A Neural Device for Searching Direct Correlations between Structures and Properties of Chemical Compounds

Igor I. Baskin

N. D. Zelinsky Institute of Organic Chemistry, Leninsky Prospect 47, Moscow 117813, Russia

Vladimir A. Palyulin and Nikolai S. Zefirov*

Department of Chemistry, Moscow State University, Moscow 119899, Russia

Received November 21, 1994[⊗]

A scheme of a neural device intended for searching direct correlations between structures and properties of organic compounds *without preliminary computation of molecular descriptors* (that are invariant with respect to renumbering atoms in a molecule) is suggested. The invariance of a property with respect to renumbering atoms in a molecule is ensured by the architecture of the neural device, which is constructed by analogy with biological vision systems. A model software of the neural device was tested on several examples. The descriptive and predictive performances of the device are shown to be comparable and even overcome the performances of using molecular descriptors, such as topological indexes and substructural descriptors, especially for analyzing heterogeneous data sets including inorganic compounds. The neural device can be advantageously used in the cases when more traditional approaches fail to work or "good" molecular descriptors have not been devised yet.

INTRODUCTION

Nowadays, the search for quantitative relationships between structures and properties of organic compounds is based to a considerable extent on using invariants of molecular graphs[1] such as topological indexes, quantum-chemical characteristics, numbers or indicators of some structural fragments in a molecule, etc. The important common feature of all those descriptors is the independence of their numerical values on renumbering atoms in a chemical structure. Obviously, a property of a chemical compound must not be changed with a chemist's decision to introduce his own numbering of vertexes in a molecular graph corresponding to the chemical structure. In order to perform quantitative "structure−activity" and "structure−property" (QSAR/QSPR) studies correctly, chemists have had to design a variety of molecular graph invariants.[1−4] In other words, the common QSAR/QSPR practice realizes a "two-step" scheme: structure → graph invariants → property.

In general, there exists an infinite number of such invariants, and a choice of a finite subset of them for conducting a QSAR/QSPR study is to some extent arbitrary. Thus the invariants chosen in such a manner appear not always to be effective for predicting a given property. Various statistical methods can provide, at the best, the selection of some more or less good subset of the initial set of invariants, which, in turn, is also chosen in an arbitrary way, by taking into account intuition and the availability of corresponding computer programs.

A possible way out would be to use a limited set of some "basic" invariants, so that all other invariants could be expanded in terms of them. And indeed, such a set has recently been found.[5−7] In spite of its theoretical importance, this set, unfortunately, seems to be too large to be used in practical QSAR/QSPR studies.

These ideas have led us to the development of an alternative approach *to perform QSAR/QSPR studies based on analyzing dependence of a property directly upon the elements of molecular graph connection table (known to uniquely describe a chemical structure) or, in general, upon the elements of any matrix describing properties of atoms and their pairs (including bonds)*. In other words, we have decided to realize a "one-step" scheme: structure → property. We have chosen artificial neural networks (see refs 8−10 for theory and refs 11−48 for applications to QSAR/QSPR studies) as a method for conducting such an analysis (preliminary communication[19]), because in the framework of this technique one can perceive relations between variables without having to specify their generic forms explicitly.[10,11]

Some related approaches deserve to be mentioned. Elrod, Maggiora, and Trenary[49] have used Ugi-Dugudji's BE-matrix[50] to represent a chemical structure in their studies of the reactivity of organic compounds. An extended form of the same matrix has also been used in West's studies of ^{31}P chemical shifts.[51] Kvasnička[52] has used a special neural network that reflects the topology of molecules in his studies of ^{13}C chemical shifts. A similar net has also been used by West in predicting phosphorus NMR chemical shifts.[53] In all these studies only local properties (*i.e.*, those that can be assigned to a single atom) have been considered, and the methods applied could hardly be extrapolated on the general case. A net for evaluating local atomic properties, ChemNet, in which each neuron corresponds to some atom in a chemical structure, has also been put forward by Kireev,[35] who, assuming that "the atomic invariant represents a molecular invariant as well", correlated molecular properties with local invariant calculated for some arbitrarily chosen atom. However the correlations reported in ref 35 are worse than the results obtained for analogous data sets using linear regression analysis with topological indexes or substructural descriptors.

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1997.

S0095-2338(94)00128-9 CCC: $14.00

DESCRIPTION OF THE NEURAL DEVICE

The goal of this paper is to suggest a scheme of the neural device specially suited for conducting the search for relationships between chemical structures and their properties without a need for a definite series of molecular graphs' invariants to be precomputed. Instead of molecular descriptors that are invariants with respect to renumbering atoms, we apply magnitudes corresponding to some atomic characteristics. In our approach, the invariance of properties with respect to renumbering atoms is reached due to some peculiarities in the architecture of the neural device but not because of preliminary reduction of a molecular graph connection table into a set of invariants.

Being a neural network, the device consists of a set of processing elements, called neurons, and a net of their connections, called synapses, through which they send signals to each other. Each connection is characterized by a real number, called connection weight, by which a signal gets multiplied on passing through it. Each neuron performs two principal operations: (i) sums up all incoming signals and (ii) forms an output signal through functional transformation of this sum

$$o_i = f(-\theta_i + \sum_j o_j w_{ij}) \qquad (1)$$

where o_i—the output signal of the ith neuron, o_j—the output signal of the jth neuron, w_{ij}—weight of the synapse (connection weight), through which the ith neuron transmits its output signal to the jth neuron, θ_i—activation threshold of the ith neuron, f—a transfer function usually taken to be a sigmoid one

$$f(x) = 1/(1 + \exp(-x)) \qquad (2)$$

The training of a neural network consists in finding such values of connetion weights and activation threshold that after imposing on it input signals describing a chemcial structure the output signals corresponding to predicted values of properties should be formed.

One neural device consists of three principal functional blocks: (1) a "*brain*", (2) a set of "*eyes*", and (3) one common *sensor field* ("retina") (see Figure 1). The sensor field perceives primary information about a chemical structure. The eyes receive all relevant information from the sensor field, process it, and form signals, which numerical values no more depend on the way atoms are numbered in a molecule, *i.e.*, such signals can be considered as invariants of a molecular graph. The brain gets signals from the eyes, processes them, and forms output signals corresponding to the properties being predicted. Hence, an intermediate set of molecular graphs invariants gets formed, which is constructed during the training rather than computed in advance, as in traditional approaches.

Consider more closely some parts of the neural device. The sensor field contains the description of a chemical structure. It is a squared matrix, the number of rows and columns in which is equal to the number of atoms in a chemical structure. Sensor neurons located on the diagonal of the matrix (atomic sensors) at the intersection of the ith row and the ith column form signals corresponding to some characteristics of the ith atom in the chemical structure. A set of possible atomic characteristics includes principal

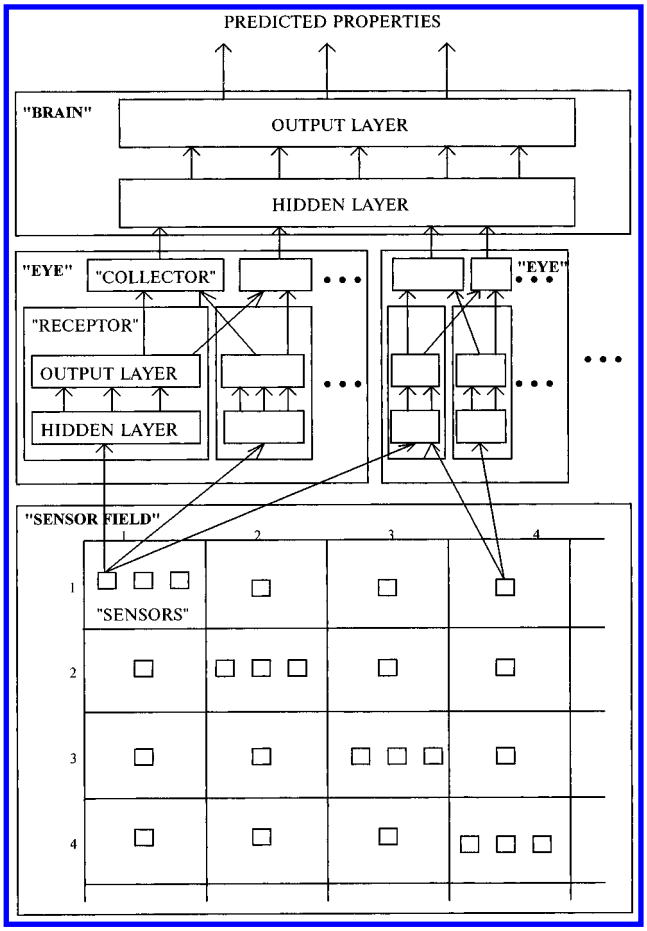

Figure 1. Architecture of the neural device.

quantum number for the corresponding chemical element, number of valence electrons, a charge on the atom, the number of hydrogen atoms attached to it, the value of its electronegativity, etc. Sensor neurons situated at the intersection of the ith row and the jth column ($i \neq j$) form signals characterizing relations between the ith and the jth atoms in the chemical structure, such as the bond order (if the atoms are linked), interatomic distance, and some others.

Each "eye" consists of (1) one or several *collectors* and (2) a set of identical "*receptors*". The important point is that all receptors have identical values of all connection weights and activation thresholds, *i.e.*, all receptors within an eye can be considered as copies of one receptor. Each receptor is a feed-forward neural network, which processes signals accepted from its rather small *receptive field*, the latter being defined as a part of the sensor field containing only a few atoms and bonds. Inside an eye, each receptor can be uniquely identified with one ordered vector $(v_1,v_2,...,v_i,...v_n)$, where n is the number of atoms in the receptive field and v_i is the ordinary number of the corresponding atom in a molecule. Such a vector will be referred to as a *receptor identifier*. In general, the number of receptors inside an eye should be equal to the number of ways the vectors could be constructed: $N!/(N - n)!$, where N is the number of non-hydrogen atoms in a chemical structure, n being the number of atoms within a receptive field (such receptors will be called *n-atomic*). For example, a three-atomic molecule can be analyzed with three one-atomic receptors with identifiers (1), (2), and (3), with six two-atomic receptors with identifiers (1,2), (2,1), (1,3), (3,1), (2,3), and (3,2), or with six three-atomic receptors with identifiers (1,2,3), (1,3,2), (2,1,3),

(2,3,1), (3,1,2), and (3,2,1). A whole neural device, with all receptors needed for analyzing a given molecule, comprises its *configuration* for the molecule. Configuration with only one receptor per each eye, which contains only mutually independent adjustable parameters, will be called *minimal*. Minimal configuration does not correspond to any particular molecule, but it rather serves as a template for the deduction of the relevant configuration for any given molecule through multiplication of receptors within eyes. It should be mentioned that the notion of the minimal configuration plays a key role in emulating the work of the neural device on an ordinary computer since only the minimal configuration of the net with a relatively small and fixed number of neurons and synapses should reside in computer memory. When training, once any adjustable parameter (connection weight or activation threshold) within a receptor takes a new value, the corresponding parameters in all other receptors within the same eye assume the same value. Because of this, training can be conceived of a minimization of an error function in the space of adjustable parameters belonging to the minimal configuration. Therefore, the minimal configuration will suffice to store all adjustable parameters of the neural device and to reproduce the whole neural device in each of its configurations.

For practical purposes, the number of receptors in an eye can be considerably reduced by imposing some additional conditions on the use of receptors. For example, one can assume that a receptor accepts signals only in the presence of a substructure within its receptive field. This enlarges the number of possible eyes but considerably reduces the number of receptors in each of them.

The signals from all the receptors inside an eye are accumulated in collectors, which are defined as neurons that sum up and transform signals received from all the receptors in the eye. Hence, the whole sensor field appears to be "seen" by an eye. Upon an arbitrary renumbering of atoms, whenever atom i acquires a new number $P(i)$, an identifier $(v_1,v_2,...,v_i,...,v_n)$ is turned to $(P(v_1),P(v_2),...,P(v_i),...,P(v_n))$. When receptors with all possible identifiers, which could be obtained in such a manner, are present in an eye, then every renumbering results in a permutation of receptors inside the eye. Since the result of the summation being made in collectors over all signals received from the receptors do not change on their permutation, then the identity of all receptors inside the eye results in the invariance of the signals being formed in the collectors with respect to the numbering of atoms in a molecule.

As it was mentioned above, each receptor is a feed-forward neural network (see, for example, ref 8), which consists of one hidden layer and one output layer. The number of hidden neurons (*i.e.*, belonging to the hidden layer) is unrestricted, while the number of output neurons is equal to the number of collectors in the eye. Each hidden neuron accepts signals from the sensors located within the corresponding receptive field, processes them, and passes the result on to every output neuron. In turn, each output neuron also processes its input signals and transmits its output signal to the corresponding collector.

The brain is also a feed-forward neural network with one hidden and one output layer. In this net, each hidden neuron receives the signals sent by every collector and after processing passes its output signal on every output neuron. Analogously, each output neuron receives the signals from the hidden layer, processes them, and forms its output signal. This signal is expected to correspond to one of the properties being predicted. The transmission of the signals over the neural device as well as its training can be described using the same mathematical expressions as in the case of an ordinary feed-forward neural network with backpropagation of errors.[8]

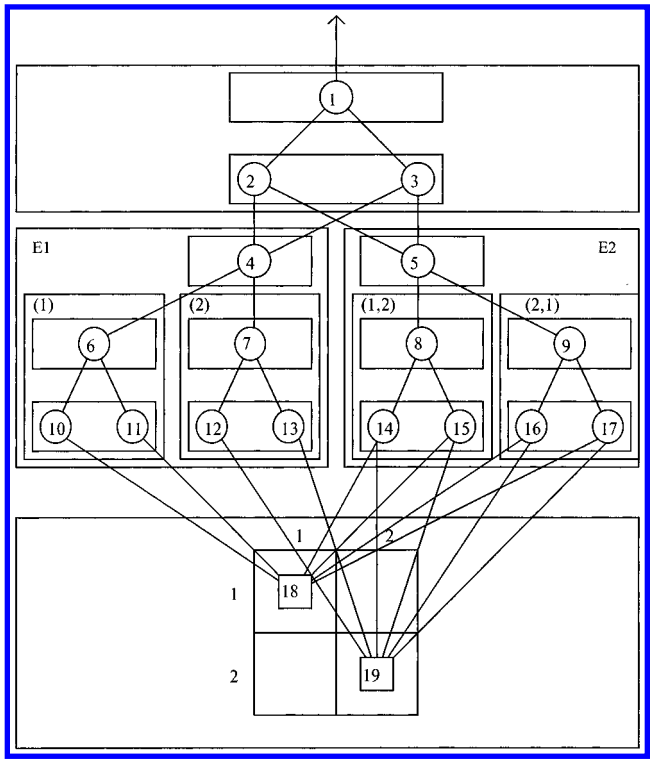

Figure 2. Neural device applied to the ethane molecule.

Hence, the structural information is processed in the neural device in four stages: (1) the primary signals that correspond to the characteristics of atoms and bonds are formed in the sensor field; (2) all the signals gathered from a receptive field are processed by the corresponding receptor; (3) signals that are invariant with respect to an arbitrary renumbering of atoms are formed in collectors; (4) the invariant signals are finally processed in the brain (see Figure 1).

It should be mentioned that the idea of using receptive fields, from which primary information would be gathered and further processed by subsequent layers of neurons to form invariants toward possible transformations of input signals, does form the basis for the *neocognitron* paradigma,[54] which is designed in compliance with neurophysiological concepts on how visual information is processed in the visual cortex.[55]

AN EXAMPLE OF THE NEURAL DEVICE

Consider an example of a neural device consisting of a brain and two eyes (named as *E1* and *E2*). Take a simple sensor field, which contains only atomic sensors (named *NH*), each of which forms only one signal corresponding to the number of hydrogens attached to that atom. Each of the receptors inside the eye *E1* receives signals from only one atomic receptor. On the contrary, each of the receptors inside the eye *E2* accepts signals from two atomic receptors corresponding to atoms that form a bond in a chemical structure. Nondiagonal elements are ignored for simplicity. Each receptor in both eyes contains two hidden and one

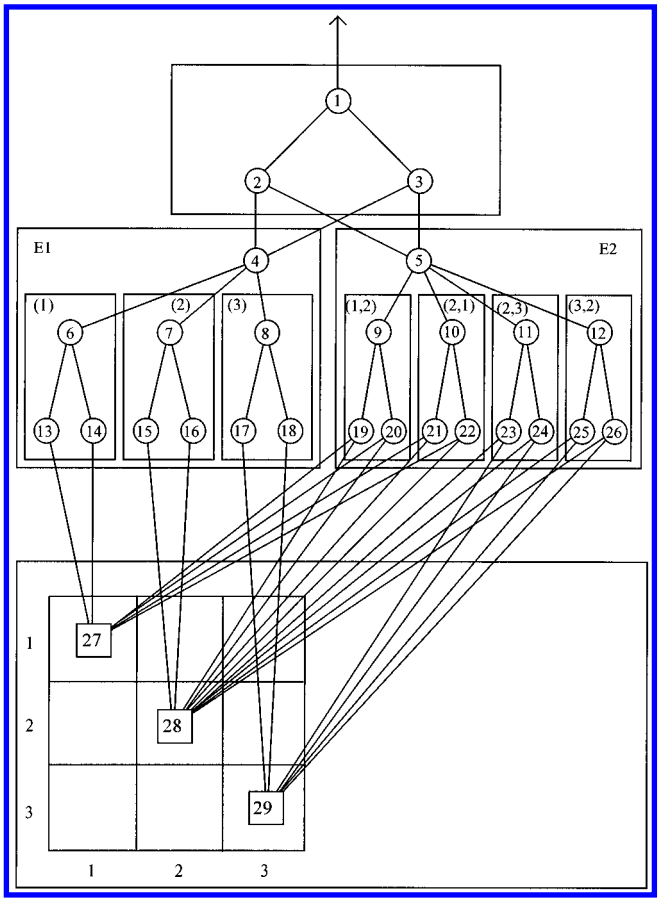

Figure 3. Neural device applied to the propane molecule.

output neuron. In correspondence with the number of output neurons, each eye contains a single collector, whose output signal is transmitted to the brain.

A configuration of the neural device for the ethane molecule is shown on Figure 2. In this case, invariance of the predicted properties with respect to renumbering atoms is provided by the following restrictions imposed on the values of connection weights ω' and activation thresholds θ': $\omega'_{4,6} = \omega'_{4,7}$; $\omega'_{6,10} = \omega'_{7,12}$; $\omega'_{6,11} = \omega'_{7,13}$; $\omega'_{5,8} = \omega'_{5,9}$; $\omega'_{8,14} = \omega'_{9,16}$; $\omega'_{8,15} = \omega'_{9,17}$; $\omega'_{10,18} = \omega'_{12,19}$; $\omega'_{11,18} = \omega'_{13,19}$; $\omega'_{14,18} = \omega'_{16,19}$; $\omega'_{14,19} = \omega'_{16,18}$; $\omega'_{15,18} = \omega'_{17,19}$; $\omega'_{15,19} = \omega'_{17,18}$; $\theta'_6 = \theta'_7$; $\theta'_8 = \theta'_9$; $\theta'_{10} = \theta'_{12}$; $\theta'_{11} = \theta'_{13}$; $\theta'_{14} = \theta'_{16}$; $\theta'_{15} = \theta'_{17}$.

In a similar manner, a configuration of the neural device for the propane molecule is given on Figure 3. The invariance of the predicted properties with respect to renumbering atoms is ensured by the following restrictions imposed on the values of connection weights ω'' and activation thresholds θ'': $\omega''_{4,6} = \omega''_{4,7} = \omega''_{4,8}$; $\omega''_{6,13} = \omega''_{7,15} = \omega''_{8,17}$; $\omega''_{6,14} = \omega''_{7,16} = \omega''_{8,18}$; $\omega''_{13,27} = \omega''_{15,28} = \omega''_{17,29}$; $\omega''_{14,27} = \omega''_{16,28} = \omega''_{18,29}$; $\omega''_{5,9} = \omega''_{5,10} = \omega''_{5,11} = \omega''_{5,12}$; $\omega''_{9,19} = \omega''_{10,21} = \omega''_{11,23} = \omega''_{12,25}$; $\omega''_{9,20} = \omega''_{10,22} = \omega''_{11,24} = \omega''_{12,26}$; $\omega''_{19,27} = \omega''_{21,28} = \omega''_{23,28} = \omega''_{25,29}$; $\omega''_{19,28} = \omega''_{21,27} = \omega''_{23,29} = \omega''_{25,28}$; $\omega''_{20,27} = \omega''_{22,28} = \omega''_{24,28} = \omega''_{26,29}$; $\omega''_{20,28} = \omega''_{22,27} = \omega''_{24,29} = \omega''_{26,28}$; $\theta''_6 = \theta''_7 = \theta''_8$; $\theta''_9 = \theta''_{10} = \theta''_{11} = \theta''_{12}$; $\theta''_{13} = \theta''_{15} = \theta''_{17}$; $\theta''_{14} = \theta''_{16} = \theta''_{18}$; $\theta''_{19} = \theta''_{21} = \theta''_{23} = \theta''_{25}$; $\theta''_{20} = \theta''_{22} = \theta''_{24} = \theta''_{26}$.

All these configurations can be deduced from the minimal one presented on Figure 4 by multiplication of receptors.

RESULTS AND DISCUSSION

Boiling Points of Alkanes. In our first experiment with a program emulator of the neural device, we have chosen

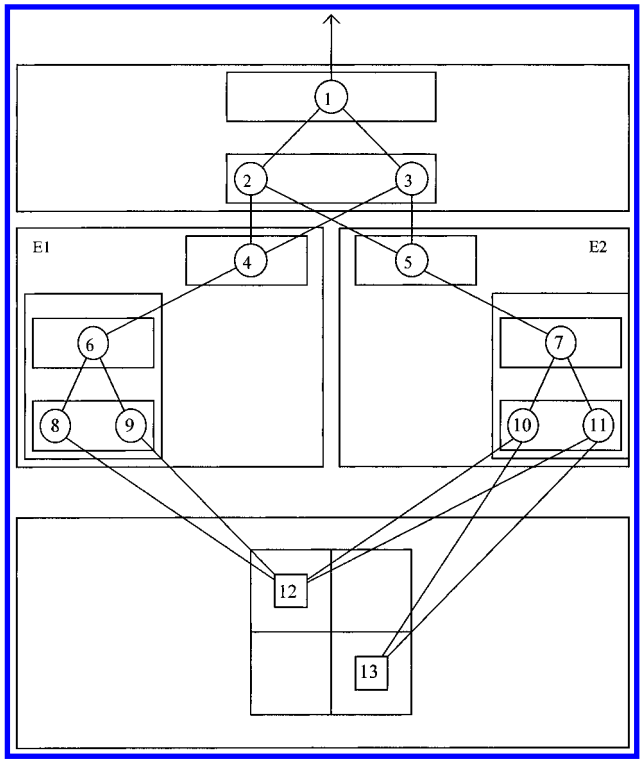

Figure 4. Minimal configuration of the neural device.

prediction of boiling points of alkanes, so far as by using this example it is possible to make comparison with results reached by authors of numerous publications. Neural networks have also been applied by several authors[18,30,35,38] for studying this problem. A set consisting of 74 alkanes[56] C_2-C_9 was divided into two parts—a training set (67 compounds) and a validation set (seven compounds). A neural network with architecture described in the foregoing example was used in this study. A "generalized delta-rule" procedure[57] in the aforementioned modification was used for training. We used a value of 0.05 for the learning rate and 0.9 for the momentum parameter (see ref 57). Training was interrupted on reaching the value of the correlation coefficient between predicted and observed boiling points of 0.994. This value is better than any correlation coefficient of the same property for the same data set with a single topological index and is comparable to the correlation coefficients that could be reached for multiple regression models including several topological indexes. The same conclusions could also be drawn by considering root-mean-square (RMS) errors on both sets (5.2 degrees on the training set and 5.1 degrees on the test set). However this result appeared to be worse than that obtained by using simple feed-forward neural networks for correlating the same properties with topological indexes or occurrence numbers of several substructures.[18] In addition, the time needed for the training appeared to be too long (up to several hours on PC). This result was reported in our preliminary communcation.[19]

To address the above-mentioned problems, we have examined the influence of the architecture of the neural device on its performance. The number of collectors within eyes has appeared to have a drastic effect on the time needed for learning (few minutes instead of hours on Pentium-100) as well as on the quality of learning. The same architecture of the neural device but with fire collectors within each of its two eyes gives an excellent value of the correlation coefficient, 0.9994, and very small RMS errors of 1.6 degrees

on the training set and 2.4 degree on the validation set. Such performance is comparable with best results reported for predicting boiling points of alkanes. Nonetheless, further increase in the number of collectors deteriorates results.

One further enhancement that affects the learning of the neural device involves the use of separate values of learning factors for eyes and brain. For stable learning, we used the value of the learning factor for eyes 10 times less than for the brain. As regards absolute values of the learning factor for the brain, its starting value of 0.25 and 0.05 at the end of learning appeared to be optimal.

Viscosity of Hydrocarbons. In the following example, a more diverse set of 81 hydrocarbons[58] C_6-C_{21}, cyclic and acyclic, saturated and unsaturated, aromatic and aliphatic, was used in evaluating their viscosity at 40 °C. Again, as in the previous example, it was split into a training set (65 compounds) and a validation set (16 compounds). A neural device containing a brain with two hidden neurons and only one eye *E2* with three hidden neurons in each receptor and five collectors was chosen for this study. After 1100 epochs of training the correlation coefficient became 0.996, and RMS error for the training set was 0.15 centipoises and for the validation set 0.18 centipoises.

Heat of Evaporation of Hydrocarbons. In the following example, a set of 267 hydrocarbons[59] C_4-C_{26}, also, as in the previous example, cyclic and acyclic, saturated and unsaturated, aromatic and aliphatic, was used for training the neural device to predict heat of formation. From this set, 54 compounds were randomly selected as a validation set, while remaining 213 compounds formed a training set. A neural device containing a brain with three hidden neurons and two eyes, *E1* and *E2*, each containing three hidden neurons in each receptor and three collectors, was used in the study. Again, as in all previous examples, only sensors which form signals corresponding to the number of hydrogens attached to a given atom were included in the sensor field. There was no "overtraining", so the learning was interrupted after 2600 epochs with a correlation coefficient of 0.996 and RMS error on the training set was 1.44 kJ/mol and 1.26 kJ/mol on the validation set. In this case, the neural device outperforms results obtained for the same data.[59]

Density of Hydrocarbons. The following example deals with predicting density of liquid hydrocarbons. A set of 141 hydrocarbons[58] C_5-C_8 (saturated and unsaturated, cyclic and acyclic, aromatic and aliphatic) was divided into a training set with 133 compounds and a validation set with 28 compounds. A neural device containing a brain with five hidden neurons and two eyes, *E1* and *E2*, each containing five hidden neurons in each receptor and five collectors, was used in the study. The sensors chosen were the same as in all previous examples. After 1700 cycles of training the value of the correlation coefficient reached 0.971, the RMS error for the training set became 0.018 g/cm^3 and 0.019 g/cm^3 for the validation set.

Heat of Solvation in Cyclohexane. Unlike all previous cases, compounds of the following set[59] belong to different classes of organic compounds. In accordance with results of preliminary studies, one compound, perfluorobenzene, was excluded as an outlier, while another 140 compounds were divided into a training set (112 compounds) and a validation set (28 compounds). A neural device with three eyes, *E1*, *E2*, and *E3*, was constructed. Eyes *E1* and *E2* are the same as described above, while *E3* contains receptors each receiving signals from three atoms linked with two bonds. To discriminate between heteroatoms, besides *NH*, one additional type of atomic sensors, which detects principal quantum number, *PQN*, was used. The brain of the neural device and all receptors were chosen to contain three hidden neurons. Furthermore, three collectors were placed in each of three eyes. After 10 000 epochs of learning, the value of the correlation coefficient amounted to 0.990, while the RMS error on the training set became 1.77 kJ/mol and on the validation set 2.46 kJ/mol. Since, as it was shown in a previous article,[59] commonly used topological indexes were incapable of giving strong correlations with the heat of solvatation, a special "solvation index" was designed[59] to be used in linear regression ($R = 0.985$ and $s = 2.1$ kJ/mol for the same data set). This example shows that the neural device can compete with using specially designed molecular descriptors.

Polarizability of Different Molecules. The following example involves diverse data set[60] containing organic compounds (up to 26 non-hydrogen atoms per molecule) belonging to different classes as well as inorganic compounds, e.g., N_2O, SO_2, H_2S, O_2, N_2, NH_3, Cl_2, etc. The data set was split at random into a training set (235 compounds) and a validation set (58 compounds). An architecture of the neural device was chosen to contain a brain with three hidden neurons and only one eye *E1* with three hidden neurons in each receptor and five collectors. Three types of atomic sensors were used: *NH*, *AR*, and *NE*. Sensor *NH* forms signals corresponding to the number of hydrogen atoms attached to a given atom, signals of sensor *AR* correspond to atomic radius, sensor *NE* gives signals corresponding to the number of electrons in the atom. The value of the correlation coefficient after 2000 epochs of learning became 0.995, and the RMS error on the training set was 0.86 cm^3 and 0.71 cm^3 on the validation set. This result is significantly better than all models we managed to build using a set of commonly used topological indexes in the framework of either multiple linear regression methodology or neural networks with simple architectures. Although group additivity methods can also be applied in this case,[60] their usage is however confined to molecules containing only those groups that are sufficiently represented in the training set, while this neural device does not require this precondition to be satisfied.

Anesthetic Pressure of Gases. The last example in this study illustrates the ability of the neural device to predict biological activity. A data base consisting of hydrocarbons, halogenhydrocarbons, and some inorganic gases, such as all noble gases, molecular nitrogen, SF_6, and N_2O, was taken from the review.[61] As in all previous examples, the data base was split into a training set (24 compounds) and a validation set (six compounds). A neural network with a brain, which contained three hidden neurons, and one eye *E1*, in which each receptor "sees" only one atom, with three hidden neurons in each receptor and five collectors was constructed. Three types of atomic sensors were used: *NH*, *PQN*, and *VE*. The first two sensors, *NH* and *PQN*, are described above; sensor *VE* detects number of valence electrons in an atom. After 4000 epochs of learning the value of the correlation coefficient became 0.990, and the RMS error on the training set was 0.18 log units (log(1/*p*)) and 0.26 log units on the validation set. For this example, we failed to find any significant correlation of the anesthetic

pressure either with topological indexes or using group contribution method.

CONCLUSIONS

Our aim was to demonstrate the possibility of constructing a neural device for correlating properties of organic compounds directly with their structures without having to select and compute topological indexes or any other molecular descriptors in advance. Instead, we use descriptors applied to individual atoms and bonds in molecules. In other words, this methodology constitutes an alternative to the use of molecular descriptors in QSAR/QSPR studies. From the other hand, this approach can easily incorporate the use of molecular descriptors if the latter are proved to be especially useful for solving some particular problem: this can be done by feeding the brain of the neural device with additional signals corresponding to values of molecular descriptors.

Since the output signals of the neural device as well as output signals of all its collectors do not depend upon atomic numbering and hence can be viewed as some molecular descriptors, the whole neural network can also be regarded as a tool for designing new molecular descriptors which would as close as possible fit molecular properties. When learning, the neural device is trying to find a way how local atomic and interatomic properties can be combined to give molecular descriptors capable of correlating with a given molecular property. Clearly there is no need to invent new descriptors if such or even better descriptors are already known. Therefore the use of the neural device in the cases when sufficiently good molecular descriptors are still not known would be worthwhile.

ACKNOWLEDGMENT

We gratefully acknowledge support from International Association for the Promotion of Cooperation with Scientists from the Independent States of the Former Sovient Union (INTAS grant No. 93-3233) and from Russian Foundation for Basic Research (Grant No. 95-03-09696). Also we wish to thank N. Arslanov, I. Atovmian, and A. Zefirov for providing us structure−property data bases.

REFERENCES AND NOTES

(1) Stankevich, M. I.; Stankevich, I. V.; Zefirov, N. S. Topological Indexes in Organic *Chemistry Usp. Khimii* **1988**, *57*, 337−366.
(2) Rouvray, D. H. Should We Have Designs on Topological Indexes? *Stud. Phys. Theor. Chem.* **1983**, *28 (Chem. Appl. Topol. Graph Theory)*, 159−177.
(3) Balaban, A. Chemical Graphs. XXXIV. Five New Topological Indices for the Branching of Tree-like Graphs. *Theor. Chim. Acta* **1979**, *53*, 355−375.
(4) Seybold, P. G.; May, M.; Bagal, U. A. Molecular Structure-Property Relationships. *J. Chem. Educ.* **1987**, *64*, 575−581.
(5) Mnukhin, V. B. Basis of Algebra of Graph Invariants. In *Mathematical Analysis and its Applications*; Rostov-na-Donu, 1983; pp 55−60.
(6) Skvortsova, M. I.; Baskin, I. I.; Slovokhotova, O. L.; Zefirov, N. S. Methodology for Building a General "Structure-Property" Model on Topological Level. *Dokl. Akad. Nauk.* **1994**, *336*, 496−499.
(7) Baskin, I. I.; Skvortsova, M. I,; Stankevich, I. V.; Zefirov, N. S. On basis of invariants of labeled molecular graphs. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 527−531.
(8) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*; Rumelhart, D. E., McClelland, J. L., Eds.; MIT Press: Cambridge, 1986.
(9) Hinton, G. E. Connectionist Learning Procedures. *Artif. Intell.* **1989**, *40*, 185−234.
(10) Kosko, B. Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence; Prentice Hall: Englewood Cliffs, NJ, 1992.
(11) Maggiora, G. M.; Elrod, D. W.; Trenary, R. G. Computational Neural Networks as Model-Free Mapping Devices. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 732−741.
(12) (a) Gasteiger, J.; Zupan, J. Neural Networks in Chemistry. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 503−527. (b) Zupan, J.; Gasteiger, J. Neural Networks for Chemists−An Introduction; VCH Publishers: 1993.
(13) Burns, J. A.; Whitesides, G. M. Feed-forward Neural Networks in Chemistry: Mathematical Systems for Classification and Pattern Recognition. *Chem. Rev. (Washington, D.C.)* **1993**, *93*, 2583−2601.
(14) Ajay, A. Unified Framework for Using Neural Networks to Build QSARs. *J. Med. Chem.* **1993**, *36*, 3565−3671.
(15) (a) Andrea, T. A.; Kalayeh, H. Application of Neural Networks in Quantitative Structure-Activity Relationships of Dihydrofolate Reductase Inhibitors. *J. Med. Chem.* **1991**, *34*, 2824−2836. (b) Andrea, T. A.; Kalayeh, H. Application of Neural Networks in Quantitative Structure-Activity Relationships. *Pharmacochem. Libr.* **1991**, *16*, 209−212. (c) Andrea, T. A. Novel Structure−Activity Insights from Neural Network Models. *ACS Symp. Ser.* **1995**, *606(Classical and Three-Dimensional QSAR in Agrochemistry)*, 282−287.
(16) (a) Aoyama, T.; Suzuki, Y.; Ichikawa, H. Neural Networks Applied to Pharmaceutical Problems. I. Method and Application to Decision Making. *Chem. Pharm. Bull.* **1989**, *37*, 2558−2560. (b) Aoyama, T.; Suzuki, Y.; Ichikawa, H. Neural Networks Applied to Structure−Activity Relationships. *J. Med. Chem.* **1990**, *33*, 905−908. (c) Aoyama, T.; Suzuki, Y.; Ichikawa, H. Neural Networks Applied to Pharmaceutical Problems. III. Neural Networks Applied to Quantitative Structure−Activity Relationship (QSAR) Analysis. *J. Med. Chem.* **1990**, *33*, 2583−2590. (d) Aoyama, T.; Ichikawa, H. Neural Networks Applied to Pharmaceutical Problems. IV. Basic Operating Characteristics of Neural Networks When Applied to Structure−Activity Studies. *Chem. Pharm. Bull.* **1991**, *39*, 358−366. (e) Aoyama, T.; Ichikawa, H. Neural Networks Applied to Pharmaceutical Problems. V. Obtaining the Correlation Indexes Between Drugs Activity and Structural Parameters Using a Neural Network. *Chem. Pharm. Bull.* **1991**, *39*, 372−378. (f) Aoyama, T.; Ichikawa, H. Neural Networks as Nonlinear Structure−Activity Relationship Analyzers. Useful Functions of the Partial Derivative Method in Multilayer Neural Networks. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 492−500. (g) Ichikawa, H.; Aoyama, T. How to See Characteristics of Structural Parameters in QSAR Analysis: Descriptor Mapping Using Neural Networks. *SAR QSAR Environ. Res.* **1993**, *1*, 115−130.
(17) Balaban, A. T.; Basak, S. C.; Colburn, T.; Grunwald, G. D. Correlation between Structure and Normal Boiling Points of Haloalkanes C_1−C_4 Using Neural Networks. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1118−1121.
(18) Baskin, I. I.; Palyulin, V. A.; Zefirov, N. S. Computational Neural Networks as an Alternative to the Linear Regression Analysis in the Studies of Quantitative "Structure−Property" Relationships for the Case of Physico-Chemical Properties of Hydrocarbons. *Dokl. Akad. Nauk* **1993**, *332*, 713−716.
(19) Baskin, I. I.; Palyulin, V. A.; Zefirov, N. S. A Methodology for Searching Direct Correlations between Structures and Properties of Organic Compounds by Using Computational Neural Networks. *Dokl. Akad. Nauk* **1993**, *333*, 176−179.
(20) Baskin, I. I.; Lyubimova, I. K.; Abilev, S. K.; Palyulin, V. A.; Zefirov, N. S. Quantitative Structure−Activity Relationship Study of Mutagenic Activity of Heterocyclic Derivatives of Pyrenes and Phenantrenes. *Dokl. Akad. Nauk* **1994**, *339*, 106−108.
(21) Bienfait, B. Applications of High-Resolution Self-Organizing Maps to Retrosynthetic and QSAR Analysis. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 890−898.
(22) Bodor, N.; Harget, A.; Huang, M.-J. Neural Network Studies. 1. Estimation of the Aqueous Solubility of Organic Compounds. *J. Am. Chem. Soc.* **1991**, *113*, 9480−9483.
(23) Brewster, M. E.; Huang, M.-J.; Harget, A.; Bodor, N. Neural Network Studies. 2. Use of a Neural Net to Estimate Oxidation Energies for Substituted Dihydropyridines and Related Heterocycles. *Tetrahedron* **1992**, *48*, 3463−3472.
(24) (a) Brinn, M. W.; Payne, M. P.; Walsh, P. T. Neural Network Prediction of Mutagenicity Using Structure−Property Relationships. *Chem. Eng. Res. Des.* **1993**, *71*, 337−339. (b) Brinn, M.; Walsh, P. T.; Payne, M. P.; Bott, B. Neural Network Classification of Mutagens Using Structural Fragment Data. *SAR QSAR Environ. Res.* **1993**, *1*, 169−210.
(25) Bruchmann, A.; Zinn, P.; Haffer, C. M. Prediction of Gas Chromatographic Retention Index Data by Neural Networks. *Anal. Chim. Acta* **1993**, *283*, 869−880.
(26) Cambon, B. Modeling the Biological Activity of PAH by Neural Networks. *Polycyclic* Aromat. Compd. **1993**, *3(Suppl.)*, 257−265.
(27) Chastrette, M.; De Saint Laumer, J. Y. Structure-Odor Relationships Using Neural Networks. *Eur. J. Med. Chem.* **1991**, *26*, 829−833.
(28) Cherqaoui, D.; Villemin, D.; Mesbah, A.; Cense, J. M.; Kvasnička, V. Use of a Neural Network to Determine the Normal Boiling Points

of Acyclic Organic Molecules Containing Heteroatoms. *AIP Conf. Proc.* **1995**, *330(E.C.C.C. 1 Computational Chemistry)*, 767−773.
(29) Domine, D.; Devillers, J.; Chastrette, M.; Karcher, W. Estimating Pesticide Field Half-Lives from a Backpropagation Neural Network. *SAR QSAR Environ. Res.* **1993**, *1*, 211−219.
(30) (a) Gakh, A. A.; Gakh, E. G.; Sumpter, B. G.; Noid, D. W. Neural Network−Graph Theory Approach to the Prediction of the Physical Properties of Organic Compounds. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 832−839. (b) Gakh, A. A.; Gakh, E. G.; Sumpter B. G.; Noid, D. W.; Trowbridge, L. D. Estimation of the Properties of Hydrofluorocarbons by Computer Neural Networks. *J. Fluorine Chem.* **1995**, *73*, 107−111.
(31) Ghoshal, N.; Mukhopadhayay, S. N.; Ghoshal, T. K.; Achari, B. Quantitative Structure-Activity Relationship Studies Using Artificial Neural Networks. *Indian J. Chem., Sect. B* **1993**, *32B*, 1045−1050.
(32) Glen, R. C.; Rose, V. S.; Lindon, J. C.; Ruane, R. J.; Wilson, I. D.; Nicholson, J. K. Quantitative Structure−Chromatography Relationships: Prediction of TLC Behavior Using Theoretically Derived Molecular Properties. *J. Planar Chromatogr.-Mod. TLC* **1991**, 432-438.
(33) Grunenberg, J.; Herges, R. Prediction of Chromatographic Retention Values (R_M) and Partition Coefficients (log P_{oct}) Using a Combination of Semiempirical Self-Consistent Reaction Field Calculations and Neural Networks. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 905−911.
(34) (a) King, R. D.; Hirst, J. D.; Sternberg, M. J. E. New Approaches to QSAR: Neural Networks and Machine Learning. *Perspect. Drug. Discovery Des.* **1993**, *1*, 279−290. (b) Hirst, J. D.; King, R. D.; Sternberg, M. J. E. Quantitative Structure−Activity Relationships by Neural Networks and Inductive Logic Programming. I. The Inhibition of Dihydrofolate Reductase by Pyrimidines. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 405−420. (c) Hirst, J. D.; King, R. D.; Sternberg, M. J. E. Quantitative Structure-Activity Relationships by Neural Networks and Inductive Logic Programming. II. The Inhibition of Dihydrofolate Reductase by Triazines. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 421−432.
(35) Kireev, D. B. ChemNet: A Novel Neural Network Based Method for Graph/Property Mapping. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 175−180.
(36) Lee, M. J.; Chen, J. T. Fluid Property Prediction with the Aid of Neural Networks. *Ind. Eng. Chem. Res.* **1993**, *32*, 995−997.
(37) Liu, Q.; Hirono, S.; Moriguchi, I. Comparison of the Functional-Link Net and the Generalized Delta Rule Net in Quantitative Structure−Activity Relationship Studies. *Chem. Pharm. Bull.* **1992**, *40*, 2962−2969.
(38) Lohninger, H. Evaluation of Neural Networks Based on Radial Basis Functions and their Application to the Prediction of Boiling Points from Structural Parameters. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 736−744.
(39) (a) Manallack, D. T.; Ellis, D. D.; Livingstone, D. J. Analysis of Linear and Nonlinear QSAR Data Using Neural Networks. *J. Med. Chem.* **1994**, *37*, 3758−3767. (b) Manallack, D. T.; Livingstone, D. J. Neural Networks and Expert Systems in Molecular Design. *Methods Princ. Med. Chem.* **1995**, *3(Advanced Computer-Assisted Techniques in Drug Discovery)*, 293−318. (c) Manallack, D. T.; Livingstone, D. J. Relating Biological Activity to Chemical Structure Using Neural Networks. *Pestic. Sci.* **1995**, *45*, 167−170.
(40) (a) Nefati, H.; Diawara, B.; Legendre, J. J. Predicting the Impact Sensitivity of Explosive Molecules Using Neuromimetic Networks. *SAR QSAR Environ. Res.* **1993**, *1*, 131−136. (b) Nefati, H.; Legendre, J.-J.; Michot, C. Impact Sensitivity Prediction by the Means of Feed-Forward Multilayer Neural Networks. *AIP Conf. Proc.* **1995**, *330-(E.C.C.C. 1 Computational Chemistry)*, 562−567.
(41) (a) Peterson, K. L. Counter-propagation Neural Networks in the Modeling and Prediction of Kovats Indexes for Substituted Phenols. *Anal. Chem.* **1992**, *64*, 379−386. (b) Peterson, R.; Fredenslund, A.; Rasmussen, P. Artificial Neural Networks as a Predictive Tool for Vapor-Liquid Equilibrium. *Comput. Chem. Eng.* **1993**, *18(Suppl.)*, 563−567. (c) Peterson, K. L. Quantitative Structure−Activity Relationships in Carboquinones and Benzodiazepines Using Counter-Propagation Neural Networks. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 896−904.
(42) Rose, V. S.; Croall, I. F.; MacFie, H. J. An Application of Unsupervised Neural Network Methodology (Kohonen-Topology Preserving Mapping) to QSAR Analysis. *Quant. Struct.-Act. Relat.* **1991**, *10*, 6−15.
(43) Salt, D. W.; Yildiz, N.; Livingstone, D. J.; Tinsley, C. J. The Use of Artificial Neural Networks in QSAR. *Pestic. Sci.* **1992**, *36*, 161−170.
(44) Smits, J. R. M.; Melssen, W. J.; Daalmans, G. J.; Kateman, G. Using Molecular Representation with Neural Networks. A Case Study: Prediction of the HPLC Retention Index. *Comput. Chem.* **1994**, *18*, 157−172.
(45) (a) So, S.-S.; Richards, W. G. Application of Neural Networks: Quantitative Structure−Activity Relationships of the Derivatives of 2,4-Diamino-5-(substituted-benzyl)pyrimidines as DHFR Inhibitors. *J. Med. Chem.* **1992**, *35*, 3201−3207. (b) So, S.-S.; Karplus, M. Evolutionary Optimization in Quantitative Structure−Activity Relationship: An Application of Genetic Neural Networks. *J. Med. Chem.* **1996**, *39*, 1521−1530.
(46) (a) Tetko, I. V.; Luik, A. I.; Poda, G. I. Application of Neural Networks in Structure−Activity Relationships of a Small Number of Molecules. *J. Med. Chem.* **1993**, *36*, 811−814. (b) Tetko, I. V.; Tanchuk, V. Yu.; Chentsova, N. P.; Antonenko, S. V.; Poda, G. I.; Kukhar, V. P.; Luik, A. I. HIV-1 Reverse Transcriptase Inhibitor Design Using Artificial Neural Networks. *J. Med. Chem.* **1994**, *37*, 2520−2526. (c) Tetko, I. V.; Tanchuk, V. Yu.; Luik, A. I. Anti-AIDS Drug Design with the Help of Neural Networks. *AIP Conf. Proc.* **1995**, *330-(E.C.C.C. 1 Computational Chemistry)*, 761−766.
(47) Weinstein, J. N.; Kohn, K. W.; Grever, M. R.; Viswanadhan, V. K.; Rubinstein, L. V.; Monks, A. P.; Scudiero, D. A.; Welch, L.; Koutsokos, A. D.; Chiausa, A. J.; Paull, K. D. Neural Computing in Cancer Drug Development: Predicting Mechanisms of Action. *Science* **1992**, *258*, 447−451.
(48) Wiese, M.; Schaper, K.-J. Application of Neural Network in the QSAR Analysis of Percent Effect Biological Data: Comparison with Adaptive Least Squares and Nonlinear Regression Analysis. *SAR QSAR Environ. Res.* **1993**, *1*, 137−152.
(49) (a) Elrod, D. W.; Maggiora, G. M.; Trenary, R. G. Application of Neural Networks in Chemistry. 1. Prediction of Electrophilic Aromatic Substitution Reactions. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 477−484. (b) Elrod, D. W.; Maggiora, G. M.; Trenary, R. G. Application of Neural Networks in Chemistry. 2. A General Connectivity Representation for the Prediction of Regiochemistry. *Tetrahedron Comput. Methodol.* **1990**, *3*, 163−174.
(50) Dugundji, J.; Ugi, I. An Algebraic Model of Constitutional Chemistry as a Basis for Chemical Computer Programs. *Top. Curr. Chem.* **1973**, *39*, 19−64.
(51) West, G. Empirical ^{31}P Spectrum Prediction by Neural Networks. In *NATO-ASI Molecular Spectroscopy: Recent Experimental and Computational Advances*; Ponta Delgada, 1992.
(52) Kvasnička, V. An Application of Neural Networks in Chemistry. Prediction of ^{13}C NMR Chemical Shifts. *J. Math. Chem.* **1991**, *6*, 63−76.
(53) West, G. M. J. Predicting Phosphorus NMR Shifts Using Neural Networks. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 577−589.
(54) (a) Fukushima, K. Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Uneffected by Shift in Position. *Biological Cybernetics* **1980**, *36*, 193−202. (b) Fukushima, K.; Miyake, S. Neocognitron: A New Algorithm for Pattern Recognition Tolerant of Deformations and Shifts in Position. *Pattern Recognition* **1982**, *15*, 455−469. (c) Fukushima, K. A. Hierarchical Neural Network Model for Associative Memory. *Biol. Cybernetics* **1984**, *50*, 105−113. (d) Fukushima, K. A Neural Network Model for Selective Attention in Visual Pattern Recognition. *Biol. Cybernetics* **1986**, *55*, 5−15.
(55) (a) Hubel, D. H.; Wiesel, T. N. Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex. *J. Physiol.* **1962**, *160*, 106−154. (b) Hubel, D. H.; Wiesel, T. N. Receptive Fields and Functional Architecture in Two Nonstriate Visual Areas (18 and 19) of the Cats. *J. Neurophysiol.* **1965**, *28*, 229−289. (c) Hubel, D. H.; Wiesel, T. N. Functional Architecture of Macaque Monkey Visual Cortex. *Proc. Roy. Soc. London Ser. B* **1977**, *198*, 1−59.
(56) Needham, D. E.; Wei, I. C.; Seybold, P. G.; Molecular Modeling of the Physical Properties of the Alkanes. *J. Am. Chem. Soc.* **1988**, *110*, 4186−4194.
(57) Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning Internal Representations by Back-Propagating Errors. *Nature* **1986**, *33*, 533−536.
(58) Rossini, F. D.; Pitzer, K. S.; Arnett, R. L.; Braun, R. M.; Pimentel, G. C. *Selected Values of Physical and Thermodynamic Properties of Hydrocarbons and Related Compounds*; Carnegie Press: Pittsburgh, PA, 1953.
(59) Antipin, I. S.; Arslanov, N. A.; Palyulin, V. A.; Konovalov, A. I.; Zefirov, N. S. Solvation Topological Index. Topological Model of Description of Disperse Interactions. *Dokl. Akad. Nauk SSSR* **1991**, *316*, 925−928.
(60) Miller, K. J. Additivity Methods in Molecular Polarizability. *J. Am. Chem. Soc.* **1990**, *112*, 8533−8542.
(61) Gupta, S. P. QSAR Studies on Drugs Acting at the Central Nervous System. *Chem. Rev.* **1989**, *89*, 1765−1800.

CI940128Y