Neural Network Method, the Tool for Studying Biological Activity of Compounds
Relationship between Infiltration Anaesthesia, Coded Structural Information, and Chromatographic Properties Applied in Homologous Series of Alkoxy-Substituted Esters of Phenylcarbamic Acids

[a]Š. HATRÍK, [b]J. LEHOTAY, and [c]J. ČIŽMÁRIK

[a]*Institute of Chemistry, Faculty of Natural Sciences, Comenius University, SK-842 15 Bratislava*

[b]*Department of Analytical Chemistry, Faculty of Chemical Technology, Slovak Technical University, SK-812 37 Bratislava*

[c]*Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Comenius University, SK-832 32 Bratislava*

Received 1 March 1995

The mathematical method of neural network was employed for studying of infiltration anaesthetic activity of five homologous series of *o*- and *m*-alkoxy-substituted morpholino-, piperidino-, perhydroazepino-, and dimethylaminoethyl esters and piperidinopropyl esters of phenylcarbamic acids, respectively. RP-HPLC capacity factors were used for the characterization of the lipophilicity of tested drugs. The three-layer perceptron, that is trained by the back propagation of errors, was successfully used for supplementing of the incomplete original data matrix and also for smoothing of the biological data. The relationships between the infiltration anaesthesia and the number of C atoms in the alkoxy side chain (LC capacity factors, respectively) for the homologous series presented the peak character, which is in agreement with the theoretical assumptions.

The locations of maxima were dependent on the position of the alkoxy side chain in the molecules of tested drugs. The maxima of infiltration anaesthesia for the homologous series of esters of phenylcarbamic acids occurred at seven and eight C atoms in the alkoxy side chain. *m*-Substituted drugs presented maxima in the range from three to five C atoms (except of piperidinopropyl esters which had a flat maximum at eight C atoms and dimethylaminoethyl esters which had two maxima at four and seven C atoms). Generally, the infiltration anaesthetic activity of the *o*-derivatives was higher than that of *m*-substituted drugs.

Generally, after discovery of a biologically active compound with the new structure, the next phase of activity optimization is commonly characterized by the varying of the basic structure to achieve the maximal biological activity [1]. The finding of mathematical relationships between the biological activity, which is a very complex quality, and a measurable characteristics (physicochemical parameters, structural information, *etc.*) of compounds can be very problematic due to the nonlinearity of dependences investigated.

Standard modelling techniques require the mathematical function to be known in advance. The advantage of the neural network model is that it does not require a knowledge of the mathematical function. The nonlinearity of a single unit transformation and a sufficiently large number of variable parameters ensure adaptation of the neural network to any relation between input and output data [2].

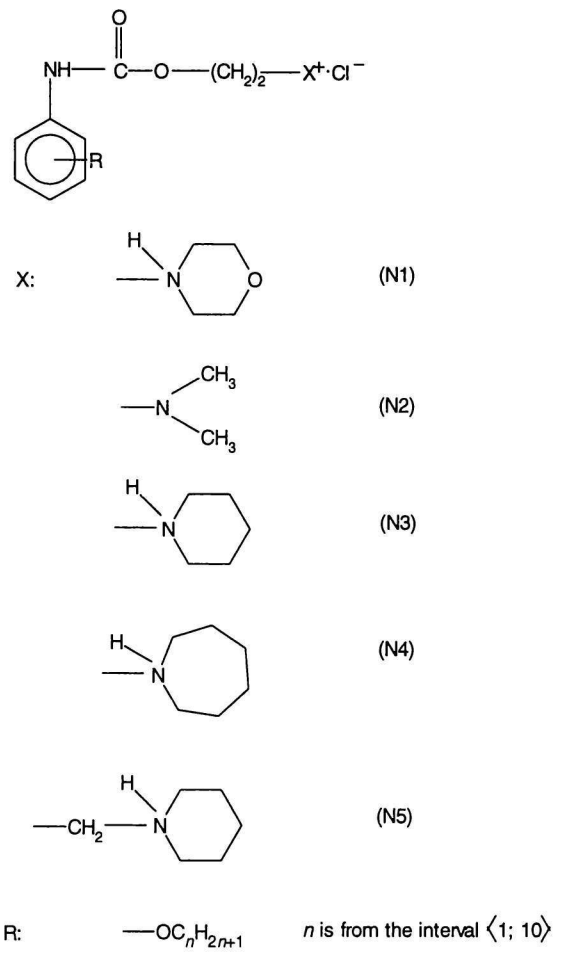
Neural networks are specialized computer systems. They process entire patterns rather than individual items of data. Neural networks are particularly appropriate for a special class of problems, which may be characterized as follows [3]:

Pattern recognition, rather than sequential data processing;

nonlinear capabilities, which can be a problem for other methods;

sufficient examples of the target data are available to allow a learning system to be considered.

Once trained, neural networks are quick and have



Formula 1. Structures of the tested drugs

a tolerance to incomplete or noisy data. This makes them helpful in a real-world environment, such as biological systems considered in this work.

In the presented project the multilayer perceptron, that is trained by the back propagation of errors, was employed. Multilayer perceptron consists of a number of interconnected processing elements organized into layers. The outputs of these elements are connected to the inputs of other elements. Relationships between elements are controlled by the weights. A training algorithm measures the error of output [4].

The homologous series of *o*- and *m*-alkoxy-substituted esters of phenylcarbamic acids, tested in this work, represent a very important group of compounds because of their high local anaesthetic activity [5—9]. The study of correlation between local anaesthetic activity, LC capacity factors, and some physicochemical properties of these compounds was published previously [10—13]. The aim of this work was to use an advanced mathematical neural network method for studying of the relationships between infiltration anaesthetic activity of the above derivatives of phenylcarbamic acid, some coded structural information, and lipophilicity represented by the RP-HPLC chromatographic capacity factors.

EXPERIMENTAL

The drugs studied (for the chemical structure see Formula 1) were prepared according to [5—9] and the values of surface anaesthesia were taken also from these papers.

The experimental details of the HPLC system used for the determination of capacity factors of tested local anaesthetics were published previously [14].

Neural network employed in presented experiments had the following architecture. Four input "neurons" which partly characterized the molecules of local anaesthetics (number of carbon atoms in the aliphatic side chain, position of this chain on the benzene ring, type of a nitrogen-containing substituent and LC chromatographic capacity factor) were used. To provide good generalization from training data to testing data, it is desirable to use the smallest number of hidden "neurons" that give satisfactory training performance. Hidden layer consists of five elements and one output element (logarithm of activity in an infiltration anaesthesia). The relationship between error (SSO) and number of hidden neurons is shown in Fig. 1. All neurons had a feed-forward layered structure with connections allowed only between adjacent layers. Before calculation data were transformed into the interval from 0 to 1. The number of hidden neurons was optimized. Schematic representation of the above neural network is shown in Fig. 2. Program of neural network was written in Turbo-Pascal 7.0. Training process took 150 min on a PC-AT 486 DX, 66 MHz.

RESULTS AND DISCUSSION

The mechanism of biological activity of a compound is commonly determined by its transport to the receptor and by the interaction with this receptor. The transport strongly depends on the lipophilicity of compound [1]. Reversed phase chromatography (especially with C 18 stationary phase) can be successfully employed for the characterization of lipophilicity of tested derivatives of phenylcarbamic acid [10—13].

The structural information of morpholino- [6], piperidino- [7], perhydroazepino- [5], and dimethylaminoethyl esters [8] and piperidinopropyl esters [9] of *o*- and *m*-alkoxy-substituted phenylcarbamic acids was coded by the following way:

Position of the alkoxy chain on the benzene ring (*o*: 1, *m*: 2);

number of C atoms in this chain (from one up to ten);

type of nitrogen-containing substituent (Formula 1) (morpholinoethyl: 1, dimethylaminoethyl: 2, piperidinoethyl: 3, perhydroazepinoethyl: 4, piperidino-

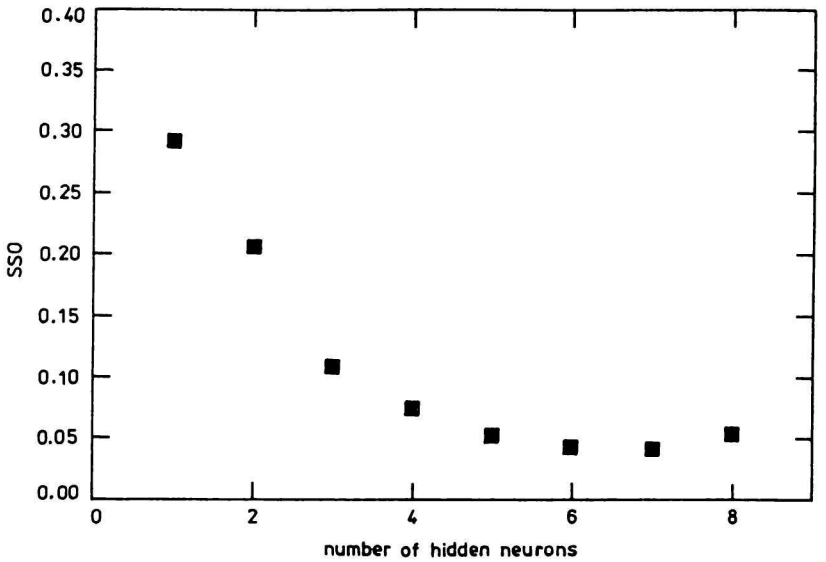


Fig. 1. Dependence of the sum-of-squares error of testing set on the number of hidden neurons (50000 iterations for each calculation).

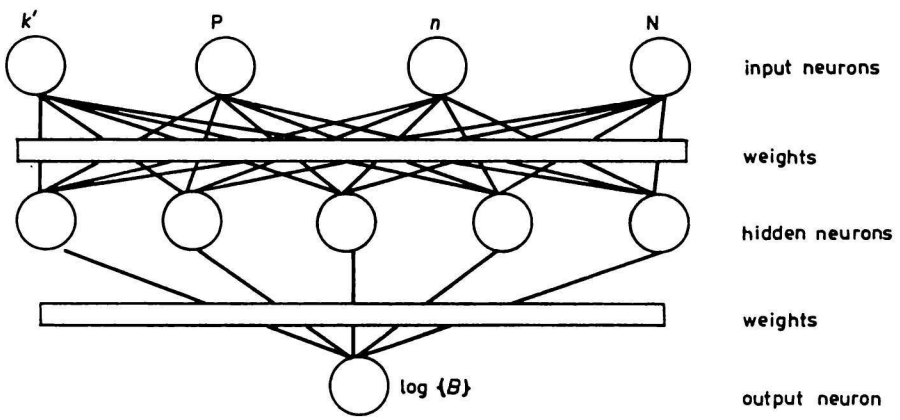


Fig. 2. Schematic presentation of neural network employed in this work. N represents the type of nitrogen-containing substituent, n is the number of C atoms in an alkoxy side chain, P is the position of this chain on the benzene ring, k' is the LC capacity factor, and $\log\{B\}$ is the infiltration anaesthesia.

propyl: 5).

The experimental data were divided into a training set of 52 local anaesthetics which was used in the learning process, and the testing set of 15 local anaesthetics. To warrant the objectivity of testing process, 15 local anaesthetics were taken uniformly from each homologous series. The learning of the neural network was accomplished by repeated cycling through the training data, presenting patterns at the input elements and indicating associations of all elements. In the testing process the output (infiltration local anaesthesia) is calculated according to the variable parameters (weights) obtained in the training process.

The outcomes of the learning and testing processes were evaluated by the following characteristics:

Average sum-of-squares error (SSO) of calculated and experimental outputs,

the gradient of function (grad),

index of correlation (I_c^2) of calculated and experimental outputs.

The above characteristics of our system are listed in Table 1. Indices of correlation both for training and testing process were around 0.9 (one is the best fit).

Table 1. The Results of Training and Testing Processes

	Training process	Testing process
SSO	0.022	0.024
grad	0.00015	–
I_c^2	0.899	0.900

Table 2. The Original Data and Data Calculated by the Neural Network for Morpholinoethyl (N1), Dimethylaminoethyl (N2), Piperidinoethyl (N3), Perhydroazepinoethyl (N4), and Piperidinopropyl Esters (N5)

		N1			N2			N3			N4			N5		
P	n	k'	log {B}		k'	log {B}		k'	log {B}		k'	log {B}		k'	log {B}	
			exp.	calc.		exp.	calc.		exp.	calc.		exp.	calc.		exp.	calc.
o-	1	0.49	N.M.	0.14	1.23	0.64	0.71	1.25	1.25	1.26	1.72	N.M.	1.08	1.74	N.M.	0.50
	2	0.55	N.M.	0.22	1.24	1.11	0.93	1.32	1.32	1.40	1.81	N.M.	1.10	1.98	0.56	0.64
	3	0.64	N.M.	0.32	1.33	1.19	1.25	1.50	1.52	1.39	2.06	1.27	1.22	2.11	0.78	0.71
	4	0.77	0.30	0.47	1.49	1.88	1.62	1.80	1.34	1.53	2.47	1.18	1.37	2.57	1.16	1.20
	5	0.90	N.M.	0.55	1.78	1.52	1.96	2.15	2.00	1.72	2.92	1.88	1.54	3.10	1.83	1.78
	6	1.15	0.90	0.81	2.08	2.37	2.24	2.76	2.02	1.91	3.58	1.42	1.77	3.71	2.24	2.21
	7	1.43	1.00	0.95	2.55	2.43	2.42	3.30	2.23	2.07	4.34	2.28	1.95	4.70	2.41	2.22
	8	1.80	0.78	1.01	3.29	2.14	2.33	4.08	2.17	1.90	5.20	N.M.	1.70	6.10	1.27	1.53
	9	2.27	1.20	0.90	4.00	1.87	1.85	5.21	1.43	1.25	6.28	N.M.	1.17	7.50	1.22	1.04
	10	2.90	0.30	0.50	5.14	N.M.	0.99	6.82	0.85	0.95	7.41	N.M.	0.99	9.75	N.M.	0.96
m-	1	0.47	N.M.	0.26	1.30	N.M.	1.10	1.05	1.20	1.21	1.66	N.M.	1.01	1.35	N.M.	0.60
	2	0.51	N.M.	0.39	1.34	1.54	1.54	1.18	1.46	1.35	1.79	N.M.	1.35	1.59	N.M.	0.59
	3	0.59	N.M.	0.44	1.56	1.63	1.76	1.41	1.55	1.70	1.92	1.97	1.92	1.74	0.52	0.65
	4	0.67	0.60	0.45	1.90	1.70	1.82	1.74	N.M.	1.98	2.35	1.60	1.74	2.31	1.73	1.53
	5	0.82	0.60	0.52	2.38	1.81	1.72	2.03	2.13	2.04	2.94	1.78	1.39	2.86	0.94	1.28
	6	1.09	0.85	0.82	2.86	1.59	1.71	2.65	1.89	1.92	3.56	1.44	1.11	3.43	0.90	1.09
	7	1.28	0.70	0.81	3.71	1.72	1.85	3.37	1.78	1.78	4.50	N.M.	0.91	4.26	1.00	0.99
	8	1.61	0.95	0.91	4.87	N.M.	1.84	4.72	N.M.	1.49	5.69	0.83	0.90	5.50	0.96	0.92
	9	2.10	N.M.	0.77	6.04	1.42	1.41	5.95	N.M.	1.20	6.90	N.M.	0.87	7.53	N.M.	0.95
	10	2.63	N.M.	0.68	7.34	N.M.	0.99	7.40	N.M.	1.03	8.41	N.M.	0.86	9.91	N.M.	0.97

N.M. – not measured, P – position of the alkoxy side chain, n – number of C atoms in the alkoxy side chain, k' – LC capacity factor, log {B} – infiltration anaesthesia.

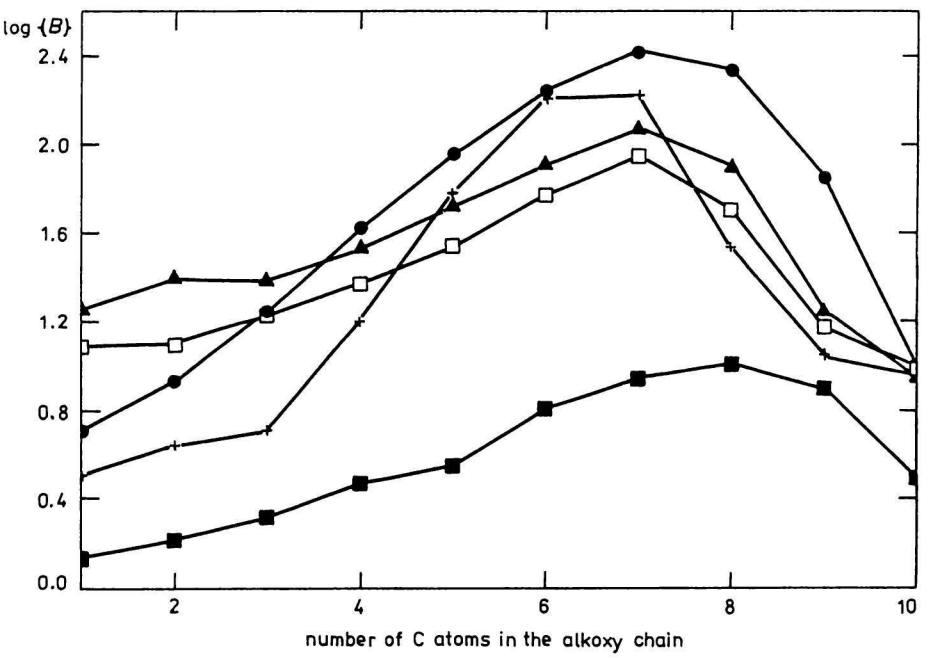


Fig. 3. The relationships between log {B} and the number of C atoms in the alkoxy side chain R for *o*-substituted drugs (see Formula 1). ■ Morpholinoethyl esters, ● dimethylaminoethyl esters, ▲ piperidinoethyl esters, □ perhydroazepinoethyl esters, + piperidinopropyl esters.

Once trained, neural network was used for supplementing of the incomplete original data matrix and also for the smoothing of the noisy biological data. The original data and data calculated by the neural network are listed in Table 2.

Figs. 3 and 4 show that the relationships between the logarithm of infiltration anaesthesia and the number of C atoms in the alkoxy side chain (resp. LC ca-

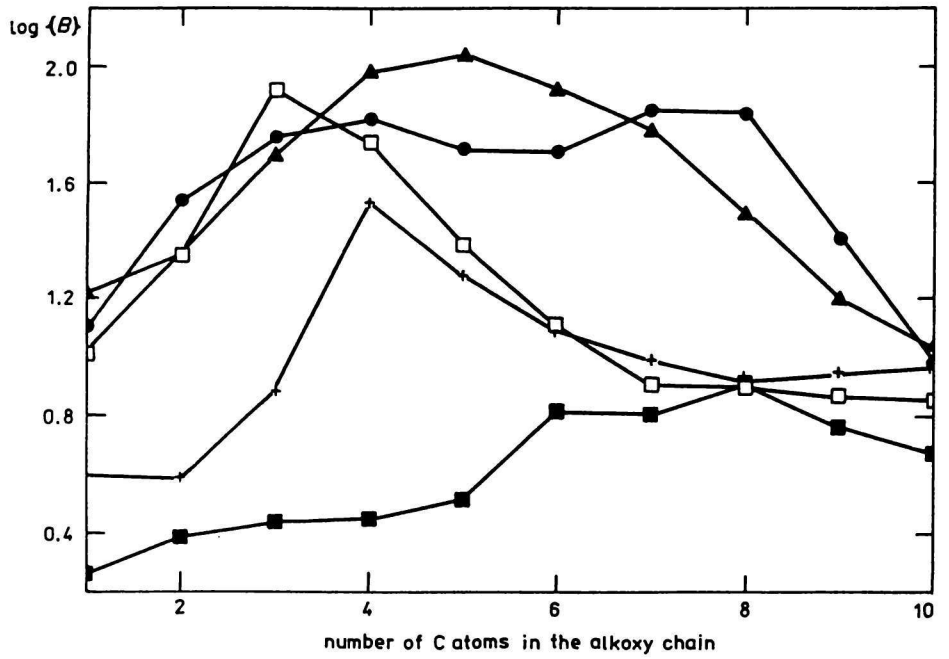


Fig. 4. The relationships between log $\{B\}$ and the number of C atoms in the alkoxy side chain R for *m*-substituted drugs (see Formula 1) of tested local anaesthetics. For curve symbols see Fig. 3.

pacity factors) for the homologous series had mostly the peak character. The first part of the curve (from the beginning up to the maximum) could be determined by increasing of biological activity with the increasing of lipophilicity (number of C atoms in the alkoxy chain). The descending part of the curve can be explained by the stereo-effect caused by the long alkoxy chain, which plays a role in the transport of drug through the cell membrane. The locations of maxima were strongly dependent on the position of the alkoxy side chain in the molecules of tested drugs (*o*- or *m*-position). The maxima of infiltration anaesthesia for the homologous series of esters of phenylcarbamic acids occurred at seven and eight carbon atoms in an alkoxy side chain. *m*-Substituted drugs presented maxima in the range from three to five carbon atoms (except of the dependence for piperidinopropyl esters which had a flat maximum at eight carbon atoms and the dependence for dimethylaminoethyl esters which had two maxima at four and seven C atoms). Generally, the infiltration anaesthetic activity of *o*-derivatives was higher than that of *m*-substituted drugs. The comparison of results reached by the classical bilinear and quadratic regression [15] and results calculated by the proposed neural network demonstrates the suitability of the neural network method for solving this problem.

The values of infiltration anaesthesia for individual homologous series of studied local anaesthetics were not always in correlation with the value of capacity factor (representing lipophilicity). Therefore, local anaesthetic activity (in this case) cannot be sufficiently characterized only by parameters corresponding with lipophilicity of tested drugs, but also by other parameters which reflect also other aspects of anaesthetical mechanism. A decreased biological activity of derivatives with a longer alkoxy substituent (despite of stronger hydrophobic interactions) can be explained as a "cut off" effect (in detail discussed by *Devínsky et al.* [16]). In the future studies we want to employ the neural network for the prediction of local anaesthesia in other homologous series of alkoxy-substituted phenylcarbamic acids which were not synthesized.

This work has demonstrated the power of neural network method for the processing of incomplete and noisy biological data. This makes it helpful in a real-world environment, such as biological systems considered in this work.

REFERENCES

1. Kuchař, M. and Rejholec, V., *Využití kvantitativních vztahů mezi strukturou a biologickou aktivitou.* (Utilization of Quantitative Relationships between Structure and Biological Activity.) P. 11. Academia, Prague, 1987.
2. Rowe, R. C., Mulley, V. J., Hughes, J. C., Nabney, J. T., and Debenham, R. M., *LC-GC 7*, 36 (1994).
3. Zupan, J. and Gasteiger, J., *Neural Networks for Chemists, An Introduction.* Verlag Chemie, Weinheim, 1993.
4. Zupan, J. and Gasteiger, J., *Anal. Chim. Acta 248*, 1 (1990).
5. Čižmárik, J., Borovanský, A., and Švec, P., *Cesk. Farm. 25*, 118 (1976).
6. Čižmárik, J., Borovanský, A., and Tumová, I., *Pharmazie 42*, 702 (1987).

7. Čižmárik, J., Borovanský, A., and Švec, P., *Acta Facultatis Pharmaceuticae Universitas Comenianae 26*, 53 (1976).
8. Čižmárik, J., Mitošinková, M., Borovanský, A., and Švec, P., *Pharmazie 33*, 509 (1978).
9. Čižmárik, J., Mazáň, Š., Novosedlíková, D., and Račanská, E., *Cesk. Farm. 41*, 130 (1992).
10. Lehotay, J., Bednáriková, A., Čižmárik, J., and Pham Thi Viet Nga, *Pharmazie 48*, 470 (1993).
11. Lehotay, J., Čižmárik, J. Pham Thi Viet Nga, and Bednáriková, A., *Pharmazie 49*, 286 (1994).
12. Bednáriková, A., Pham Thi Viet Nga, Lehotay, J., and Čižmárik, J., *Pharmazie 48*, 947 (1993).
13. Pham Thi Viet Nga, Čižmárik, J., Lehotay, J., and Bednáriková, A., *Cesk. Farm. 42*, 27 (1993).
14. Čižmárik, J., Lehotay, J. Pham Thi Viet Nga, and Bednáriková, A., *Pharmazie 48*, 149 (1993).
15. Čižmárik, J., Kráľová, K., and Loos, D., *Cesk. Farm. 44*, 154 (1995).
16. Devínsky, F., Kopecká-Leitmanová, A., Šeršeň, F., and Balgavý, P., *J. Pharm. Pharmacol. 42*, 790 (1990).

Translated by Š. Hatrík