Volume Learning Algorithm Artificial Neural Networks for 3D QSAR Studies

Igor V. Tetko,*[,†,‡] Vasyl V. Kovalishyn,[†] and David J. Livingstone[§]

Biomedical Department, Institute of Bioorganic & Petroleum Chemistry, Murmanskaya 1, Kiev-660, 253660 Ukraine, Laboratoire de Neuro-Heuristique, Institut de Physiologie, Université de Lausanne, Rue du Bugnon 7, Lausanne, CH-1005 Switzerland, ChemQuest, Delamere House, 1, Royal Crescent, Sandown, Isle of Wight, PO36 8LZ U.K., and Centre for Molecular Design, University of Portsmouth, Portsmouth, Hants, PO1 2EG U.K.

Received February 20, 2001

The current study introduces a new method, the volume learning algorithm (VLA), for the investigation of three-dimensional quantitative structure−activity relationships (QSAR) of chemical compounds. This method incorporates the advantages of comparative molecular field analysis (CoMFA) and artificial neural network approaches. VLA is a combination of supervised and unsupervised neural networks applied to solve the same problem. The supervised algorithm is a feed-forward neural network trained with a back-propagation algorithm while the unsupervised network is a self-organizing map of Kohonen. The use of both of these algorithms makes it possible to cluster the input CoMFA field variables and to use only a small number of the most relevant parameters to correlate spatial properties of the molecules with their activity. The statistical coefficients calculated by the proposed algorithm for cannabimimetic aminoalkyl indoles were comparable to, or improved, in comparison to the original study using the partial least squares algorithm. The results of the algorithm can be visualized and easily interpreted. Thus, VLA is a new convenient tool for three-dimensional QSAR studies.

Introduction

Comparative molecular field analysis (CoMFA)[1,2] is a 3D QSAR method that efficiently handles shape-dependent pharmacodynamic interactions. Among the many innovative approaches, such as Hans-Dieter Höltje's calculations of interaction energies in hypothetical drug−receptor complexes,[3] the active analogue approach of Garland Marshall,[4] or the GRID program of Peter Goodford,[5] CoMFA is the most widely practiced.[6−8] Usually, analysis of CoMFA field variables is performed by partial least squares (PLS).[9] In recent years, a number of other statistical methods were used in CoMFA studies, such as genetic algorithms[10] or *k*-nearest neighbor methods.[11] PLS gives good results when the correlation between activity and variables is linear. Though there are studies describing nonlinear versions of PLS,[12−14] the type of nonlinear dependence retrieved by this method is often limited to squared and cross-terms of parameters and thus may not always be adequate to find proper relationships between structure and activity of the analyzed molecules.

Artificial neural networks (ANNs) is one group of methods that are increasingly being used in drug design to study QSAR.[15−17] This method is able to elucidate structure−activity relationships and take into account any nonlinear character of these relationships. Thus, this method can be of significant interest in 3D QSAR studies. There are basically two different types of neural networks.

The first type, an unsupervised neural network, realizes training without the teacher. This means that the target values are not known or absent and neural network learning consists of the detection and clustering of input samples according to internal relationships among them. Principal methods in this field were devised by Hopfield,[18] Kohonen,[19,20] and Fukushima.[21,22] As a rule, the unsupervised neural algorithms are represented by large networks of hundreds of neurons. They are used to solve a great variety of problems from models of associative memory[23] to comparison of geometric and electronic properties of molecular surfaces[24] and the selection of training and test sets.[20] Applications of this method in QSAR studies were recently reviewed in ref 25. These methods are well adapted for the analysis of a large number of input parameters.

The second type, a supervised neural network (SNN), is used to establish relationships between input and target variables and represents training with a teacher. The feed-forward neural networks trained with the back-propagation algorithm are the most widely used methods in this field.[26,27] In the case where the dependencies between analyzed descriptors and molecular parameters are nonlinear, the SNN can produce more accurate models than linear regression methods. This can be very important for practical applications and the design of new compounds. These networks are usually used for the analysis of several tens of molecular descriptors.[28−30] The application of SNNs to a data set with a large number of input parameters, such as the thousands of input data generated by the CoMFA method, is complicated. First, the speed of a neural network is very low when dealing with a large number of input parameters. Second, SNN can have a low generalization ability due to the overtraining problem,[30] which becomes more crucial as the number of inputs

* Address for correspondence: Igor V. Tetko, Institut de Physiologie, Université de Lausanne, Rue du Bugnon 7, CH-1005 Lausanne, Switzerland. Tel: ++41-21-692.5534. Fax: ++41-21-692.5505. E-mail: itetko@eliot.unil.ch.
† Institute of Bioorganic & Petroleum Chemistry.
‡ Université de Lausanne.
§ ChemQuest and University of Portsmouth.

10.1021/jm010858e CCC: $20.00
Published on Web 06/15/2001

increases. It is also important to mention a specific problem when applying SNNs to CoMFA parameters. Since all input variables are used as independent parameters for neural network training, there is a substantial loss of information about the 3D structure of molecules. In addition, the presence of correlation effects between input variables, which is the usual case for CoMFA parameters, can have a negative impact on the generalization of SNNs.

The objective of this study is to develop a combination of unsupervised and supervised neural networks, the volume learning algorithm (VLA), that combines the best features of both networks to provide an easy interpretation of the steric and electrostatic requirements of a set of analyzed molecules at the receptor site. Since the PLS algorithm is effectively the exclusive analytical method used in CoMFA studies, a new alternative approach can give a different insight into the analyzed problem and thus can be used to complement the results obtained by the traditional techniques.

The proposed procedure (VLA) defines cluster zones in space around molecules using the self-organizing map of Kohonen (SOM)[19] and then uses the mean values of these clusters for the training of a feed-forward back-propagation neural network. This approach decreases the number of input parameters required for neural network training by several orders, preserves the spatial structural information of molecules, and calculates neural network models with high generalization ability.

Data Sets

The efficiency of the new approach was studied using a series of cannabimimetic aminoalkyl indoles (AAI).[31] The AAIs, developed from the lead compound pravadoline, represent a novel class of cannabinoid receptor agonists capable of binding with the cannabinoid CB_1 receptor. The structures of the analyzed compounds and their activities are listed in Table 1. The activity of the compounds was taken as their ability to bind with the receptor, reported as a constant pK_i. This data set was recently carefully analyzed in ref 31 using the traditional CoMFA approach. It was known from the literature[32] that pravadoline and its analogues have pK_a's of 4.5∼6.0 for the morpholine nitrogen of AAI. Thus, this atom would exist largely in the unprotonated form at physiological pH. However, it is also possible that there exists within the receptor binding site a specific residue capable of protonating the morpholine nitrogen or, indeed, that the pH of the microenvironment of the binding site could be such that this nitrogen would be protonated. That is why the authors[31] considered for their analysis two data sets. The first set consisted of molecules in which the morpholine nitrogen was unprotonated, whereas set 2 contained the corresponding protonated form. The 64 AAI derivatives were used in the training set and 6 compounds in the test data set (Table 1). Both data sets were analyzed using the standard CoMFA procedure. Detailed information on molecular modeling and structural alignment can be found in ref 31. In the current work we did not perform any such modeling but just used the CoMFA field values from the previous study provided to us by Prof. W. Welsh and Dr. J.-Y. Shim. Use of the same input values for PLS and neural network training made possible a straightforward comparison of the results of both methods.

In brief, each molecule with a fixed spatial pattern was placed inside a 3D cubic lattice divided into a grid with 2 Å spacing. Compounds **13** and **46** were the most potent ligands in the studied series and thus are expected to contain the key structural features of the highly potent AAI compounds. The only difference between them was the presence of a −4OMe group at the naphthyl moiety of compound **46**. Either of these compounds might have been used as a template molecule for alignment, but compound **46** was selected since it was also used as a template in the original work of Shim et al.[31] All compounds were aligned onto the template molecule by root-mean-square fitting onto three common atoms. The atoms were selected as follows: for the unprotonated model, the indole N_1 atom, the C_α or H_α atom of the C_2, and the O atom of the C_3 carbonyl; for the protonated model, the C_α atom of the N_1 side chain, the indole C_3 atom, and the C_1' atom on the C_3 carbonyl. This alignment scheme was selected by Shim et al.,[31] since it provided both reasonably good overlap of the putative biologically relevant pharmacophore elements and statistically significant 3D QSAR models from CoMFA.

The steric (van der Waals) and electrostatic (Coulombic) energies were calculated from the interaction between all the atoms of each molecule and a probe atom, represented by an sp^3 carbon with a +1 charge, in each unit of the grid.[31] This analysis calculated two data sets with $P = 1430$ and $P = 1716$ parameters for the protonated and unprotonated sets, respectively. The field values were truncated, as is usual in CoMFA studies, at 30 kcal/mol.

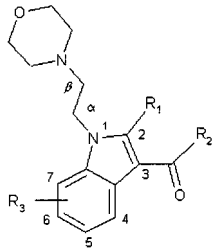
Supervised Artificial Neural Networks

The feed-forward neural network trained with the back-propagation algorithm was used as the SNN.[33] The neural networks had five neurons in one hidden layer selected as indicated in the Results and Discussion section. The bias neuron was presented on the input and on the hidden layer.

The input parameters of the SNN were mean values (simple average) of CoMFA field parameters calculated for detected clusters of grid points. The clustering procedure was performed using SOM, as described in the next section. The objective of SNN learning was to correlate the input parameters and the analyzed activities of molecules and to build a model with high generalization ability.

Avoidance of overfitting/overtraining has been shown to be an important factor for the improvement of generalization ability in neural network studies.[30,34] The early stopping over ensemble technique was used in the current study to overcome this problem. A detailed description of this approach can be found elsewhere.[30,34] In brief, each analyzed artificial neural network ensemble was composed of $M = 100$ networks. The values calculated for analyzed cases were averaged over all M neural networks, and their means were used to compute statistical coefficients with targets. A subdivision of the initial training set into two equal learning/validation subsets was used. The first set was used to train the

Table 1. Structures and pK_i Values of Cannabimimetic Aminoalkyl Indoles



					pK_i (calculated)			
					unprotonated		protonated	
compd	R_1	R_2	R_3	pK_i (obsd)	PLS[a]	VLA[b]	PLS	VLA
				Test Set				
pravadoline	Me	*p*-OMe-phenyl	H	−3.40	−3.26	−2.90	−3.22	−2.90
2	H	*p*-OMe-phenyl	H	−2.49	−2.02	−2.36	−2.01	−1.92
4	H	7-benzofuryl	H	−1.15	−1.51	−1.56	−1.18	−1.77
9	H	1-naphthyl	H	−0.78	−1.54	−1.44	−1.50	−1.54
11	Me	1-naphthyl	H	−1.18	−1.70	−1.76	−1.88	−1.56
12[c]	H	1-naphthyl	H	−0.38	−0.48	−0.87	−0.36	−0.42
				Training Set				
8[d]	H	1-naphthyl	H	−1.61	−1.12	−1.73	−0.22	−1.12
13[e]	H	1-naphthyl	H	0.02	−1.13	−0.77	−0.82	−0.52
15	Me	*o*-OMe-phenyl	H	−2.80	−2.99	−2.64	−2.46	−2.49
16	Me	*p*-OMe-phenyl	6-Me	−3.17	−2.58	−2.52	−2.63	−2.89
17	Me	*p*-OMe-phenyl	7-Me	−3.08	−3.04	−2.84	−3.19	−2.95
18	Me	*p*-OMe-phenyl	6-OMe	−3.06	−3.02	−2.54	−2.71	−2.96
19	Me	*p*-OMe-phenyl	7-F	−2.65	−2.89	−2.86	−2.56	−2.99
20	Me	*p*-OMe-phenyl	6-Br	−2.61	−2.50	−2.75	−2.51	−2.94
21[f]	Me	*p*-OMe-phenyl	H	−2.69	−3.01	−2.93	−2.97	−3.09
22	Me	*m*-Me-phenyl	H	−2.68	−2.88	−2.73	−2.82	−2.77
23	Me	*p*-Me-phenyl	H	−3.15	−2.40	−2.83	−3.00	−2.61
24	Me	*p*-Cl-phenyl	H	−3.02	−2.72	−3.30	−3.04	−2.89
25	Me	*p*-Et-phenyl	H	−2.38	−2.21	−2.59	−2.81	−2.11
26	Me	*p*-*i*-Pr-phenyl	H	−2.95	−3.00	−2.70	−3.17	−2.66
27	Me	*o*,*m*-di-Me-phenyl	H	−2.47	−2.51	−2.10	−2.75	−3.01
28	Me	*m*,*p*-di-Me-phenyl	H	−2.04	−2.22	−2.47	−1.89	−2.10
29	Me	1-naphthyl	6-Me	−0.94	−1.04	−0.62	−1.13	−1.58
30	Me	1-naphthyl	7-OMe	−0.90	−0.95	−1.46	−0.39	−1.16
31	Me	1-naphthyl	6-Br	−0.90	−1.00	−0.81	−1.15	−1.62
32[g]	Me	1-naphthyl	H	−1.47	−1.43	−1.68	−1.61	−1.69
33(*R*)[h]	Me	1-naphthyl	H	−1.01[n]	−1.01	−1.50	−0.85	−0.69
33(*S*)[h]	Me	1-naphthyl	H	−1.01[n]	−1.06	−1.76	−1.14	−1.49
34[i]	Me	1-naphthyl	H	−1.67	−1.74	−1.56	−2.04	−1.86
35[j]	Me	1-naphthyl	H	−2.48	−2.22	−1.87	−2.13	−1.69
36	H	1-naphthyl	6-Me	−0.44	−0.89	−0.51	−0.53	−1.55
37	H	1-naphthyl	5-F	−1.44	−1.64	−1.48	−1.50	−1.56
38	H	1-naphthyl	5-Br	−2.17	−1.74	−2.05	−1.55	−2.11
39	H	1-naphthyl	5-OH	−1.74	−1.85	−1.76	−1.55	−1.76
(+)-40[h]	H	1-naphthyl	H	−1.95	−1.57	−1.34	−1.92	−1.40
(−)-41[h]	H	1-naphthyl	H	−1.42	−1.32	−1.71	−1.18	−1.35
42(*R*)[k]	H	1-naphthyl	H	−1.99[n]	−1.74	−1.75	−1.99	−1.66
42(*S*)[k]	H	1-naphthyl	H	−1.99[n]	−1.84	−1.61	−1.90	−1.60
43	Cl	1-naphthyl	H	−0.90	−1.36	−1.40	−1.26	−0.86
44	H	4-Me-1-naphthyl	H	−0.35	−0.33	−0.38	−0.43	−0.14
45	Me	4-Me-1-naphthyl	H	−0.67	−0.63	−0.75	−0.97	−0.57
46	H	4-OMe-1-naphthyl	H	−0.04	−0.18	−0.38	0.60	−1.08
47	Me	4-OMe-1-naphthyl	H	−1.64	−1.64	−1.02	−1.15	−1.02
48	H	4-OH-1-naphthyl	H	−0.43	−0.53	−0.73	−0.62	−0.76
49	Me	4-CN-1-naphthyl	H	−1.07	−1.05	−1.24	−1.03	−0.97
50	Me	4-Br-1-naphthyl	H	−0.74	−0.92	−1.11	−1.04	−0.90
51	Me	2-naphthyl	H	−2.01	−2.28	−2.32	−2.11	−2.01
52	H	2-naphthyl	H	−1.67	−1.54	−1.29	−1.48	−1.44
53	Me	2-quinolinyl	H	−2.42	−2.15	−2.28	−2.06	−2.16
54	Me	4-quinolinyl	H	−2.35	−2.25	−1.81	−1.97	−1.62
55	Me	5-quinolinyl	H	−2.82	−2.38	−1.91	−2.24	−1.91
56	Me	6-quinolinyl	H	−2.60	−2.50	−2.29	−2.75	−2.49
57	H	6-quinolinyl	H	−2.59	−2.55	−2.41	−2.49	−2.35
58	Me	7-quinolinyl	H	−2.15	−2.26	−2.33	−2.63	−2.32
59	Me	8-quinolinyl	H	−2.34	−1.94	−2.05	−1.97	−2.15
60	Me	2-benzofuryl	H	−2.52	−2.44	−2.74	−2.36	−2.54
61	Me	3-benzofuryl	H	−1.94	−1.81	−2.19	−2.08	−2.62
62	H	4-benzofuryl	H	−1.28	−1.54	−1.79	−1.37	−1.45
63	Me	5-benzofuryl	H	−2.45	−2.58	−2.65	−2.80	−2.54

Table 1 (Continued)

compd	R_1	R_2	R_3	pK_i (obsd)	pK_i (calculated) unprotonated PLS[a]	unprotonated VLA[b]	protonated PLS	protonated VLA
			Training Set					
64	H	5-benzofuryl	H	−2.04	−2.69	−2.59	−2.31	−2.57
65	H	6-benzofuryl	H	−2.38	−2.34	−1.73	−2.13	−1.92
66	H	7-benzofuryl	H	−1.75	−2.47	−2.42	−2.02	−2.37
67	H	9-anthracenyl	H	−3.37	−3.32	−1.88	−3.43	−2.91
68	H	1,2,3,4-H_4-naphthyl	H	−1.48	−1.42	−1.95	−1.51	−1.29
69[e]	Me	*p*-OMe-phenyl	H	−2.74	−2.93	−2.29	−2.79	−2.66
70(*R*)[l]	Me	*p*-OMe-phenyl	H	−2.59	−2.64	−2.35	−2.50	−2.47
71[e]	H	*p*-OMe-phenyl	H	−0.88	−0.79	−1.61	−0.76	−1.19
72[l]	Me	1-naphthyl	H	−0.71	−0.60	−0.66	−0.80	−0.25
73[e]	Me	1-naphthyl	H	−0.63	−0.52	−0.91	−0.58	−0.66
74[m]	H	1-naphthyl	H	−0.72	−0.64	−0.54	−0.80	−0.53

[a] PLS = partial least squares results from ref 31. [b] VLA = volume learning algorithm results calculated using a combination of steric and electrostatic fields. [c] [3-(4-Me-Morpholinyl)]methyl instead of 2-(4-morpholinyl)ethyl for the N1 side chain. [d] [2-(1,4-Di-Me-piperazinyl)]methyl instead of 2-(4-morpholinyl)ethyl for the N1 side chain. [e] [2-(1-Me-Piperidinyl)]methyl instead of 2-(4-morpholinyl)ethyl for the N1 side chain. [f] 2-[4-(1-Thiomorpholinyl)]ethyl instead of 2-(4-morpholinyl)ethyl for the N1 side chain. [g] 2-(1-Piperidinyl)ethyl instead of 2-(4-morpholinyl)ethyl for the N1 side chain. [h] 1-Me-2-(4-Morpholinyl)ethyl instead of 2-(4-morpholinyl)ethyl for the N1 side chain. [i] 2-[4-(2-Me-Morpholinyl)]ethyl instead of 2-(4-morpholinyl)ethyl for the N1 side chain. [j] 2-[4-(3-Me-Morpholinyl)]ethyl instead of 2-(4-morpholinyl)ethyl for the N1 side chain. [k] 2-Me-2-(4-Morpholinyl)ethyl instead of 2-(4-morpholinyl)ethyl for the N1 side chain. [l] [2-(1-Me-Pyrrolidinyl)]methyl instead of 2-(4-morpholinyl)ethyl for the N1 side chain. [m] (2-Piperidinyl)methyl instead of 2-(4-morpholinyl)ethyl for the N1 side chain. [n] Reported as racemate, but separately treated in CoMFA models.

neural network while the second one was used to monitor the training process as measured by root-mean-square error (RMS). An early stopping point determined as a best fit of a network to the validation set was used to terminate the neural network learning. It has been shown that a neural network trained up to an early stopping point provides better prediction ability than a network trained to the error minimum for the learning set.[30] Thus, statistical parameters calculated at the early stopping point were used. The learning process limits $N_{all} = 10\,000$ iterations (total number) or $N_{stop} = 2000$ iterations (local number) following the last decrease of RMS for the validated set in the early stopping point.

It has been shown that pruning algorithms[35,36] may be used to optimize the number of input parameters for SNN learning and to select the most significant ones. These algorithms operate in a manner similar to stepwise multiple regression analysis and exclude on each step one input parameter that was estimated to be nonsignificant. The pruning algorithms were used in the current study to determine significant clusters of input data points of the analyzed molecules as described in refs 35 and 36.

Self-Organizing Map of Kohonen

The main role of a SOM network is to create a nonlinear projection of high-dimensional input samples $\mathbf{X_s}$ to a lower two-dimensional output SOM space that is represented by a two-dimensional array of neurons. The SOM training is performed in such a way that input vectors with similar properties in the high dimensional space are mapped to the same (or to the nearby) neurons on the two-dimensional space. Thus, by considering all input vectors projected to the same output neuron it is possible to determine clusters of vectors having similar properties in the high-dimensional space. Since all vectors from the same cluster are similar, only one vector need be used to represent the properties of all the vectors from this cluster. The use of one instead of a number of vectors provides a compression of the input data. Such compression was used in the current study to decrease the number of CoMFA field parameters. To do this, each grid point of the CoMFA map was considered as a vector in the space of the analyzed molecules. In a simple case, the dimension of this vector is equal to the number of molecules in the training set, and its coordinates correspond to the field values calculated at this grid point for different molecules. The vectors that have similar properties in this space were projected by the SOM algorithm to the same neuron and formed one cluster. Thus, this procedure subdivided the input space of grid points into a number of separate clusters, and training of the SNN was performed as described in the previous section.

It should be pointed out here that the clustering procedure was only used to provide a subdivision of the grid space into a number of disjoint clusters, i.e., to compress the input space of the parameters. After this subdivision, there was no further need of the SOM neural networks. The detected clusters can also be used to compress the input data of a new set of molecules.

Another important point is that different vector components can be used to perform the clustering of the input space. For example, it was possible to use only electrostatic, steric, or a combination of both these field parameters as components of the vectors $\mathbf{X_s}$. More sophisticated components, for example, the input to hidden weights of the SNN (i.e., the weights connecting a given grid point to the hidden layer neuron) could also be used.

Let us describe in more detail an algorithmic implementation and the parameters of the clustering procedure used in this study. The source code of the SOM algorithm was downloaded from an anonymous ftp server[37] and integrated into the data analysis program. In the current study, SOM neural networks with a hexagonal lattice type of neuron array were used. It has been suggested that this lattice type is more effective for visual display and analysis than a rectangular one.[19]

As was already mentioned, the purpose of the unsupervised analysis performed by the SOM network was

to find clusters of input parameters. The input data set used to train the SOM consisted of k input samples corresponding to the number of the field parameters (e.g., $k = 1430, 1716$). Each input sample consisted of m values corresponding to the field values of this grid point calculated for each analyzed molecule:

$$\mathbf{X}_s = (x_{s1}, x_{s2}, ..., x_{sm}) \quad s = 1, 2, ..., k \qquad (1)$$

Each SOM neuron j had a number of weights m corresponding to the number of molecules. These m weights formed an m-dimensional weight vector

$$\mathbf{W_j} = (w_{j1}, w_{j2}, ..., w_{jm}) \quad j = 1, 2, ..., n \qquad (2)$$

where n was the number of neurons in the Kohonen network. At the beginning of network training, the neuron weights were initialized using random values.

Since the algorithm itself has thoroughly been described elsewhere,[19,20,37] the process of SOM learning is only briefly discussed here. During training, each input vector, $\mathbf{X}$, was compared to all of the $\mathbf{W_j}$, obtaining the location of the closest match $\mathbf{W_c}$ (given by $||\mathbf{X} - \mathbf{W_c}|| = \min_j \{||\mathbf{X} - \mathbf{W_j}||\}$ where $||\mathbf{a}||$ denotes the norm of vector $\mathbf{a}$). In the learning process, the weights $\mathbf{W_c}$ of the matching (winning) neuron were adjusted in such a way that they became even more similar to the input vector $\mathbf{X}$. The weights of all the other neurons in the network were also adjusted but to an amount that decreased with increasing topological distance from the winning neuron. The amount of adjustment was determined by the learning rate, α, and the neighborhood radius, σ, which were monotonically decreasing with the time of SOM training. Such training produced a self-organization of the neuron weights. All input samples were represented to the SOM network thousands of times, and the neuron weights were gradually adjusted to approximate the input vector as well as possible. The quality of a SOM was estimated by a quantization error $\boldsymbol{E_c}$, that was calculated at the end of training as a mean value of $||\mathbf{X} - \mathbf{W_c}||$ over all $\mathbf{X}$ vectors.

Training a Kohonen map consisted of two phases.[18] The first phase, of 100 000 iterations, was used to roughly order the weight vectors of the map neurons. During the second phase, of 40 000 iterations, the values of the weight vectors were fine-tuned.[19,37] The initial learning rate and neighborhood radius of the SOM were selected to be $\alpha_1 = 0.6$, $\sigma_1 = 2/5(xy)^{0.5}$ and $\alpha_2 = 0.15$, $\sigma_2 = 2/5\sigma_1$ for the first and the second phase, respectively, where x and y correspond to the size of the SOM map. These parameters were selected based on the guidelines and examples of the SOM_PAK manual.[37] It was found that decreasing the number of iterations increased the quantization error $\boldsymbol{E_c}$, while no improvement of the mapping was found with an increase in the number of iterations. Variation of the learning rates, α, and the neighborhood radii, σ, for about 50% of their magnitudes did not influence the quantization error $\boldsymbol{E_c}$.

After termination of the learning process, all data samples were analyzed by the SOM network one by one, and the winning neuron was determined for each input vector using the same method as described earlier. The result of the network self-organizing analysis was a map that projected k input samples on the two-dimensional array of SOM neurons.

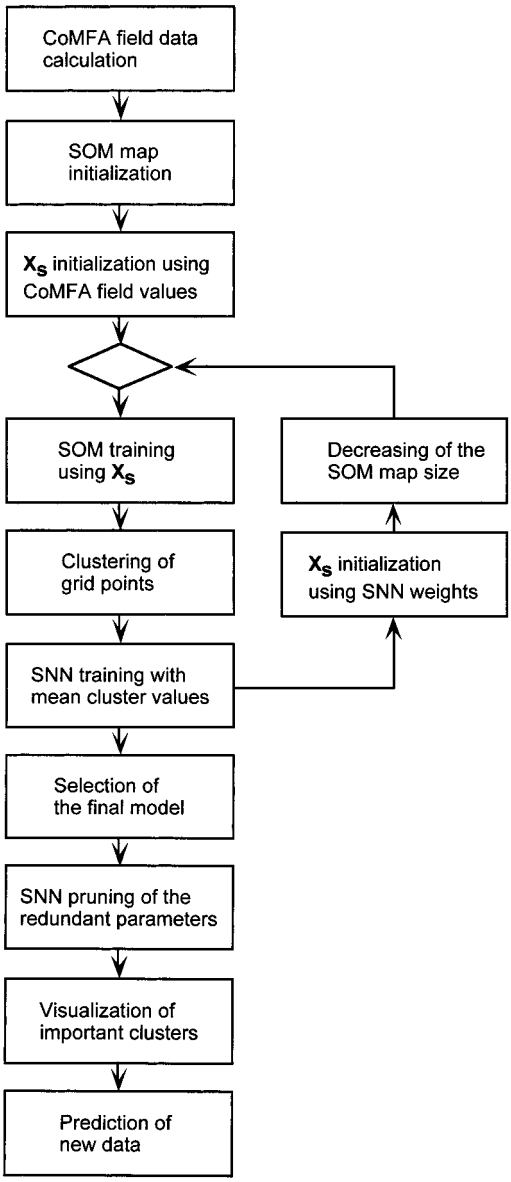


Figure 1. Flowchart of the volume learning algorithm.

Volume Learning Algorithm Artificial Neural Network

The volume learning algorithm (VLA) realized a recurring iterative application of supervised (SNN) and unsupervised (SOM) algorithms (Figure 1). The general idea of the algorithm was to use SOM to perform partitioning of the input data parameters in clusters and then to use the mean values of these clusters for SNN training. The optimal size of the map was determined during the work of VLA according to the minimum RMS error calculated for the validation set at the early stopping point. The initial size of the Kohonen map in the SOM algorithm was $(x, y) = (14, 12)$ or 168 nodes, and it was selected by a rule of thumb. Further studies have determined that the optimal sizes of maps that gave the best final predictions were, in fact, in the range from $(x, y) = (8, 6)$ to $(x, y) = (4,2)$. Thus, the initial size was sufficient to select the optimal maps.

The algorithm included a number of iterations that were used to find an optimal partitioning of the input CoMFA parameters on the number of clusters.

(1) The first step consisted of preparation of an initial learning data set for the SOM algorithm training. The number of data samples corresponded to the number of input variables calculated by CoMFA. Thus, there were $P = 1430$ or $P = 1716$ input data samples for the protonated and unprotonated data sets, respectively. During the very first iteration of VLA, the components of the vectors used in the learning set of SOM contained the CoMFA parameters of the analyzed field values. The dimension of each data sample $d_i = 64$ corresponded to the number of molecules in the training set. However, during the next steps of VLA learning, the CoMFA parameters were substituted by the weights of trained SNN networks. In this case each data sample contained input-to-hidden neurons weights corresponding to this input parameter in the ensemble of supervised neural networks. Notice that each input neuron was connected to the H hidden neurons and only these weights were considered. Since there were $M = 100$ of SNNs in the ensemble, the dimension of each data sample was $d_i = MH$. It was found that the use of the SNN weights provided improved results compared to the analysis based on the partitioning using the CoMFA parameters.

(2) The second step included initialization and training of the Kohonen network of size (x, y). The initial size of this map was (x, y) = (14, 12). This size was decreased with iterations of the algorithm as indicated at step 4. The training by the SOM algorithm partitioned the initial CoMFA parameters into a number of clusters. All small clusters that contained less than a minimal number of points, $n_{\min}$, were identified, and neighbor clusters with a minimal distance $||\mathbf{X_i}, \mathbf{X_j}||$ were joined together. The number $n_{\min} = 8$ was selected after a few trials as discussed in the Results and Discussion section.

(3) The clusters of input CoMFA data points calculated by the SOM were analyzed. The mean values of input parameters from each cluster formed a new data set for the SNN learning. Thus, the number of input variables of the SNN corresponded to the number of clusters determined by the SOM algorithm. An ensemble of 100 neural networks was trained. The optimized neural network weights that calculated a minimum validated RMS error at the early stopping point were saved and then were used for SOM neural network training at step 2.

(4) The size of the Kohonen network was decreased from (x, y) to ($x - 1$, $y - 1$) for successive iterations of the algorithm. The steps 1−4 were repeated until the size of the Kohonen map did not decrease up to (4, 2). The partitioning that provided minimum RMS was selected as the optimal one.

(5) The last step included application of the pruning algorithms that were used to select a set of the relevant input parameters. The optimized clusters of parameters were used to visualize the regions of molecular parameters that were found to be important for the analyzed activity of molecules.

In the case of the combined fields, the electrostatic and steric fields were analyzed separately using the SOM algorithm. The detected clusters formed a joint partition that was used to train the neural networks.

Results and Discussion

A preliminary study was performed to determine an optimum number of hidden neurons in the SNNs. This study used a Kohonen SOM of the maximal size. The input data for the SOM were represented by steric parameters of the unprotonated set. The number of hidden neurons was investigated by the examination of neural networks with 2, 3, 4, 5, 7, and 10 hidden neurons. It was found that the RMS error calculated for the validation set was decreasing (i.e., RMS = 0.65 ± 0.02, 0.64 ± 0.03, 0.64 ± 0.04, 0.63 ± 0.02) when the number of neurons in the hidden layer was changed from 2 to 5. However, further increase in the number of hidden neurons from 5 to 7 and 10 did not improve the prediction ability of the neural networks (i.e., RMS = 0.66 ± 0.03, 0.66 ± 0.03). Thus the number of neurons in the hidden layer was selected to be 5. This number also provided the maximum speed of SNN calculations.

Two types of analysis were performed. In the first type, the clustering for all iterations of VLA was performed using only CoMFA parameters for Kohonen network training. In the second type of analysis, the neural network weights from the SNN were used as inputs for the Kohonen neural network starting after the second iteration of the algorithm. Such analysis was performed for steric, electrostatic, and a combination of steric and electrostatic fields for both data sets (Table 2).

The use of the neural network weights as input for the Kohonen neural network significantly improved performance of the VLA. Indeed, there were 12 cross-validated coefficients calculated by VLA for the analyzed fields (including joint set), two data sets (i.e., leave-one-out results for training and prediction results for the test set), and two models (protonated and unprotonated). In 10 out of 12 cases, the performance of VLA trained with neural network weights, as input for the Kohonen network, was superior to that of VLA trained using the initial CoMFA parameters. Thus, according to the nonparametric sign criterion[38] with $p < 0.01$, VLA predictions using neural network weights as input for training of the Kohonen neural network were improved in comparison to the same analysis using only the initial CoMFA parameters.

The difference in the performance of both these approaches was due to an improvement of SOM mapping when using neural network weights compared to use of the initial CoMFA parameters. The optimal SOM maps calculated with neural weights usually contained a smaller number of clusters compared to those calculated using the original CoMFA parameters. This fact clearly indicated that a clustering of initial CoMFA parameters was a much more complex task for the SOM algorithm (and since a large number of different clusters was found) than a similar analysis of SNN weights. The initial parameters contained a lot of information that was not important or relevant to the analyzed activity of the molecules. By contrast, the weights of a trained neural network contained information that was more relevant to the analyzed activity. Indeed, the magnitudes of these weights were determined during neural network learning to produce a maximum fit between calculated and observed activities. That is why the use of SNN weights provided a better clustering by the SOM

Table 2. Cross-Validated q^2 Values Calculated Using the Volume Learning and PLS Algorithms for Different Fields[a]

	volume learning algorithm						PLS[b]	
	clustering using initial field params		clustering using ANN weights		pruning results			
field	clusters	cross-validated q^2	clusters	cross-validated q^2	clusters	cross-validated q^2	latent variables	cross-validated q^2
Aminoalkyl Indoles (Unprotonated)								
steric	28	0.47 ± 0.02 (0.63 ± 0.03)	14	0.78 ± 0.01 (0.72 ± 0.02)	10	0.78 ± 0.01 (0.71 ± 0.02)	5	0.53 (0.76)
electrostatic	16	0.28 ± 0.03 (0.70 ± 0.02)	8	0.43 ± 0.02 (0.78 ± 0.02)	4	0.49 ± 0.02 (0.79 ± 0.03)	4	0.31 (0.56)
steric + electrostatic	83	0.39 ± 0.04 (0.78 ± 0.02)	16	0.75 ± 0.02 (0.76 ± 0.02)	10	0.76 ± 0.03 (0.74 ± 0.03)	6	0.56 (0.83)
Aminoalkyl Indoles (Protonated)								
steric	10	0.54 ± 0.03 (0.49 ± 0.03)	15	0.72 ± 0.02 (0.81 ± 0.01)	6	0.73 ± 0.02 (0.81 ± 0.02)	5	0.59 (0.75)
electrostatic	14	0.48 ± 0.01 (0.86 ± 0.02)	8	0.61 ± 0.02 (0.75 ± 0.02)	5	0.63 ± 0.02 (0.76 ± 0.03)	4	0.45 (0.67)
steric + electrostatic	40	0.52 ± 0.02 (0.66 ± 0.03)	16	0.77 ± 0.01 (0.74 ± 0.02)	6	0.77 ± 0.01 (0.74 ± 0.02)	6	0.59 (0.78)

[a] PLS = partial least squares; ANN = artificial neural network. The leave-one-out results for the training set and, in parentheses, for the test set are shown. [b] Results updated from ref 31.

Table 3. Cross-Validated q^2 for Different Minimal Numbers of Points n_{min} per Cluster Using Steric Unprotonated Data Set

n_{min}	calculated clusters	LOO[a]	test set
1	24	0.63 ± 0.03	0.70 ± 0.03
3	15	0.77 ± 0.01	0.73 ± 0.02
4	24	0.71 ± 0.02	0.73 ± 0.04
5	28	0.71 ± 0.03	0.57 ± 0.03
7	19	0.74 ± 0.01	0.74 ± 0.03
8	14	0.78 ± 0.01	0.72 ± 0.02
10	7	0.84 ± 0.02	0.67 ± 0.02
15	5	0.81 ± 0.02	0.66 ± 0.02
20	5	0.73 ± 0.02	0.56 ± 0.04
30	6	0.71 ± 0.02	0.48 ± 0.04

[a] LOO = leave-one-out results for the training set.

algorithm and, consequently, a better performance of VLA. The idea to cluster neural network weights was inspired by a similar method used to cluster input samples in the efficient partition algorithm[34] and by the data partitioning algorithm of ref 20.

A minimum number of points, n_{min}, per cluster was also an important parameter for VLA. A detailed study of this parameter was carried out for the steric fields of the unprotonated set as shown in Table 3. The optimal performance of VLA was detected for $n_{min} = 7-15$. A smaller value of this parameter tended to produce a large number of clusters with low prediction ability. This effect was probably due to the overfitting of neural networks trained with a larger number of input parameters. An increase of this value beyond the optimal range also decreased performance of the algorithm. However, in this case the decrease was due to the low spatial resolution of data that was lost in the clustering. Thus, this number was selected to be $n_{min} = 8$, and it was used for all calculations reported in this article.

Interesting results were provided by an analysis of the distribution of parameters in clusters calculated by the SOM algorithm (Table 4). Most input parameters were collected in one large cluster while the number of parameters in the other clusters was in a range from 8 to 73 with a mean of 20 and a median of 14 parameters, respectively. The largest cluster predominantly contained input parameters for the 3D space occupied by the molecules and, correspondingly, with truncated field values. From a chemical point of view such clusters were the least informative.

The importance of the detected clusters for the observed activity was evaluated using pruning methods during the last step of the algorithm. The pruning eliminated a number of clusters detected by the VLA procedure. Clusters with the largest number of parameters were detected as nonsignificant and were eliminated by the pruning algorithms.

For the training set, steric parameters provided higher statistical coefficients compared to the electrostatic parameters for both unprotonated and protonated sets. The importance of steric interactions for the cannabinoid activity of the analyzed molecules was also identified by analysis of the joint set containing both steric and electrostatic fields. Only 3 out of 10 clusters selected by the pruning methods for the unprotonated model contained electrostatic parameters while no significant clusters with electrostatic parameters were found for the protonated model. However, even if the clusters with electrostatic parameters were found to be important for the activity of molecules in the unprotonated model, their addition did not change its prediction ability. The cross-validated q^2 for steric and steric plus electrostatic fields were within confidence intervals for both test and training sets. This result indicated the importance of the steric field for the analyzed activity.

A similar conclusion about these fields was found in the PLS studies.[31] Indeed, the individual contributions from steric/electrostatic fields were 71%/29% and 80%/20%, respectively, for the unprotonated and protonated models.[31] In fact, PLS results calculated for steric and steric plus electrostatic fields were the same for the protonated model while only a small improvement of results was observed for the unprotonated model. Thus, the results of both VLA and PLS methods indicated that the variation in binding affinity among the AAI is dominated by steric interactions at the receptor site. This is in accordance with the known importance of the hydrophobic components of the AAI[39] and the classical cannabinoids[40] for the cannabimimetic activity.

Prediction ability of all methods for the test set of six molecules significantly varied, and it tended to be higher

Table 4. Distribution of Parameters in Clusters Detected by the Volume Learning Algorithm

	number of parameters in each cluster														
field	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
						Aminoalkyl Indoles (Unprotonated)									
steric	1398	73*	48*	40*	31*	23	21*	15	14*	12*	11*	11*	10	9*	
electrostatic	1596	31*	22*	20*	13	12*	12	10							
steric + electrostatic	1566	58*	39*	16*	11*	10*	8*	8*							
	1635	17*	17	10*	10	9	9*	9							
						Aminoalkyl Indoles (Protonated)									
steric	1077	70*	63*	35*	32	25*	21*	19	15	15	13	12	12	11*	10
electrostatic	1321	22*	20*	18*	18*	11	10	10*							
steric + electrostatic	1250	71*	38*	22*	15	14*	11*	9*							
	1349	15	13	12	12	11	10	8							

* Clusters selected by pruning methods. Total numbers of input parameters were $P = 1430$ and $P = 1716$ for protonated and unprotonated sets, respectively.

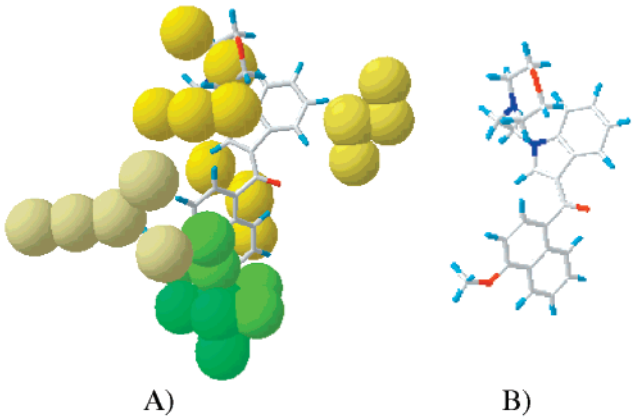


Figure 2. Steric (A) contour plots for the protonated model calculated by the pruning methods. Compound **46** (B) is used as a reference molecule. Enhanced binding is associated with increasing (green) and decreasing (yellow) steric bulk.

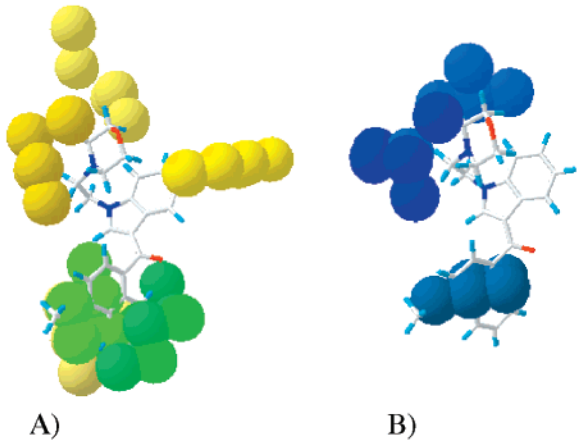


Figure 3. Steric (A) and electrostatic (B) contour plots for the unprotonated model calculated by the pruning methods. Compound **46** is used as a reference molecule. Enhanced binding is associated with increasing (green) and decreasing (yellow) steric bulk and with decreasing (blue) negative charges.

than that for the training set. This result is probably due to a smaller structural diversity of this set compared to the training set.

Visualization of the calculated results represented an important problem in assessment of the performance of VLA and comparison with PLS. Since the original *xyz*-coordinates were not directly used to cluster the grid points in the SOM algorithm, there were some disruptions in the neighborhood relations amid points in 3D space, i.e., some detected clusters contained points separated in space.

This happened mainly for clusters that were detected as nonsignificant and were pruned by ANNs. On the contrary, the clusters found to be significant as a rule contained grid points (shown as balls) from a compact region. The visualization procedure was performed as follows. At first, the geometrical center of each cluster was found by calculating the average value of the *x*-, *y*-, and *z*-coordinates. The nearest five or six grid points to the geometrical center of each cluster were shown as balls. If several points could be shown, preference was given to the point connected to the already shown points. Such a procedure reasonably mapped the position of small clusters (composed of 10−15 grid points) but was not optimal for the large ones.

The plots for the protonated (Figure 2) and unprotonated data sets (Figure 3) depicted regions around the molecules where enhanced CB_1 cannabinoid receptor binding affinity was associated with increasing (green) and decreasing (yellow) steric bulk and with decreasing (blue) negative charge.

The calculated results were consistent with known QSAR for the aminoalkyl indoles. For example, the green color around the naphthyl moiety indicated that the presence of bulky substituents at the R2 position increased binding inside a large pocket of the receptor. This is the case when, for example, the naphthyl moiety (compounds **45** and **47**) replaced the substituted phenyl moiety (pravadoline and compound **23**, respectively). The shape of the pocket was similar to that of 1-naphthyl, since there was a decrease in the pK_i if the 1-naphthyl group (compounds **9**, **11**) was changed to a 2-naphthyl (compounds **52**, **51**), quinolyl (compounds **53**−**59**), benzofuryl (compounds **4**, **60**−**66**), or 9-anthracenyl (compound **67**). Two different clusters with light- and dark-green colors (the light color corresponds to larger increase in energy) contribute to the green region around the naphthyl moiety. This suggests that, while the increase of steric bulk is in general favorable, the position of substituents was also important for the activity of molecules.

The electrostatic interactions also influenced the activity of the molecules for the unprotonated model. A blue region near C4′ of the naphthyl moiety (Figure 3B) was consistent with a similar region found by Shim et al. (see Figure 7A of ref 31). It indicates that enhanced CB1 cannabinoid receptor binding affinity could be associated with the presence of a methoxy (compound **46**) or a bromo (compound **50**) substituent at this

position.[31] However, the region detected by VLA is much wider and also includes a part of space around the naphthyl moiety. It is interesting that PLS (Figure 7A of ref 31) detected the importance of electrostatic interactions at this region only for the unprotonated, but not for the protonated, model. However, the steric and not electrostatic interactions, which were already described, were dominating at this region. For example, the presence of a Me subsitituent at the C4′ position, compound **44**, increased pK_i compared to the unsubstituted compound **9**. Compound **44** is more active than compound **48**, which has a hydroxy group at the C4′ position.

The presence of the yellow clusters closely surrounding the heterocyclic ring in the N1 side chain (see Figure 3A) indicate sterically forbidden regions. The presence of substituents at these regions significantly decreased activity of the molecules. This was the case for the methyl group of the piperazinyl ring (compound **8**) as well as for the methyl groups of compound **42(*R*)** and **42(*S*)** (methyl groups in the second position of the N1 side chain). Similarly, the methyl groups of compounds **34** and **35** at the second and third position of morpholine ring, respectively, also enter this region. The interactions with these sterically forbidden regions decreased the activity of these compounds compared to their analogue, compound **9**, that does not have such substituents. For the unprotonated model, this region also contained two blue clusters that indicated the importance of electrostatic interactions at this region. This emphasizes the importance of the basicity of this structural element for binding.

Similar sterically forbidden regions near the heterocyclic ring in the N1 side chain were detected by the PLS method for the protonated model, while only one small region consisting of a couple of grid points was detected for the unprotonated model (Figure 7 of ref 31). This indicated some instability of the calculated PLS models. On the contrary, the results calculated by VLA were quite similar for both protonated and unprotonated models. Both VLA and PLS algorithms calculated different orientations and sizes of contour maps for protonated and unprotonated models. This could be a result of different alignment schemes used for both models.

The presence of a sterically forbidden region near the C5 position explains a decrease of the activity of compounds that have a substituent at this position (compounds **37**−**39** compared to **9**). This region was not detected by PLS models.

In general, the places of localization of the identified regions in most cases were similar to the areas described in ref 31, though they had different form and size. Unfortunately, no definitive statement can be made regarding whether these compounds interact with the receptor in a protonated or unprotonated form, since both models calculated comparable statistical results. To some extent, the protonated model looks more attractive, since only a steric field was needed to explain the activity of the analyzed compounds, and thus the model itself is simple. However, such a conclusion can be biased, since the effect of the electrostatic field played only a minor part in the AAI ligand−receptor interactions for the unprotonated model too.

At the present stage of its development, the VLA can be applied after PLS analysis. The results calculated by both methods can be compared. The similar regions detected by both methods will provide an additional cross-validation of the results of the PLS method. On the contrary, discrepancies in the calculated maps should be analyzed to understand the limitations of each method. Further development of VLA will be aimed at an improvement of the visualization of the detected clusters.

Conclusion

This study presents a neural network system, VLA, for 3D QSAR studies. This algorithm clustered input parameters of the molecules and used their mean values to correlate the activity of analyzed molecules with their spatial and electronic structure. This significantly decreased the number of input parameters and made it possible to calculate 3D QSAR models with high prediction ability. An application of pruning algorithms estimated significance of the detected clusters and identified the most important regions of the input parameters determining the analyzed biological activity. The results calculated by the proposed approach were comparable to or superior to those of PLS. Moreover, both algorithms correctly determined the dominant role of steric interactions for the AAI cannabinoid receptor binding activity. The regions crucial for activity of the analyzed molecules detected by VLA were similar to those of PLS.

Acknowledgment. This study was partially supported by INTAS-Ukraine Grants 95-0060 and 96-1115. The authors thank William J. Welsh and Joong-Youn Shim (St. Louis, MO) for providing us the CoMFA data sets. We thank anonymous reviewers for their criticism and objective remarks that allowed us to improve the clarity of the article.

References

(1) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative molecular field analysis (COMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.

(2) Cramer, R. D.; DePriest, S. A.; Patterson, D. E.; Hecht, P. The developing practice of comparative molecular field analysis. In *3D QSAR in Drug Design*; Kubinyi, H., Ed.; Leiden: ESCOM, 1993; pp 443−485.

(3) Höltje, H.-D.; Anzali, S.; Dall N.; Höltje, M. Binding Site Models. In *3D QSAR in Drug Design*; Kubinyi, H., Ed.; Leiden: ESCOM, 1993; pp 320−335.

(4) Marshall, G. R. Binding-Site Modeling of Unknown Receptors. In *3D QSAR in Drug Design*; Kubinyi, H., Ed.; Leiden: ESCOM, 1993; pp 80−116.

(5) Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *28*, 849−857.

(6) Golbraikh, A.; Bernard, P.; Chretien, J. R. Validation of protein-based alignment in 3D quantitative structure−activity relationships with CoMFA models. *Eur. J. Med. Chem.* **2000**, *35*, 123−136.

(7) Palomer, A.; Pascual, J.; Cabre, F.; Garcia, M. L.; Mauleon, D. Derivation of pharmacophore and CoMFA models for leukotriene D(4) receptor antagonists of the quinolinyl(bridged)aryl series. *J. Med. Chem.* **2000**, *43*, 392−400.

(8) Zhang, S. X.; Feng, J.; Kuo, S. C.; Brossi, A.; Hamel, E.; Tropsha, A.; Lee, K. H. Antitumor agents. 199. Three-dimensional quantitative structure−activity relationship study of the colchicine binding site ligands using comparative molecular field analysis. *J. Med. Chem.* **2000,** *43*, 167−76.

(9) Wold, S.; Johansson, E.; Cocci, M. PLS−Partial Least Squares Projection to Latent Structures. In *3D QSAR in Drug Design*; Kubinyi, H., Ed.; Leiden: ESCOM, 1993; pp 523−563.

(10) So, S. S.; Karplus, M. Three-dimensional quantitative structure−activity relationships from molecular similarity matrixes and genetic neural networks. 1. Method and validations. *J. Med. Chem.* **1997**, *40*, 4347−4359.

(11) Hoffman, B.; Cho, S. J.; Zheng, W.; Wyrick, S.; Nichols, D. E.; Mailman, R. B.; Tropsha, A. Quantitative structure−activity relationship modeling of dopamine D(1) antagonists using comparative molecular field analysis, genetic algorithms-partial least-squares, and K nearest neighbor methods. *J. Med. Chem.* **1999**, *42*, 3217−3226.
(12) Berglund, A.; Wold, S. INLR, Implicit Non-Linear Latent Variable Regression. *J. Chemom.* **1997**, *11*, 141−156.
(13) Berglund, A.; Wold, S. A serial extension of multiblock PLS. *J. Chemom.* **1999**, *13*, 461−471.
(14) Eriksson, L.; Johansson, E.; Lindgren, F.; Wold, S. GIFI−PLS: Modeling of nonlinearities and discontinuities in QSAR. *Quant. Struct.-Act. Relat.* **2000**, *19*, 345−355.
(15) Kovesdi, I.; Dominguez-Rodriguez, M. F.; Orfi, L.; Naray-Szabo, G.; Varro, A.; Papp, J. G.; Matyus, P. Application of neural networks in structure−activity relationships. *Med. Res. Rev.* **1999**, *19*, 249−269.
(16) Jalali-Heravi, M.; Parastar, F. Use of artificial neural networks in a QSAR study of anti-HIV activity for a large group of HEPT derivatives. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 147−154.
(17) Manallack, D. T.; Livingstone, D. J. Neural networks in drug discovery: have they lived up to their promise? *Eur. J. Med. Chem.* **1999**, *34*, 195−208.
(18) Hopfield, J. J. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 2554−2558.
(19) Kohonen, T. *Self-organisation Maps*; Springer-Verlag: Berlin, 1995.
(20) Simon, V.; Gasteiger, J.; Zupan, J. A. Combined Application of Two Different Neural Network Types for the Prediction of Chemical Reactivity. *J. Am. Chem. Soc.* **1993**, *115*, 9148−9159.
(21) Fukushima, K. Cognitron: a self-organizing multilayer neural network. *Biol. Cybernetics* **1975**, *20*, 121−136.
(22) Fukushima, K. Neocognitron: a self-organizing multilayer neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybernetics* **1980**, *36*, 193−202.
(23) Hertz, J.; Krogh, A.; Palmer, R. G. *Introduction to the Theory of Neural Computation*; Addison-Wesley: Reading, MA, 1991.
(24) Anzali, S.; Barnickel, G.; Krug, M.; Sadowski, J.; Wagener, M.; Gasteiger, J.; Polanski, J. The comparison of geometric and electronic properties of molecular surfaces by neural networks: application to the analysis of corticosteroid-binding globulin activity of steroids. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 521−534.
(25) Anzali, S.; Gasteiger, J.; Holzgrabe, U.; Polanski, J.; Sadowski, J.; Teckentrup, A.; Wagener, M. The Use of Self-Organizing Neural Networks in Drug Design. *Perspect. Drug Discovery Des.* **1998**, *9−11*, 273−299.
(26) Zupan, J.; Gasteiger, J. *Neural Networks for Chemistry and Drug Design: An Introduction*, 2nd ed.; VCH: Weinheim, 1999.
(27) 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.
(28) Maddalena, D. Applications of Artificial Neural Networks to Quantitative Structure Activity Relationships. *Expert Opin. Ther. Pat.* **1996**, *6*, 239−251.
(29) Devillers, J. *Neural Networks in QSAR and Drug Design*; Academic Press: London, 1996.
(30) Tetko, I. V.; Livingstone, D. J.; Luik, A. I. Neural Network Studies. 1. Comparison of Overfitting and Overtraining. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 826−833.
(31) Shim, J. Y.; Collantes, E. R.; Welsh, W. J. Three-dimensional quantitative structure−activity relationship cannabimimetic (aminoalkyl) indoles using comparative molecular analysis. *J. Med. Chem.* **1998**, *41*, 4521−4532.
(32) D'Ambra, T. E.; Estep, K. G.; Bell, M. R.; Eissenstat, M. A.; Josef, K. A.; Ward, S. J.; Haycock, D. A.; Baizman, E. R.; Casiano, F. M.; Beglin, N. C.; Chippari, S. M.; Crego, J. D.; Kullnig, R. K.; Daley, G. T. Conformationally Restrained Analogues of Pravadoline: Nanomolar Potent, Enantioselective, (Aminoalkyl)indole Agonists of the Cannabinoid Receptor. *J. Med. Chem.* **1992**, *35*, 124−135.
(33) Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning internal representations by error propagation. In *Parallel Distributed Processing*: *Explorations in the Microstructure of Cognition*; Rumelhart, D. E., McClelland, J. L., Eds; The MIT Press: Cambridge, MA, 1986; pp 318−362.
(34) Tetko, I. V.; Villa, A. E. P. Efficient Partition of Learning Data Sets for Neural Network Training. *Neural Networks* **1997**, *10*, 1361−1374.
(35) Tetko, I. V.; Villa, A. E. P.; Livingstone, D. J. Neural network studies. 2. Variable selection. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 794−803.
(36) Kovalishyn, V. V.; Tetko, I. V.; Luik, A. I.; Kholodovych, V. V.; Villa, A. E. P.; Livingstone, D. J. Neural Network Studies. 3. Variable Selection in the Cascade-Correlation Learning Architecture. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 651−659.
(37) Kohonen, T.; Hinninen, J.; Kangas, J.; Laaksonen, J. SOM_PAK: The self-organizing map program package is obtainable via anonymous ftp from the Internet address "cochlea.hut.fi" (130.233.168.48).
(38) Korn, G. A., Korn, T. M. *Mathematical Handbook for Scientists and Engineers*; McGraw-Hill Inc.: New York, 1961.
(39) (a) Eissenstat, M. A.; Bell, M. R.; D'Ambra, T. E.; Alexander, E. J.; Daum, S. J.; Ackerman, J. H.; Gruett, M. D.; Kumar, V.; Estep, K. G.; Olefirowicz, E. M.; Wetzel, J. R.; Alexander, M. D.; Weaver, J. D., III; Haycock, D. A.; Luttinger, D. A.; Casiano, F. M.; Chippari, S. M.; Kuster, J. E.; Stevenson, J. I.; Ward, S. J. Aminoalkylindoles: Structure−Activity Relationships of Novel Cannabinoid Mimetrics. *J. Med. Chem.* **1995**, *38*, 3094−3105. (b) D'Ambra, T. E.; Eissenstat, M. A.; Abt, J.; Ackerman, J. H.; Bacon, E. R.; Bell, M. R.; Carabateas, P. M.; Josef, K. A.; Kumar, V.; Weaver, J. D., III; Arnold, R.; Casiano, F. M.; Chippari, S. M.; Haycock, D. A.; Kuster, J. E.; Luttinger, D. A.; Stevenson, J. I.; Ward, S. J.; Hill, W. A.; Khanolkar, A.; Makriyannis, A. C−Attached Aminoalkylindoles: Potent Cannabinoid Mimetics. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 17−22.
(40) Melvin, L. S.; Milne, G. M.; Johnson, M. R.; Subramaniam, B.; Wilken, G. H.; Howlett, A. G. Structure−Activity Relationships for Cannabinoid receptor-Binding and Analgesic Activity: Studies of Bicyclic Cannabinoid Analogues. *Mol. Pharmacol.* **1993**, *44*, 1008−1015.

JM010858E