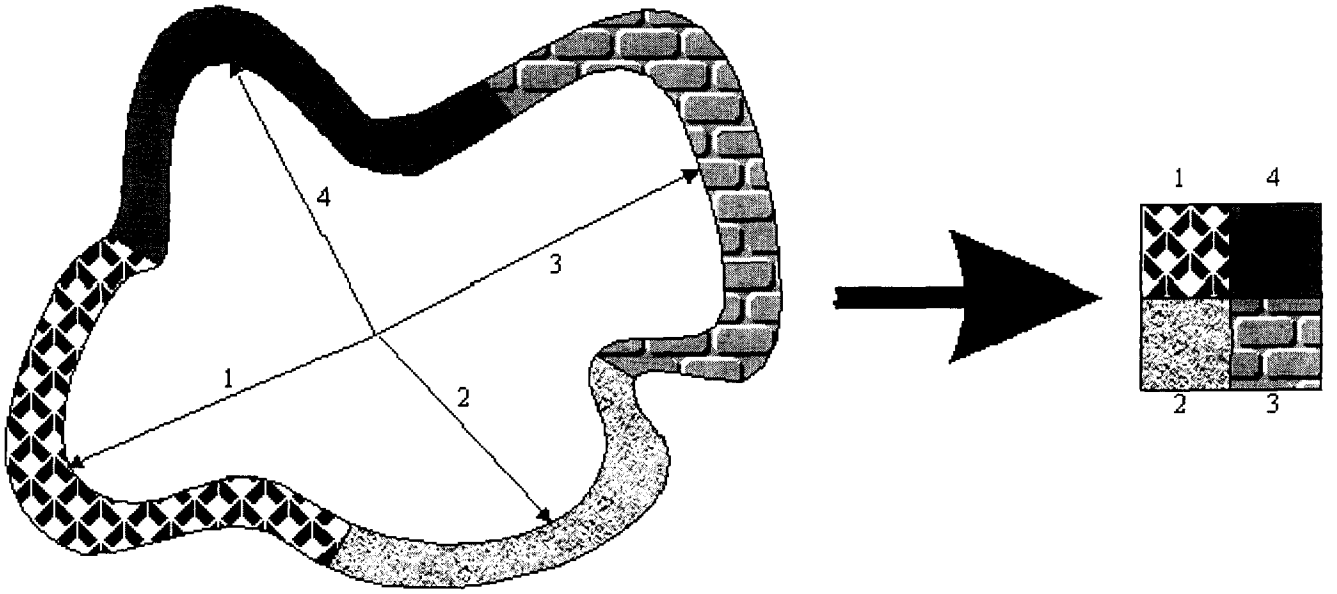

Figure 3. Constructing the feature map. Points on the surface of the molecule are compared with the four weight vectors. The unit associated with each weight vector takes on the feature of the molecular surface in the area surrounding its weight vector.

FEATURE MAPS AND SIMILARITY

Feature maps were made for the ring structures of a series of agonists for the histamine H2 receptor. These molecules are listed in Figure 4.

Basic molecular structures were built and the ethylamine side chain common to all of the molecules was replaced with a methyl group to avoid conformational differences. The structures of these derived compounds were then minimized using MOPAC.[11] Partial atomic charges were calculated empirically with RATTLER[12] and the electrostatic potentials at a series of points on the van der Waals surface were determined by NEMESIS.[13] These points were randomly shuffled and used to train a Kohonen neural network as described above.

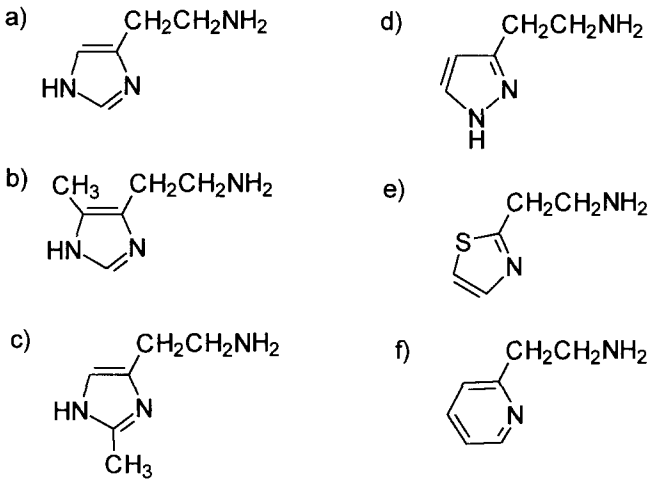

Figure 4. Series of histamine H2 receptor agonists, with their H2 activities relative to histamine (100): (a) histamine (100); (b) 4-methylhistamine (43); (c) 2-methylhistamine (4); (d) betazole (2); (e) 2-pyridylethylene (2); (f) 2-thiazoylethylamine (2). (Ganellin, C.R. In: Proceedings of 6th International Symposium on Medicinal Chemistry, September 1978, p. 167.)

Color Plate 1 shows the feature maps of the ring structures from Figure 4. Each map has been replicated four times, and the four copies laid edge to edge. This highlights the absence of boundaries in the two-dimensional array as discussed in Figure 1. The electrostatic potential has been color coded on the same scale for each map. The maps have been qualitatively ranked according to their similarity to the histamine ring map. This ranking corresponds to that of agonist activity at the histamine H2 receptor.[14]

FEATURE MAPS WITH A DIFFERENCE

An alternative method for comparing the molecules is to make a "comparison map." The histamine ring is used to build a self-organized map as described above. Then surface vectors from the other structures are used to generate feature maps. As a result, the electrostatic potential on the surface of each molecule is superimposed on the same map (the map of the histamine ring) (see Figure 5). Once again the maps can be ranked according to similarity, as in Color Plate 2.

The comparison maps in Color Plate 2 can readily be subtracted from one another to produce difference maps. The difference maps in Color Plate 3 confirm the ranking of maps in Color Plate 2.

A difference parameter can be calculated, on the basis of the sum of the magnitude of the difference between the electrostatic potential at corresponding units of the maps being compared. Difference parameters are listed in Color Plate 3. Ordering of these values also corresponds to the ranking of biological activity.

CONCLUSIONS

The Kohonen neural network provides a method for reducing the dimensionality of a problem. Here the network has

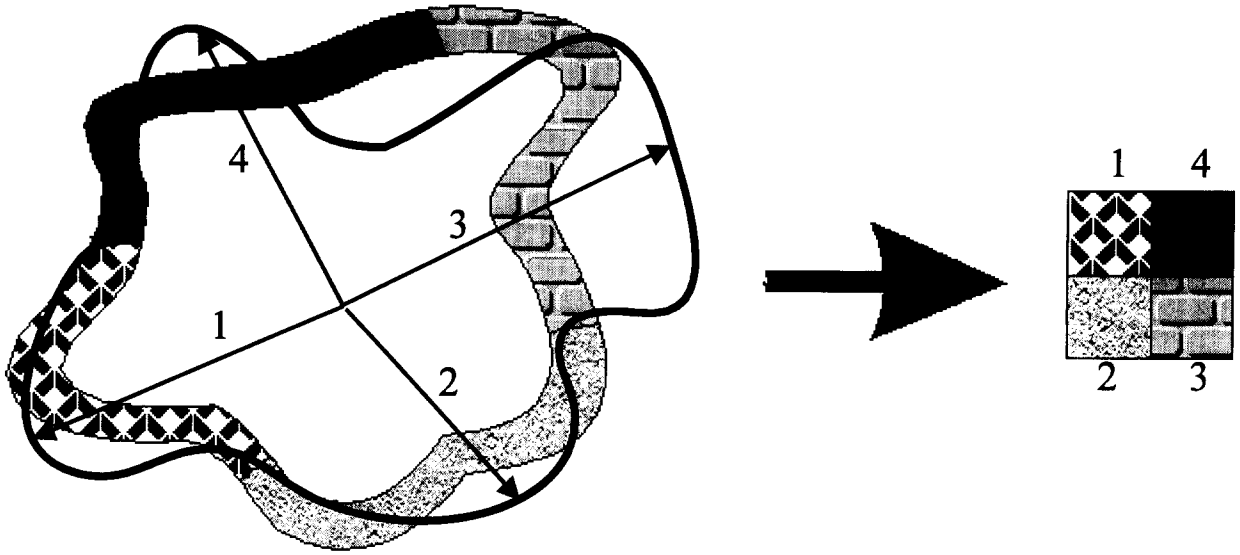

Figure 5. Feature mapping to produce a "comparison map." Features from one molecule are laid down on an array that has been organized for a different molecule.

been used to generate two-dimensional maps of the electrostatic potential surrounding a series of molecules. These maps can be compared and ranked qualitatively according to similarity. A new method has also been proposed for producing comparison maps of molecular electrostatic potential. These maps have also been ranked according to similarity. In both cases the ranking corresponds to that of molecular biological activity. The method may prove to be useful in similarity studies for drug design.

ACKNOWLEDGMENTS

The author is grateful for the support he receives from Kobe Steel, Ltd., and from the ORS award scheme through CVCP.

REFERENCES

1 Gasteiger, J. and Li, X. Mapping the electrostatic potential of muscarinic and nicotinic agonists with artificial neural networks. *Angew. Chem. Int. Ed. Engl.* 1994, **33(6),** 643–646

2 Kohonen, T. Self-organized formation of topologically correct feature maps. *Biol. Cybernetics* 1982, **43,** 59–62

3 Kohonen, T. Analysis of a simple self-organizing process. *Biol. Cybernetics* 1982, **44,** 135–140

4 Kohonen, T. *Self-Organization and Associative Memory*. Springer-Verlag, Berlin, 1988

5 Kohonen, T. The self-organizing map. *Proc. IEEE* 1990, **78(9),** 1464–1480

6 Kohonen, T. Things you haven't heard yet about the self-organizing map. *IEEE Conf. Neural Networks* 1993, **III,** 1147–1156

7 Hertz, J., Krogh, A., and Palmer, R.G. *Introduction to the Theory of Neural Computing*. Addison-Wesley Publishing Co., 1991

8 Gasteiger, J. and Zupan, J. Neural networks in chemistry. *Angew. Chem. Int. Engl.* 1993, **32,** 503–527

9 Simon, V., Gasteiger, J., and Zupan, J. A combined application of two different neural network types for the prediction of chemical reactivity. *J. Am. Chem. Soc.* 1993, **115,** 9148–9159

10 Zupan, J. and Gasteiger, J. *Neural Networks for Chemists: An Introduction*. WCH, Weinheim, 1993

11 Stewart, J.J.P. *MOPAC*, Q.C.P.E. 455

12 Oxford Molecular, Ltd., The Magdalen Centre, Oxford Science Park, Sanford on Thames, Oxford, OX1 4GA. *RATTLER*, 1991

13 Oxford Molecular, Ltd., The Magdalen Centre, Oxford Science Park, Sanford on Thames, Oxford, OX1 4GA. *NEMESIS*, 1994

14 Ganellin, C.R. Chemical characterization of agonists acting at histamine receptors. In: *Proceedings of 6th International Symposium on Medicinal Chemistry*, September 1978, pp. 167–177