ELSEVIER

Dyes and Pigments 49 (2001) 51–63

DYES and PIGMENTS

www.elsevier.com/locate/dyepig

The use of self-organising neural networks in dye design

A.J. Greaves [a,*], J. Gasteiger [b]

[a]*Department of Colour Chemistry, University of Leeds, Leeds LS2 9JT, UK*
[b]*Computer-Chemie-Centrum, University of Erlangen-Nürnberg, Germany*

Received 19 October 2000; accepted 15 December 2000

Abstract

The mapping of molecular surfaces is of particular interest to dye chemists for numerous reasons, none more so than the prediction of dye-substrate binding. Self-organising neural networks have been used to map the hydrogen bonding, electrostatic and hydrophobic 3D molecular surface potentials of a series of dyes. The results indicate that the hydrogen bonding potential, the molecular electrostatic potential and their combination are useful in classifying the dyes and that the hydrogen bonding potential is a useful molecular descriptor of substantivity. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Dye; Substantivity; Neural network; Dye design; Prediction; QSAR

1. Introduction.

Neural networks have seen an explosion of interest over the last few years, and are being successfully applied across an extraordinary range of problem domains, in areas as diverse as finance, medicine, engineering, physics and even colour chemistry [1]. Indeed, anywhere that there are problems of prediction, classification or control, neural networks are being introduced. This sweeping success can be attributed to their power and ability to model extremely complex functions.

One complex problem encountered by molecular scientists is the visualisation of a 3D molecular surface. It becomes even more complex when one tries to compare the molecular surface property of large numbers of molecules. Hypothetically, one could place all the molecules onto a 3D viewer, rotate them in 3D and then compare their surface properties. This technique is immensely laborious and prone to many problems. The projection of a 3D molecular surface onto a 2D plane quite obviously simplifies the process of comparison. A Kohonen neural network is capable of projecting a 3D molecular surface into a 2D-plane/topological map with maximum preservation of the topology and geometry of the multidimensional information.

Kohonen networks are used quite differently from other neural networks. The latter are designed for supervised learning tasks, where the training data set contains cases featuring input variables together with the associated outputs (and the network must infer a mapping from the inputs to the outputs). Kohonen networks are designed primarily for unsupervised learning, where the training data set contains only input variables. At first this may seem strange. However, Kohonen networks attempt to

* Tel.: +44-113-233-2811; fax: +44-113-233-2947.
E-mail address: a.greaves@leeds.ac.uk (A.J. Greaves).

PII: S0143-7208(00)00110-8

learn the structure of the data. One possible use is, therefore, in exploratory data analysis. The Kohonen network can learn to recognise clusters of data, and can also relate similar classes to each other. The user can build up an understanding of the data, which is used to refine the network. As classes of data are recognised, they can be labelled, so that the network becomes capable of classification tasks and eventually predictive tasks.

A Kohonen network has only two layers: the input layer, and an output layer of radial units (also known as the topological map layer). The units in the topological map layer are laid out in space — typically in two dimensions. Kohonen networks are trained using an iterative training procedure, which arranges the network so that units representing centres close together in the input space (i.e. the molecular space) are also situated close together on the topological map. You can think of the network's topological layer as a crude two-dimensional grid, which must be folded and distorted into the *N*-dimensional input space, so as to preserve as far as possible the original structure. Clearly any attempt to represent an *N*-dimensional space in two dimensions will result in loss of detail. Even so, the technique can be worthwhile in allowing the user to visualise data that might otherwise be impossible to understand. Once the network has been trained to recognise structure in the data, it can be used as a visualisation tool to examine the data (simply by applying a property to be visualised to each case/neuron).

The mapping of 3D molecular surfaces into 2D involves inputting the 3D atomic co-ordinates (Cartesian co-ordinates) of a 3D-geometry optimised molecule into a Kohonen network (Fig. 1). The number of inputs is determined by the number of variables of the molecule that is used as input to the network.

The size of the 2D plane is operator dependent, though generally the size is similar to the number of molecules. For example, a molecular surface which has been defined by using 10,000 Cartesian co-ordinates would have a starting architecture of $3\times100\times100$ where 3 is the number of Cartesian co-ordinates and $100\times100=10{,}000$. Like all neural networks some trial and error is needed to optimise the architecture.

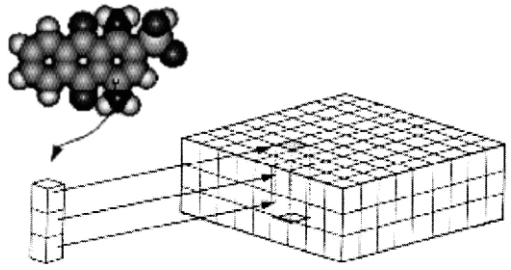


Fig. 1. Mapping of a 3D molecular surface using Cartesian co-ordinates as inputs into a Kohonen network.

The mapping of molecular surfaces is of particular interest to dye chemists for a number of reasons. Prediction of dye properties such as substantivity, exhaustion, fixation and even properties such as aggregation and solubility would be of immense benefit. Several linear modelling techniques, such as multiple linear regression analysis, have been used in attempts to predict the properties of dye molecules [2–6]. Unfortunately, none of these linear techniques have been successful enough to be of significant use. Understanding the properties of dyes and having the ability to predict the properties would have huge consequences on the synthesis and development of novel dye molecules for applications such as textile dyeing and printing and paper printing to applications as diverse as antibody-dye molecular recognition in techniques such as photodynamic therapy in biological systems. Many of the properties of the dye molecules will inevitably be due to the surface characteristics of the dye and it is for this reason that neural networks and molecular surface potentials are being applied to dye design. In this paper an account of the classification of acid, direct and reactive dyes is developed and the molecular electrostatic, hydrophobic and hydrogen bonding surface potentials related to the substantivity of the dyes for a biomass. The procedure for determining the substantivity has been reported elsewhere [7]. Greaves et al. have already demonstrated that the prediction of the substantivity of anionic, water soluble dyes using linear modelling techniques is immensely complex [8–10]. The complexity of this problem was deemed a suitable challenge for self-organising neural network approaches.

2. Mapping of molecular surface properties

If a 2D-topological/Kohonen map of the molecular surface is generated via the procedure outlined above, then individual surface characteristics such as molecular electrostatic, hydrophobicity and hydrogen bonding potentials can be plotted on the map and the surface characteristics visualised in 2D. Each of the surface characteristics is discussed below.

3. Molecular electrostatic potentials

The molecular electrostatic potential (MEP) on a molecular surface is particularly important because it is on the molecular surface that molecules come into intimate contact with chemical reagents or substrates. MEPs can be calculated from quantum mechanical, semi-empirical or ab initio methods. They can also be calculated from a simple, classical point charge model where each atom bears a partial atomic charge [11]. The partial atomic charges are obtained by partial equalisation of the orbital electronegativity. Experimentally, the MEP is obtained by moving a unit positive charge across the surface and, at various points, calculating the MEP by [Eq. (1), calculation of the molecular electrostatic potential]:

$$\mathrm{MEP}_j = \sum_{i}^{\text{atoms}} q_i / r_{ji} \qquad (1)$$

where

q = partial atomic charge and
r = distance between atoms i and j.

This latter method is the method used herein. The MEP measurement is taken at the same point as the Cartesian co-ordinate that is used as input into the neural network. Molecular surfaces that are projected onto Kohonen maps are generally visualised by colour. The MEP value is visualised on the 2D Kohonen map as a colour, different values of MEP being assigned different colours. For example, if the Cartesian co-ordinates of an oxygen atom are represented as red in the 3D molecular surface then the neurons that are occupied by the oxygen co-ordinates should be assigned as red, so that a comparison of 3D and 2D can be conducted (Fig. 2).

In other words, if the molecular surface property under investigation is the MEP then the value of the MEP at the Cartesian co-ordinate is input into the neuron in the network that specifies that Cartesian co-ordinate.

One extremely useful application of this technique is in the search for isosteric groups in research. It is possible to computationally search huge databases of molecules by this technique and identify molecules (by pattern recognition of the Kohonen map) that are isosteric (Fig. 3). Alternatively, one could search databases for dyes of similar molecular surface properties to improve (for example) substantivity.

4. Hydrogen bonding potential

In the same way that the MEP of a molecular surface can be projected onto a 2D plane, so can the hydrogen bonding potential (HBP). The calculation of HBPs is highly complex and detailed elsewhere [12]. What is important is that the HBP value is visualised on the 2D Kohonen map as a colour, different values of HBP being assigned different colours.

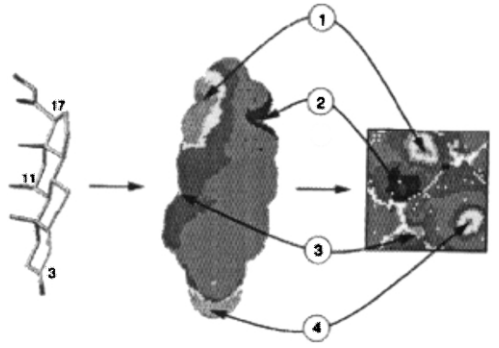


Fig. 2. Mapping of a 3D molecular structure onto a Kohonen map.

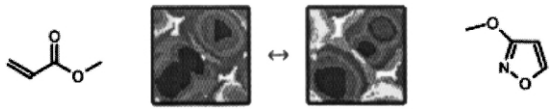


Fig. 3. Kohonen maps of the molecular electrostatic potentials of two isosteric molecules showing the similar surface MEPs.

5. Hydrophobicity potential

In the same way that the MEP and HBP of a molecular surface can be projected onto a 2D plane the hydrophobicity (HPP) can. Again, the calculation of HPPs is highly complex and follows the approach of Heiden [13]. The HBP value is visualised on the 2D Kohonen map as a colour, again different values of HBP being assigned different colours.

Combinations of the surface potentials can also be used to give a greater understanding of the dye property under investigation. For example, if analysis demonstrate that the MEP and HBP are important at describing the data then a combination of the two may improve the analysis.

6. Mapping of numerous 3D molecular surfaces using a Kohonen neural network

Previously, each neural network was trained with information from one molecular surface of one molecule. The maps generated were replications of one molecular surface. It is possible to map entire dataset of molecules into one Kohonen network. This enables the classification and hence understanding of large data sets. Every input mapped into the network will consist of an entire molecule (and not just the Cartesian co-ordinates of one point). In order to analyse data in this way the inputs have to be represented by the same number of variables. If the inputs are molecules, one has to come up with the same number of variables/descriptors irrespective of the size of the molecule and the number of atoms in the molecule. In other words, molecules of different sizes cannot be mapped onto the same network. Autocorrelation is used to attain this requirement.

7. Autocorrelation of molecular surface properties

Experimentally, a set of randomly distributed points on the molecular surface is generated. Then, all the distances between the surface points are calculated and sorted into pre-set intervals by [Eq. (2) calculation of auto-correlation reactors]:

$$A(d) = 1/L \sum_{ij} p(i)p(j) \quad d_l \leqslant d_{ij} \geqslant d_u \tag{2}$$

where $p(i)$ and $p(j)$ are property values at points i and j respectively (a property value could be partial charge); d_{ij} is the distance between the points i and j; L is the total number of distances in the interval $[d_l, d_u]$ represented by d. For a series of distance intervals with upper and lower bounds, d_u and d_l, a vector of autocorrelation (A) is obtained. It is a condensed representation of the distribution of the property p on the molecular surface.

The important point is that the autocorrelation vectors can be used as molecular descriptors in quantitative structure–property relationship (QSPR) studies.

8. Interpretation of Kohonen maps

Mapping of the autocorrelation vectors on a 2D plane should show, if they are suitable descriptors of the data, clustering of neurons. Simplistically, the clustering represents the degree of similarity between molecules. When used in this way a Kohonen network is simply a classification technique. The shapes of clusters can be quite elaborate and to someone inexperienced in the field, difficult to appreciate. It should be noted that clustering is the term given to clustering of neurons containing similar dyes and should not be confused with the clustering of molecules within one neuron.

The relationship between the molecules and the property under investigation is interpreted on a Kohonen 2D map in two ways. The first way is by the number of violations. Violations occur when the learning process cannot differentiate between the surface property of two or more dyes. They are represented on a coloured Kohonen map by criss-cross squares. Basically, when two dyes have the same (non-differentiable) surface characteristics then they are placed into the same neuron in the network. This can be detrimental to any analysis. For example, if one is trying to use the surface property MEP to describe the differences between two molecules and they are placed in the same neuron, then obviously the description is not possible. However, the lack of differentiation can be a positive thing because it enables the identification

of molecules that are structurally different but have the same surface property. The second way of interpreting the relationship between the molecules and the property under investigation on a Kohonen map is by the number of empty neurons, which are represented as blank or white squares on the coloured maps. Blank squares imply missing data. Ideally, all neurons should be filled because the architecture of the network is generally based around the principle "one input–one neuron". For example, if 100 inputs are used then a starting network would be 10×10 and each neuron should contain one dye. Empty neurons imply that there are neurons occupied by more than one dye, which in turn implies that the descriptor is not differentiating between the dyes. In practice, all neurons are rarely occupied. This is because data reduction from 20 dimensions (if 20 autocorrelation vectors are used) to 2 dimensions is immensely difficult.

9. General procedure

The sulphonic acid form of the dyes were 3D geometry optimised using CORINA [14]. The MEP and HBP of the van der Waals' surfaces were generated using PETRA [15] and the HPP using LOGP [16]. SURFACE [17] was used to calculate autocorrelation vectors and finally, KMAP [18] to project the autocorrelation vectors onto a 2D map. Through experience, a point density of 10 points/angstrom-square on the surface (10 measurements of the property taken for every angstrom-square on the surface) and a distance interval of 1 angstrom for each autocorrelation vector was proven to be optimal.

The mapping and ultimately the prediction of a property of a dye using 3D molecular surface properties proceeds in three steps:

9.1. Dye classification

Training of the network to produce a 2D map, which enables recognition of the structure in the data. This is basically a screening technique, where one is screening to see if the network can differentiate the dyes on the basis of their 3D molecular surface property. In other words, using the "one input–one neuron" principle, the differentiation is assessed by the number of violations and empty neurons as previously described. If the network can differentiate the dyes then it can be used to conduct a dye property classification.

9.2. Dye property classification

Application of a 3D molecular surface property to the 2D map and then visualisation to examine the data for clustering. The visualisation is usually done using colours, different colours representing different vales of the molecular surface potential. This is basically a screening technique, where one is looking to see if the network can cluster the dyes on the basis of their 3D molecular surface property and% substantivity. If the network can cluster the dyes then the network can be used to do a QSAR analysis. At this stage the dye property under investigation is generally divided into a small number of categories such as high, medium and low, the high, medium and low representing e.g. high, medium and low substantivity of a dye. The number of categories used depends on the amount of data available and the accuracy required from the model. The greater the number of categories the greater the predictive ability of the model.

9.3. QSAR analysis

Input of the auto-correlation vectors as independent variables into a QSAR analysis and development of a model for use in the prediction of the dye property under investigation.

This procedure is used below to develop a model suitable for the prediction of the substantivity of a series of anionic, water-soluble dyes onto a biomass.

10. Development of a model to predict the substantivity of a series of anionic, water-soluble dyes on a biomass

The structures of the dyes used in this study are given in the Appendix. Twenty autocorrelation vectors were chosen because the average length of all of the dye molecules was calculated to be approximately 20 angstroms. A 20×8×8 architecture was used for all Kohonen networks.

11. Classification of dyes using molecular electrostatic potential

The even spread of the dyes, the small number of empty neurons and the small number of violations in the numbered Kohonen map (Fig. 4) implied that the MEP of the surface of the dyes is sufficiently different and demonstrated that the MEP was a useful descriptor of the dyes in this study.

12. Classification of dyes using hydrogen bonding potential

As before, the even spread of the dyes, the small number of empty neurons and the small number of violations in the numbered Kohonen map (Fig. 5) implied that the HBP of the surface of the dyes is sufficiently different and demonstrated that the HBP was a useful descriptor of the dyes in this study.

13. Classification of dyes using hydrophobicity potential

The uneven spread of the dyes, the large number of empty neurons and the large number of violations in the numbered Kohonen map (Fig. 6) implied that the HPP of the surface of the dyes is sufficiently different and demonstrated that the HPP was a not a useful descriptor of the dyes in this study.

There were 15 dyes assigned to the bottom left corner neuron. The HPP descriptor was not differentiating between these. This bottom left corner neuron actually accounted for 9 of the 10 dyes based on the anthraquinone chromophore, which implied that the HPP was extremely poor at differentiating the hydrophobicity potential of dyes based on the anthraquinone chromophore.

The inability to differentiate the HPP of dyes meant that this descriptor would have been no use at

46	54	58	33		60		8
16		23	57		2	1 3 4	
11	43 44	14 48		51	50	49	63
41		20	31	55	34	26	47
40	42	22	30		45 59		56
52 52		5 6 7	18	21	35	15	17
62	36		13	10	9 12	24	61
32 37	50	39	29	38 39	25	28 27	19

Fig. 5. The numbered Kohonen map of the hydrogen bonding potential indicating which dyes are placed into which neurons.

2 49	13 48	22 50 51	55 59	31	37	21	3
46	38	30 29		34	20	62	1 4
10 45 47		57 54	39	26		61 63	
		24	56	52	58		18
25		7	28		41		32
27 17		5 6		42 44	43		35
15	8	36		60			53
16	11 19		14	9 12	33		23 40

Fig. 4. The numbered Kohonen map of the molecular electrostatic potential indicating which dyes are placed into which neurons.

	22		39		12 13		9 10
54 35						21	11
		50	25 27		18	15 19	
32			20		14 23		24
	50 51					16 46	17
55	59		44	33 42 53	29 40		
	31		45 57	56	41	47	28
1 2 3 4 30 8 5 7 26 34 36 37 60 62 63 13		38 39		52	43 48		49 61

Fig. 6. The numbered Kohonen map of the hydrophobicity potential indicating which dyes are placed into which neurons.

describing an experimentally determined property, such as substantivity. Combinations of the MEP and HPP, HBP and HPP and MEP, HBP and HPP were also poor at differentiating the dyes. The results of these combinations implied that the HPP was the dominant feature in the combinations because no separation of the 15 dyes assigned to the bottom left corner neuron occurred.

Since HPP and its combinations with MEP and HBP were poor at differentiating the dyes, the descriptor was not used in subsequent experiments aimed to describe substantivity.

14. Classification of dyes using a combination of molecular electrostatic potential and hydrogen bonding potential

Not surprisingly, the even spread of the dyes, the small number of empty neurons and the small number of violations in the numbered Kohonen map (Fig. 7) demonstrated that a combination of the MEP and HBP was a useful descriptor of the dyes in this study.

15. Classification of substantivity

The MEP, HBP and a combination of the MEP and HBP were shown to differentiate the dyes well and therefore can be used as descriptors for the classification of the experimental data. The HPP was poor at differentiating the dyes and, therefore, could not be used as a descriptor.

The 63 dyes were classified into 4 distinct substantivity groups: high, medium, low and zero. The specification of each individual group is shown in Table 1.

The substantivity results are shown in the Appendix A. A 20×8×8 architecture was used throughout so that a comparison of the descriptors could be made. Each different shade on the following shaded Kohonen maps represents a different group.

Fig. 8 shows:

1. the MEP was a poor descriptor of the substantivity because no clustering of the dye groups occurred;
2. the HBP is a moderate descriptor of the substantivity because there was some evidence of clustering;
3. a combination of both the MEP and the HBP were a poor descriptor of the substantivity because there was no evidence of clustering.

There appeared to be some clustering in the HBP map, which implied that the auto-correlation vectors of the HBP would be useful QSAR descriptors

10 43 45	44	13 58		22	5 29	14	11
2	46	36				28	7
47 49	62	55	63	19	15 35	33	24
18 30 23	52 56	53		25	8		39
17	59		16 50	37			31
34	32	61		39	41 57	12	9
21		20	40				27
1 3 4	60	42	50	26	54	51	38

Fig. 7. The numbered Kohonen map of the molecular electrostatic potential and hydrogen bonding potential indicating which dyes are placed into which neurons.

Table 1
Specification for the individual groups

Dye group	% Substantivity	Representation
High substantivity	66.1–100	
Medium substantivity	33.1–66.0	
Low substantivity	0.1–33.0	
Zero substantivity	0	

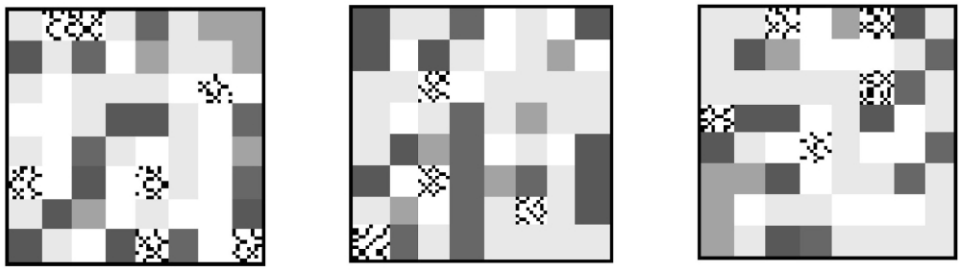


Fig. 8. Kohonen maps of the molecular electrostatic potential.

of the substantivity of this series of water-soluble dyes. The MEP and a combination of the MEP and HBP showed no clustering and thus were not used in any subsequent QSAR study.

16. Prediction of substantivity using 3D molecular surface properties and a back propagation algorithm

Classification of the results by the Kohonen networks demonstrated that the hydrogen bonding potential was the only molecular surface potential descriptor that showed promise as a descriptor of substantivity. Consequently, this was the only molecular surface property used in the development of a model to predict substantivity.

The auto-correlation vectors of the hydrogen bonding potential 3D molecular surface property were used as input into a multilayer perceptron network using a feed-forward back propagation algorithm. This type of network is the most popular network architecture in use today and account for 80% of all practical applications of NNs [19]. Since the output data (the substantivity results) was known, a supervised learning network was used involving a back propagation algorithm.

The architecture was $20\times10\times1$. The total number of dyes was divided into a training set (T) and a verification set (V). The verification set was approximately 10% of the total number of dyes and the computer software randomly assigned the dyes contained in this set. A number of experiments were conducted with each configuration and the best network (in terms of error) was retained. The network was deemed to be performing satisfactorily if the training and verification errors (RMS errors) were between 0.05 and 0.20. The predicted substantivity results were plotted against the observed substantivity results and the results analysed statistically.

Table 2
The result of the back propagation neural network using the auto-correlation vectors of the hydrogen bonding potential as input and the % substantivity as output

No. of dyes		RMS error				
T	V	T	V	r	r^2	P
59	6	0.17	0.07	0.83	0.69	0.00

The software packages STATISTICA [20] and STATISTICA NEURAL NETWORKS [21] were used for all analyses in this section.

The result of the optimal back propagation neural network is shown in Table 2.

The correlation coefficient (r) result of the experiment was deemed highly significant because the P level was less than 0.01 (1%). The correlation coefficient (r) was of a magnitude that demonstrated that the auto-correlation vectors of the hydrogen bonding potential of the 3D molecular surface were important in describing the substantivity of the dyes. This is further evidence to substantiate that the hydrogen bonding ability of water-soluble dyes is important in the substantivity of the dyes.

17. Conclusions

The auto-correlation vectors of the hydrogen bonding potential, the molecular electrostatic potential and their combination were found to be useful descriptors of anionic, water-soluble dyes. The auto-correlation vectors of the hydrophobicity potential were not useful descriptors because they could not differentiate between the dyes.

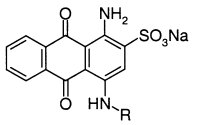
The hydrogen bonding potential appeared to be a useful descriptor of the substantivity of the dyes studied. The correlation coefficient was of a magnitude that demonstrated that the auto-correlation vectors of the hydrogen bonding potential of the 3D molecular surfaces of anionic, water soluble dyes can be used to predict the substantivity of the dyes for a biomass.

Acknowledgements

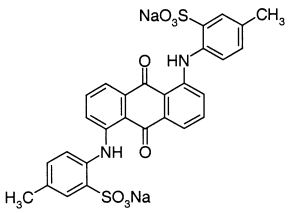
Dr. M.G. Hutchings at BASF (UK) is thanked for his valuable time and comments.

Appendix

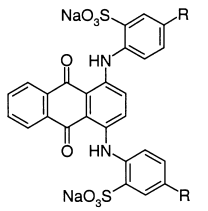
The % substantivity is shown in parentheses.



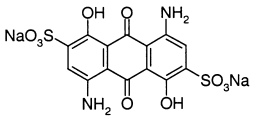
Dye 1 (42.0%): R = C_6H_5
Dye 2 (46.0%): R = 4-Me, 3-CH_3CONH-C_6H_3
Dye 3 (33.0%): R = cyclohexyl
Dye 4 (45.0%): R = 4-Me-C_6H_4



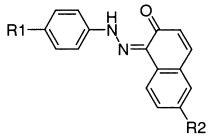
Dye 5 (11.0%)



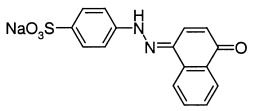
Dye 6 (0.0%): R = CH_3
Dye 7 (84.0%):R = C_4H_9



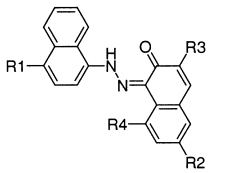
Dye 8 (0.0%)



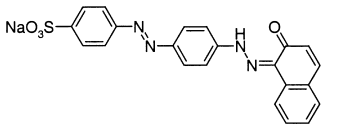
Dye 9 (22.0%): R1 = SO_3Na, R2 = H
Dye 10 (29.0%): R1 = H, R2= SO_3Na



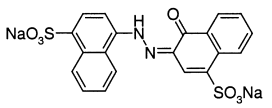
Dye 11 (11.0%)



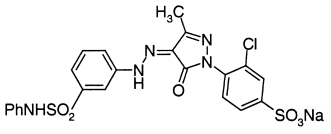
Dye 12 (79.0%): R1=SO_3Na, R2-4= H
Dye 13 (90.0%): R2= SO_3Na, R1,3,4= H
Dye 14 (0.0%): R1,2= SO_3Na, R3,4= H
Dye 15 (11.0%): R2,3= SO_3Na, R1,4= H
Dye 16 (0.0%): R1-3= SO_3Na, R4= H
Dye 17 (0.0%): R2-4= SO_3Na, R1= H



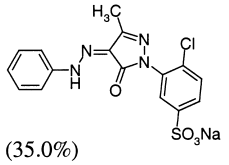
Dye 18 (96.0%)



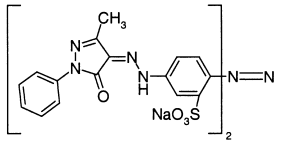
Dye 19 (4.0%)



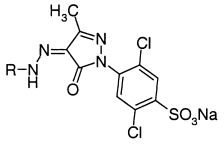
Dye 20 (22.0%)



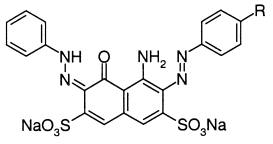
Dye 21 (35.0%)



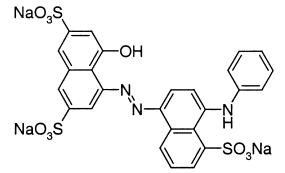
Dye 22 (40.0%)



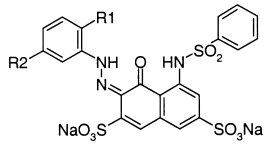
Dye 23 (0.0%): R= 4-SO_3Na-C_6H_4
Dye 24 (6.0%): R= 2-(1-SO_3Na)-naphthyl



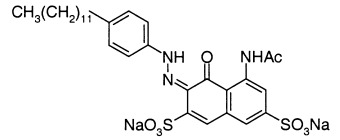
Dye 25 (8.0%): R= NH_2



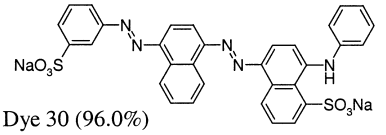
Dye 26 (9.0%)



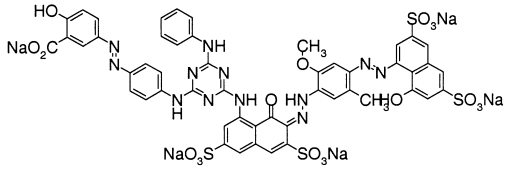
Dye 27 (22.0%): R1,2= H
Dye 28 (26.0%): R1= PhO, R2=Cl.



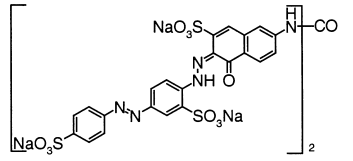
Dye 29 (96.0%)



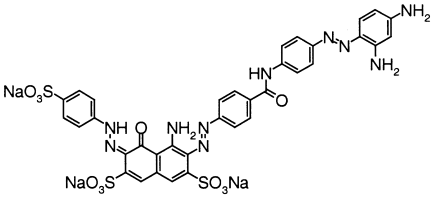
Dye 30 (96.0%)



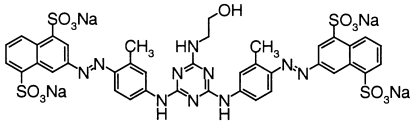
Dye 31 (22.0%)



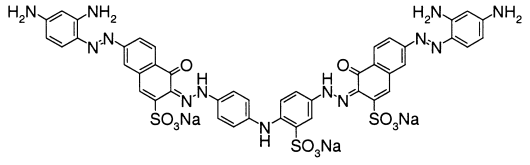
Dye 32 (12.0%)



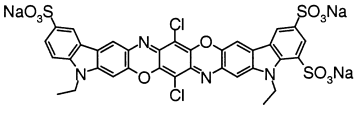
Dye 33 (53.0%)



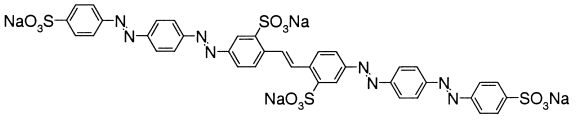
Dye 34 (8.0%)



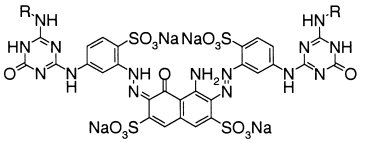
Dye 35 (89.0%)



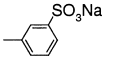
Dye 36 (25.0%)

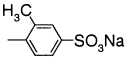


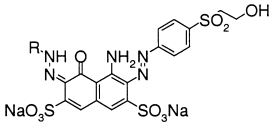
Dye 37 (7.0%)



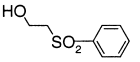
R

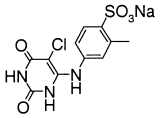
Dye 38 (10.5%) 

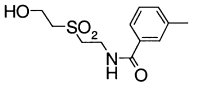
Dye 39 (12.9%) 

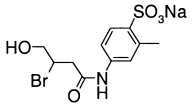


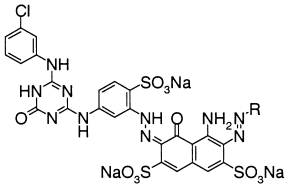
R

Dye 40 (0.4%) 

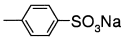
Dye 41 (6.2%) 

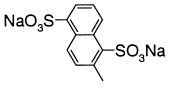
Dye 42 (0.0%) 

Dye 43 (4.8%) 



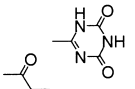
R

Dye 44 (10.0%) 

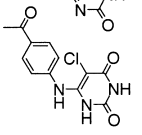
Dye 45 (7.1%) 

R

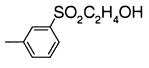
Dye 46 (3.2%)



Dye 47 (5.0%)

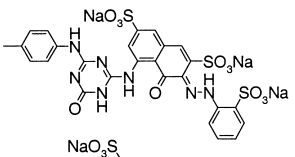


Dye 48 (2.1%)

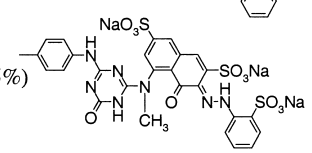


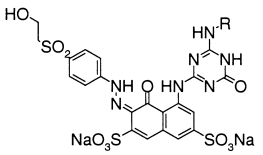
Dye 49 (7.3%)

Dye 50 (16.4%)



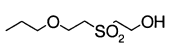
Dye 51 (14.5%)





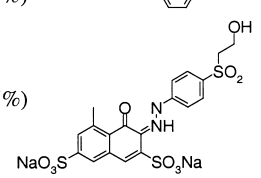
R

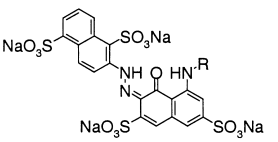
Dye 52 (0.0%)



Dye 53 (0.0%)

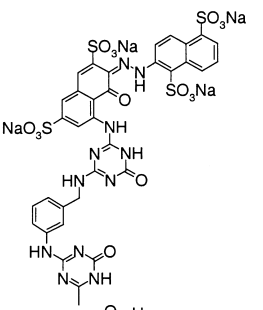
Dye 54 (5.7%)



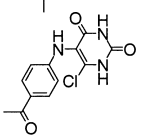


R

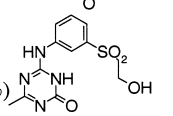
Dye 55 (11.8%)



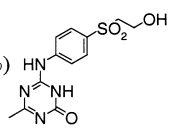
Dye 56 (0.0%)



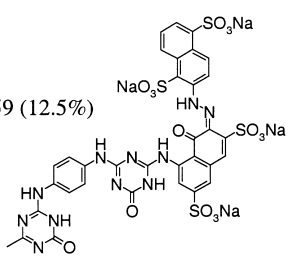
Dye 57 (1.4%)

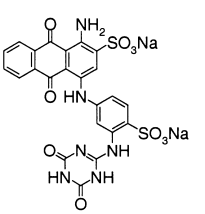


Dye 58 (1.3%)

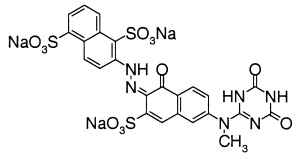


Dye 59 (12.5%)

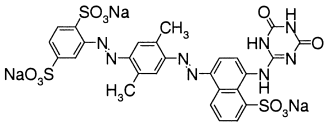




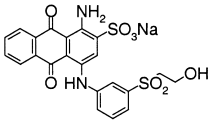
Dye 60 (19.6%)



Dye 61 (0.0%)



Dye 62 (3.0%)



Dye 63 (25.3%)

References

[1] Gasteiger J, Zupan J. Neural networks in chemistry and drug design. Wiley-VCH; 1999.

[2] Timofei S, Kurunczi L, Schmidt W, Simon Z. Dyes and Pigments 1996;32:25.

[3] Timofei S, Fabian WMF. Journal of Molecular Structure (Theochem) 1996;362:155.

[4] Timofei S, Kurunczi L, Schmidt W, Simon Z. Dyes and Pigments 1995;29:251.

[5] Timofei S, Kurunczi L, Fabian WMF, Suzuki T, Muresan S. Dyes and Pigments 1997;34:181.

[6] Timofei S, Kurunczi L, Schmidt W. Quant Struct–Act Relat 1995;14:444.

[7] Greaves AJ, Churchley JH, James AE, Hutchings MG, Phillips DAS. Water Research 2000;34(5):1673.

[8] Greaves AJ, Churchley JH, Hutchings MG, Phillips DAS, Taylor J. J Soc Dyers and Colourists 1999;115:363–5.

[9] Greaves AJ, Churchley JH, Hutchings MG, Phillips DAS, Taylor J. J Soc Dyers and Colourists 2000;116:222–8.

[10] Greaves AJ, Churchley JH, Hutchings MG, Phillips DAS, Taylor J. J Soc Dyers and Colourists 2000;116:279.

[11] Gasteiger J, Li X. Angew Chem, Int Ed Engl 1994;33:643.

[12] Vedami S, Huhta DW. JACS 1990;112:4759.

[13] Heiden W, Moeckel G, Brickmann J. J Comput-Aided Mol Des 1993;7:503.

[14] Gasteiger J, Corina software. Computer-Chemie-Centrum, Institute for Organic Chemistry, University of Erlangen-Nurnberg, Nagelsbachstrasse 25, D-91052 Erlangen, Germany (www2.ccc.uni-erlangen.de).

[15] Gasteiger J, PETRA software (1995). Computer-Chemie-Centrum, Institute for Organic Chemistry, University of Erlangen-Nurnberg, Nagelsbachstrasse 25, D-91052 Erlangen, Germany (www2.ccc.uni-erlangen.de).

[16] Gasteiger J, Logp software: Computer-Chemie-Centrum, Institute for Organic Chemistry, University of Erlangen-Nurnberg, Nagelsbachstrasse 25, D-91052 Erlangen, Germany (www2.ccc.uni-erlangen.de).

[17] Gasteiger J, Surface software version 1.0 (1994). Computer-Chemie-Centrum, Institute for Organic Chemistry, University of Erlangen-Nurnberg, Nagelsbachstrasse 25, D-91052 Erlangen, Germany. (www2.ccc.uni-erlangen.de).

[18] Gasteiger J, KMAP version 2.1 (1996). Computer-Chemie-Centrum, Institute for Organic Chemistry, University of Erlangen-Nurnberg, Nagelsbachstrasse 25, D-91052 Erlangen, Germany. (www2.ccc.uni-erlangen.de).

[19] Westland S. J Soc Dyers and Colourists 1998;114:309.

[20] STATISTICA version 5.1: Statsoft Inc, USA.

[21] STATISTICA NEURAL NETWORKS version 1: Statsoft Inc, USA.