

2D Autocorrelation modeling of the negative inotropic activity of calcium entry blockers using Bayesian-regularized genetic neural networks

Julio Caballero,^a Miguel Garriga^b and Michael Fernández^{a,*}

^a*Molecular Modeling Group, Center for Biotechnological Studies, Faculty of Agronomy, University of Matanzas, 44740 Matanzas, Cuba*

^b*Plant Biotechnology Group, Center for Biotechnological Studies, Faculty of Agronomy, University of Matanzas, Matanzas, CP 44740, Cuba*

Received 28 September 2005; revised 24 November 2005; accepted 22 December 2005

Available online 26 January 2006

Abstract—Negative inotropic potency of 60 benzothiazepine-like calcium entry blockers (CEBs), Diltiazem analogs, was successfully modeled using Bayesian-regularized genetic neural networks (BRGNNs) and 2D autocorrelation vectors. This approach yielded reliable and robust models whilst by means of a linear genetic algorithm (GA) search routine no multilinear regression model was found describing more than 50% of the training set. On the contrary, the optimum neural network predictor with five inputs described about 84% and 65% variances of 50 randomly selected training and test sets. Autocorrelation vectors in the nonlinear model contained information regarding 2D spatial distributions on the CEB structure of van der Waals volumes, electronegativities, and polarizabilities. However, a sensitivity analysis of the network inputs pointed out to the electronegativity and polarizability 2D topological distributions at substructural fragments of sizes 3 and 4 as the most relevant features governing the nonlinear modeling of the negative inotropic potency.

© 2006 Elsevier Ltd. All rights reserved.

1. Introduction

In the last two decades, the development of compounds capable of blocking calcium channels has provided much benefit to heart disorders. Due to their vascular properties, calcium entry blockers (CEBs) are used in the treatment of hypertension as well as in ischemic heart diseases.^{1,2} On the other hand, CEBs reduce the slow calcium current with a consequently direct negative inotropic effect by blocking myocardial calcium channels.³ However, vasodilatory response may affect cardiac performance by reflex-mediated changes masking negative inotropic activity. The best-known CEBs which are widely used as therapeutic agents in cardiovascular diseases are dihydropyridines (DHPs), such as Nifedipine, and non-DHPs, represented by Verapamil and Diltiazem.

Dihydropyridines are an important class of drugs able to block the Ca^{2+} currents through voltage-dependent L-type channels. They elicit the therapeutic effects by reversibly blocking Ca^{2+} influx through L-type calcium channels (LCCs Cav1) found in cardiac and vascular smooth muscle.⁴ In particular, splice variants Cav1.2a and Cav1.2b are associated with the heart and smooth muscle, respectively.⁵ On the other hand, Diltiazem, a 1,5-benzothiazepine, has affinity for the L-type calcium channels like the other CEBs, but it is less potent on the peripheral smooth muscle than on myocardial tissue. This weak selectivity produces negative chronotropic and inotropic effects that reduce the myocardial oxygen demand.⁶ In particular, the selective chronotropic and inotropic activity of new compounds might open new perspectives in the search for more effective drugs for the control of cardiac arrhythmias and led attention to the synthesis of novel Diltiazem-like CEBs.^{7–10}

Computational-based rational design of drugs has increased in the last decade. Most of those approaches are focused on quantitative structure–activity relationship (QSAR) studies, using different kinds of molecular

Keywords: Bayesian regularization; QSAR; Myocardial activity; Artificial neural networks; Genetic algorithm.

* Corresponding author. Tel.: +53 45 26 1251; fax: +53 45 25 3101; e-mail addresses: michael.fernandez@umcc.cu; michael_llamosa@yahoo.com

descriptors for encoding chemical information.^{11,12} After computing a set of descriptors, multivariate linear or/and nonlinear relationships are established between a reduced subset of variables and the biological activity, leading to a mathematical model.

Since interactions between a chemical and a biological system are often nonlinear by nature, artificial neural network (ANN) methodology has been successfully applied in QSAR studies of biological activities yielding, in most of the cases, better results than multilinear regression analysis (MRA).^{13–22} Besides the nonlinearity existing between biological activities and the computed molecular descriptors, another major problem arises when the number of calculated variables exceeds the number of compounds in the data set, so that one is dealing with an undetermined problem where undesirable overfitting can result.¹³ This problem can be handled by implementing a feature selection routine that determines which of the descriptors has a significant influence on the activity of a set of compounds. Genetic algorithm (GA), rather than forward or backward elimination procedures, has been successfully applied for feature selection in QSAR studies when the dimensionality of the data set is high and/or the interrelations between variables are convoluted.^{14,17,20–23} In this context, we recently introduced Bayesian-regularized genetic neural networks (BRGNNs) as a computational approach that combines Bayesian-regularized neural networks (BRANNs) and GA for building optimum nonlinear QSAR models.^{21–23}

Graph-theoretical and topological approaches have been included in several QSAR studies.^{12,13,20,23–31} Among these methods, 2D spatial autocorrelations^{24–27} have been successfully used in the last decades for modeling log *P*-values,²⁸ biological activities,^{26,27} for pharmaceutical^{29,30} and toxicological research.³¹ In addition, in recent works, our group has obtained outstanding results when such chemical code was used in combination with ANN approach in biological QSAR studies.^{20,23} Such results have inspired us to go forward on the application of these molecular descriptors on the study of biological interactions.

In this work, we employed Bayesian-regularized genetic neural networks (BRGNNs) for nonlinear modeling of the negative inotropic potency of 60 Diltiazem-like CEBs (Table 1). BRGNN approach was used for building an optimum neural network model using 2D autocorrelation vectors for encoding the chemical information. In order to gain in performance in both robustness of predictions and speed of computation for the ANNs, BRGNNs incorporate Bayesian regularization that was implemented in a Levenberg–Marquardt algorithm for error minimization during supervised training of full-connected feed-forward ANNs.

2. Results and discussion

In a first attempt, we searched for linear dependences between the negative inotropic potency of the studied

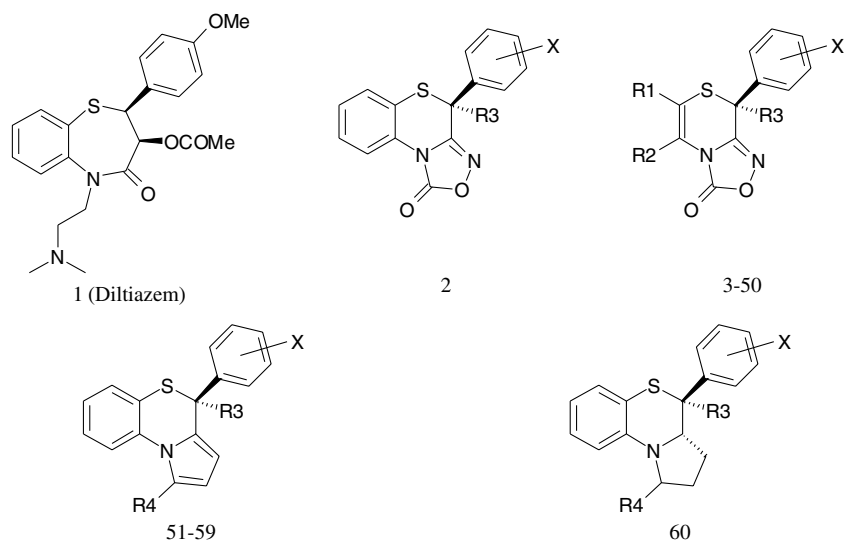
CEBs and the 2D spatial autocorrelation vectors we used for structural information encoding. In this sense, linear subspace in the data set was explored using multilinear regression models with number of variables ranging from 3 to 8 inside a GA routine. However, even after deleting some compounds from the training set as outliers, no multilinear regression model with $R^2_{\text{train}} > 0.5$ was found. This result evidences we are dealing with a phenomenon having high nonlinear components. Therefore, a computational method able to establish reliable nonlinear relationships between the chemical information encoded in the 2D spatial autocorrelation vectors and pEC₅₀, a very complex and multi-factor-dependent biological activity, is here mandatory. So, in this work from this point to further the modeling of the negative inotropic potency of the studied CEBs was faced using BRGNN approach.

2.1. Bayesian-regularized genetic neural networks (BRGNN) simulations

Similar to the multilinear GA search, in the BRGNN framework the nonlinear sub-space in the data set was explored varying the number of network inputs from 3 to 8. In this connection, optimum neural network predictors were found having five 2D autocorrelation vectors as network inputs. Concerning this, Figure 1 depicts plots of data fitting for optimum predictor but varying the number of hidden nodes in the network. As can be observed, improvement in training and test sets fitting was increased with the increment of the number of hidden neurons until an optimum architecture was reached. This behavior of the ANNs is due to the Bayesian regularization, which adjusts the network weights and biases to an extent in which the network outputs fit the targets (biological activity) in an optimum way, yielding a practically architecture-independent network.

Furthermore, Table 2 shows statistics and variables for optimum BRGNN predictor with different number of hidden nodes. The variables in the nonlinear models represent, *GATS8v* the Geary autocorrelation of a topological structure of lag 8 weighted by van der Waals volumes, *GATS8e* the Geary autocorrelation of a topological structure of lag 8 weighted by Sanderson electronegativities, *GATS3e* the Geary autocorrelation of a topological structure of lag 3 weighted by Sanderson electronegativities, *GATS6e* the Geary autocorrelation of a topological structure of lag 6 weighted by Sanderson electronegativities, and *MATS4p* the Moran autocorrelation of a topological structure of lag 4 weighted by polarizabilities. By inspection of Table 2 it can be observed that Bayesian regularization yielded quite stable and reliable networks. However, among those neural networks the optimum predictor was BRGNN 4 with 4 hidden nodes (neurons in the hidden layer) having adequate average values of R^2_{train} and R^2_{test} for training and test sets fitting about 0.84 and 0.65, respectively.

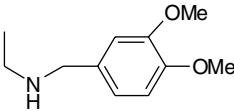
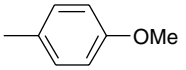
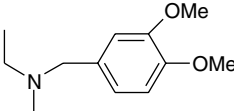
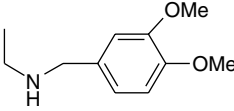
Concerning the possibility of chance correlations, following the method used by So and Karplus in Ref. 14, we performed a randomization test. Randomized values

Table 1. Chemical structures of CEBs and experimental, predicted, and residual negative inotropic potency according to model BRGNN 4

Compound	R1	R2	R3	R4	X	pEC ₅₀		
						Experimental	Predicted	Residual
1	—	—	—	—	—	0.102	−0.035	0.138
2	—	—	OH	—	4-Cl	0.284	0.394	−0.110
3	H	CH ₃	OH	—	4-Cl	0.097	0.238	−0.141
4	H	H	OH	—	4-Cl	−0.610	−0.618	0.009
5	CH ₃	CH ₃	OH	—	4-Cl	−0.152	−0.031	−0.122
6	CH ₃	H	OH	—	4-Cl	−0.236	−0.233	−0.002
7	H	CH ₃	OH	—	H	0.638	0.257	0.382
8	H	CH ₃	OH	—	4-Br	0.495	0.383	0.112
9	H	CH ₃	OH	—	4-F	−0.083	−0.086	0.003
10	H	CH ₃	OH	—	4-CH ₃	0.244	0.131	0.113
11	H	CH ₃	OH	—	4-Ph	−0.423	−0.358	−0.065
12	H	CH ₃	OH	—	4-CN	0.119	−0.012	0.131
13	H	CH ₃	OH	—	4-OCH ₃	0.284	0.306	−0.022
14	H	CH ₃	OH	—	4-NO ₂	−0.217	−0.271	0.054
15	H	CH ₃	OH	—	3-Cl	0.187	0.332	−0.145
16	H	CH ₃	OH	—	3-CH ₃	0.143	0.292	−0.150
17	H	CH ₃	OH	—	3-OCH ₃	0.310	0.226	0.084
18	H	CH ₃	OH	—	3-NO ₂	−0.246	−0.193	−0.052
19	H	CH ₃	OH	—	3-CF ₃	−0.336	−0.277	−0.060
20	H	CH ₃	OCH ₃	—	4-Cl	−0.188	−0.066	−0.122
21	H	CH ₃	OCH ₂ CH(CH ₃) ₂	—	4-Cl	0.252	0.328	−0.076
22	H	H	OH	—	H	−0.818	−0.839	0.020
23	H	CH ₃	OCH ₂ CH(CH ₃) ₂	—	4-Br	0.509	0.485	0.024
24	H	CH ₃	OCH ₂ CH ₃	—	4-Cl	0.569	0.580	−0.012
25	H	CH ₃	OCH ₂ CH ₃	—	4-Br	1.398	1.161	0.237
26	H	CH ₃	O(CH ₂) ₂ CH ₃	—	4-Cl	0.051	0.118	−0.067
27	H	CH ₃	O(CH ₂) ₂ CH ₃	—	4-Br	0.387	0.135	0.253
28	H	CH ₃	O(CH ₂) ₃ CH ₃	—	4-Cl	−0.267	0.032	−0.299
29	H	CH ₃	O(CH ₂) ₃ CH ₃	—	4-Br	−0.009	0.042	−0.050
30	H	CH ₃	O(CH ₂) ₅ CH ₃	—	4-Cl	0.194	0.173	0.021
31	H	CH ₃	O(CH ₂) ₅ CH ₃	—	4-Br	0.081	0.187	−0.106
32	H	CH ₃	O(CH ₂) ₈ CH ₃	—	4-Cl	0.102	0.156	−0.054
33	H	CH ₃	O(CH ₂) ₈ CH ₃	—	4-Br	0.456	0.449	0.007
34	H	CH ₃	O(CH ₂) ₁₁ CH ₃	—	4-Cl	0.060	0.041	0.020
35	H	CH ₃	O(CH ₂) ₁₁ CH ₃	—	4-Br	0.886	0.413	0.473
36	H	CH ₃	O(CH ₂) ₁₄ CH ₃	—	4-Cl	−0.053	−0.095	0.042
37	H	CH ₃	O(CH ₂) ₁₄ CH ₃	—	4-Br	0.377	0.287	0.090
38	H	CH ₃	O(CH ₂) ₁₇ CH ₃	—	4-Br	0.013	0.166	−0.153
39	H	CH ₃	OCH(CH ₃) ₂	—	4-Cl	0.658	0.537	0.120
40	H	CH ₃	OCH(CH ₃) ₂	—	4-Br	0.569	0.703	−0.135
41	H	CH ₃	OCH ₂ CH=CH ₂	—	4-Cl	0.432	0.216	0.216
42	H	CH ₃	OCH ₂ CH=CH ₂	—	4-Br	0.260	0.275	−0.015
43	H	CH ₃	OCH ₂ CH=CHCH ₂ CH ₃ Z-isomer	—	4-Cl	−0.117	0.211	−0.328

(continued on next page)

Table 1 (continued)

Compound	R1	R2	R3	R4	X	pEC ₅₀		
						Experimental	Predicted	Residual
44	H	CH ₃	OCH ₂ CH=CHCH ₂ CH ₃ Z-isomer		4-Br	0.201	0.232	−0.032
45	H	CH ₃	OCH ₂ CH=CH(CH ₂) ₇ CH ₃ Z-isomer		4-Cl	0.060	−0.007	0.067
46	H	CH ₃	OCH ₂ CH=CH(CH ₂) ₇ CH ₃ Z-isomer		4-Br	−0.100	0.358	−0.459
47	H	CH ₃	OCH ₂ Ph		4-Cl	0.131	0.124	0.007
48	H	CH ₃	OCH ₂ Ph		4-Br	0.229	0.148	0.081
49	H	CH ₃	OCH ₂ -cyclohexyl		4-Cl	0.620	0.521	0.099
50	H	CH ₃	OCH ₂ -cyclohexyl		4-Br	0.721	0.588	0.133
51	—	—	H	CH ₂ N(CH ₃) ₂	H	0.076	0.110	−0.034
52	—	—	H		4-OCH ₃	0.347	0.317	0.030
53			OCOCH ₃	H	H	0.252	0.234	0.017
54	—	—	CH ₂ N(CH ₃) ₂		4-OCH ₃	0.036	0.010	0.026
55	—	—	H		4-OCH ₃	−0.176	−0.073	−0.103
56	—	—			2,4-Di-OCH ₃	0.420	0.550	−0.130
57	—	—	H	CH ₂ N(CH ₃) ₂	4-OCH ₃	0.620	0.687	−0.067
58	—	—	OCOCH ₃	H	4-OCH ₃	0.051	−0.035	0.086
59	—	—	Ph	CH ₂ N(CH ₃) ₂	H	0.137	0.155	−0.018
60	—	—	H	H	4-OCH ₃	−0.079	−0.114	0.035

Compounds 1–22 were taken from Ref. 9.

Compounds 23–50 were taken from Ref. 10.

Compounds 51–60 were taken from Ref. 8.

were given to the dependent variable (biological activities) and networks were trained using this randomized target and the real set of five 2D spatial autocorrelation vectors (independent variables). By repeating these processes 500 times no correlation was found between R^2 values of training and test sets, similar to the results of So and Karplus.¹⁴

The correlation matrix of the descriptors representing the optimum BRGNN inputs is shown in Table 3. As can be observed, the BRGNN approach selected non-correlated variables, so different information is brought to the model by each autocorrelation vector. In addition, the fact that autocorrelation vectors weighted by different properties (van der Waals volume, atomic masses, and polarizabilities) occurred in the optimum nonlinear predictor pointed out we are dealing with a very complex activity that depends on different types of molecular interactions.

Aiming to figure out the impact of each input in the predictor BRGNN 4, we performed a sensibility analysis in

two ways.³² Descriptor under study was removed from the BRGNN together with its corresponding weights and the statistical coefficient R^2_{train} for training set was analyzed. Comparison between these R^2 values and that calculated when no descriptor was removed gave an idea of the importance of the descriptor removed. On the other way, mean of the absolute deviation values Δmi between the observed and estimated value for all compounds was calculated. Finally, the contribution factor C_i^{32} of descriptor i ($i = 1-5$) is given by:

$$C_i = \frac{100 \times \Delta mi}{\sum \Delta mi} \quad (1)$$

The results of the sensitivity analysis appear in Table 4. As can be observed, both approaches for measuring the descriptors' importance agree well. In this sense, the higher decreases in R^2 corresponded to the higher values of C_i . Among the five descriptors in the model BRGNN 4, descriptors with highest impacts ($C_i > 20\%$ and $R^2_{\text{train}} < 0.5$) are *GATS3e* and *MATS4p* which are autocorrelation vectors weighted by Sanderson electroneg-

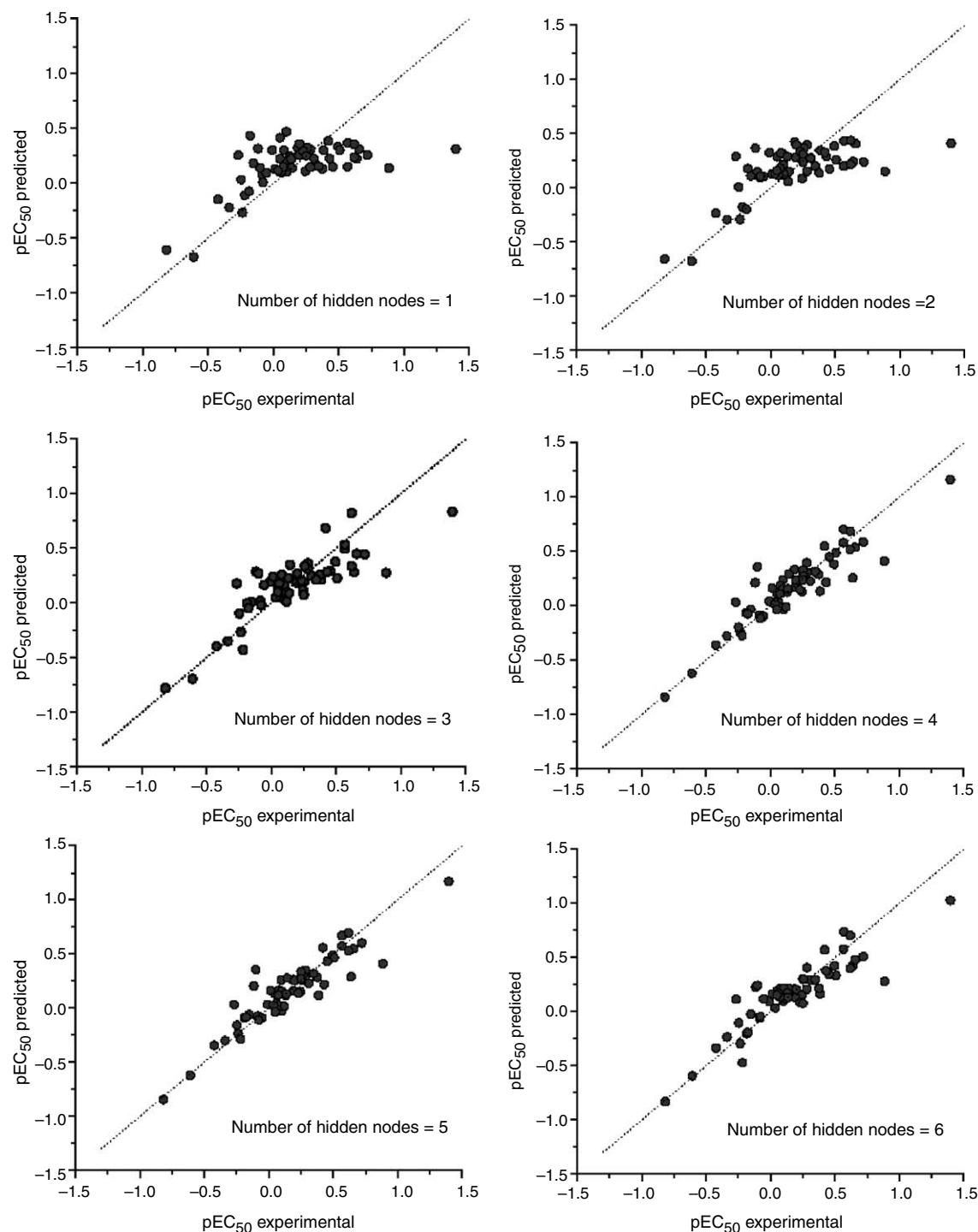


Figure 1. Plots of predicted versus experimental negative inotropic potency (pEC_{50}) of the CEBs by the nonlinear predictors trained with optimum five 2D autocorrelation vectors yielded by BRGNN but varying the number of hidden nodes. The dotted lines are an ideal fit with the respective intercept and slopes equal to zero and one.

activities and polarizabilities representing topological substructures of sizes 3 and 4, respectively, in the CEB molecule.

In addition to the sensitivity analyses that allows getting a general notion about the impact of each descriptor in an ANN model, the black-box nature of three-layer ANNs has been ‘deciphered’ in a recent report of Guha

et al.³³ Their method allows understanding how an input descriptor is correlated to the predicted output by the network and consists of two parts. First, the nonlinear transform for a given neuron is linearized. Afterward, the magnitude in which a given neuron affects the downstream output is determined. Next, a ranking scheme for neurons in the hidden layer is developed. The ranking scheme is carried out by determining the square contri-

Table 2. Statistics of the optimum nonlinear predictors for the negative inotropic activity of the Diltiazem-like CEBs

Descriptors	BRGNN model	Hidd nod	Num par	Average over 50 training/test partitions				
				Opt par	R^2_{train}	S_{train}	R^2_{test}	S_{test}
<i>MATS4p</i> , <i>GATS8v</i> , <i>GATS3e</i> , <i>GATS6e</i> , <i>GATS8e</i>	1	1	8	6	0.352	0.302	0.336	0.602
	2	2	15	9	0.498	0.261	0.488	0.501
	3	3	22	15	0.712	0.201	0.609	0.402
	4	4	29	20	0.842	0.123	0.652	0.343
	5	5	36	22	0.840	0.126	0.602	0.410
	6	6	43	22	0.839	0.125	0.570	0.435

Optimum neural network predictor appears in bold letter.

Hidd nod represents the number of hidden nodes, Num par represents the number of neural network parameters, Opt par represents the average optimum number of neural network parameters yielded by the Bayesian regularization, R^2_{train} and R^2_{test} are average square correlation coefficients of 50 randomly selected training and test sets fitting, respectively, S_{train} and S_{test} are average standard deviations of 50 randomly selected training and test sets fitting, respectively.

Table 3. Correlation matrix of the inputs of the optimum BRGNN predictor

	<i>MATS4p</i>	<i>GATS8v</i>	<i>GATS3e</i>	<i>GATS6e</i>	<i>GATS8e</i>
<i>MATS4p</i>	1	0.009	0.087	0.264	0.008
<i>GATS8v</i>		1	0.02	0.014	0.168
<i>GATS3e</i>			1	0.262	0.004
<i>GATS6e</i>				1	0.162
<i>GATS8e</i>					1

Table 4. Results of the sensitivity analysis for evaluating the impact of each descriptor in model BRGNN 4

Model	Removed descriptor	$C\%_{\text{train}}$	R^2_{train}
BRGNN 4	<i>MATS4p</i>	22.8	0.413
	<i>GATS8v</i>	18.8	0.606
	<i>GATS3e</i>	24.4	0.390
	<i>GATS6e</i>	16.9	0.716
	<i>GATS8e</i>	17.1	0.698
	None	100	0.811

$C\% > 20\%$ and $R^2_{\text{train}} < 0.5$ appear in bold letter.

bution values (SCV) for each hidden neuron (see Ref. 33 for details). This method for ANN model interpretation is similar in manner to the partial least-squares interpretation method for linear models described by Stanton.³⁴ Results of the ANN deciphering study appear in Table 5. The reported effective weight matrix for the optimum model BRGNN 4 reflects that it is very difficult the interpretation of the interactions between descriptors and the biological activity under study. First, three of

Table 5. Effective weight matrix for the 5-4-1 BRGNN 4 model developed for the inotropic activity of the Diltiazem-like CEBs^a

Inputs	Hidden neuron			
	3	4	1	2
<i>MATS4p</i>	1.841	−1.549	0.393	−0.973
<i>GATS8v</i>	−1.965	−0.307	2.328	−0.127
<i>GATS3e</i>	−1.655	1.557	0.764	0.044
<i>GATS6e</i>	−1.436	−1.428	1.495	1.006
<i>GATS8e</i>	0.280	0.071	−1.038	−0.196
SCV	0.366	0.332	0.264	0.039

^a The columns are ordered by the SCVs for the hidden neurons, shown in the last row.

the four hidden neurons (3, 4, and 1) in the model have about the same contribution to the model (SCV = 0.366, 0.332, and 0.264). Second, the most weighted descriptors (*MATS4p*, *GATS8v*, *GATS3e*, and *GATS6e*) in these three most important hidden neurons have effective weights with different signs. These facts confirmed the complexity of the vectorial subspace resembled for the autocorrelation vectors used as input of the optimum predictor BRGNN 4.

2.2. Model's interpretation

In spite of several reports having been focused on QSAR studies on ion channel blocker activity,^{17,44} little has been done on the modeling of myocardial activity of CEBs. In this sense, Budriesi et al.¹⁰ reported a 3D QSAR study using GRID descriptors on a series of Diltiazem analogs they synthesized and other previously published analog compounds,⁹ all of them including in our work. They identified some pharmacophoric features that favorably affect the inotropic potency: one positively charged center, three lipophilic groups, and two hydrogen-bonding acceptor groups. This fact well agrees with our results taking into account our optimum BRGNN model present five 2D spatial autocorrelation vectors gathering structural information related to van der Waals volumes, electronegativities, and polarizabilities distribution on the chemical structure, pointing out to multifactor receptor–drug interactions. We concluded that the BRGNN approach was able to gather different interactions governing the myocardial activity of the studied CEBs, in a way that a unique nonlinear model satisfactorily comprises compounds displaying extreme high (compound **25**) or low (compounds **4** and **22**) activities.

Interpreting a QSAR model in terms of the specific contribution of substituents and other molecular features to the modeled activity is always a difficult work. In spite of the obvious limited information from 2D representations of molecules, 2D autocorrelation descriptors have found successful applications in the performance of ANN-based QSAR modeling in previous reports of our group.^{20,23} Although more evident information can be obtained from ‘traditional’ descriptors, such as log *P* or p*K*_a, several authors pointed out problems with the definition of an adequate chemistry space using

them.^{11c} In general, ‘traditional’ descriptors often are highly correlated and cannot distinguish the details of important substructural differences. Likewise, 3D descriptors are highly dependent on conformations of molecules; consequently, an adequate three-dimensional application requires the identification of active conformations (the receptor-bound conformation) among multiple molecular shapes of the compounds.

The 2D autocorrelation descriptors represent the topological structure of the compounds, but are more complex in nature when compared to the classical topological descriptors. The computation of these descriptors involves the summations of different autocorrelation functions corresponding to different structural lags and leads to different autocorrelation vectors corresponding to the lengths of substructural fragments. Bearing in mind this aspect, the interpretation of 2D autocorrelation descriptors is uneasy.

Basically, the pool of 2D autocorrelation descriptors defines a wide 2D space. On behalf of a greater applicability, physicochemical properties (atomic masses, atomic van der Waals volumes, atomic Sanderson electronegativities, and atomic polarizabilities) were inserted as weighting components. As a result, these descriptors address the topology of the structure or parts thereof in association with a specific physicochemical property. For a sound application, we found that a very few structural keys yield better performances than the complete set. In point of fact, the BRGNN method selected an optimum descriptor combination, which includes van der Waals volumes, electronegativities, and polarizabilities as the most relevant key features. This result illustrates that a certain distribution of these properties is necessarily required for typifying the negative inotropic potency of the studied CEBs.

To the best of our knowledge, our study is the first comprehensive study regarding the modeling of negative inotropic activity of Diltiazem-like CEBs by 2D QSAR. The complexity of the biological property under investigation led us to the use of ANN modeling rather than a linear approach, making model interpretation harder. However, by means of a sensitivity analysis of the neural network inputs and an effective weight study, we gained a deeper insight into the BRGNN model. Those analyses showed the complexity of the developed model mainly governed by highly nonlinear interactions among the autocorrelation vectors and the inotropic activity. Although, from this analyses was derived the fact that autocorrelations of electronegativities and polarizabilities at substructural fragments of size 3 and 4 in the CEB molecule are the most relevant features governing the nonlinear modeling of the negative inotropic potency. This fact may be viewed in terms of association of activity information content with structural fragments of such size and could be related with the great importance of an adequate molecular size and/or shape for a proper matching into the receptor. However, further deciphering of the information content of these descriptors is very complex as their computations involve integration of the structural fragments and due to this it is

not possible to traverse backward from a higher state to a lower one.¹³ These facts make our model mainly interesting as a predictive tool rather than for designing or an approach for explanation of the structure property trends that are encoded within the model. Regarding this, our BRGNN predictor should be useful for predictive purposes, as an ‘in silico’ filter, to screen for new potent myocardial active CEBs.

3. Concluding remarks

Since biological phenomena are complex by nature, in this work the negative inotropic potency of 60 CEBs, Diltiazem analogs, was successfully modeled using a hybrid approach that combines GA and BRANNs named BRGNN. 2D spatial autocorrelation vectors demonstrated again to encode relevant structural information from the studied compounds that highly correlate in a nonlinear way with the myocardial activity. Non-reliable linear relationship was found by means of GA search routine even deleting some compounds as outliers. However, BRGNN approach yields a reliable predictor describing about 84% and 65% of 50 training and test set partitions’ variance, respectively.

Furthermore, a sensitivity analysis of the inputs of the optimum BRGNN yielded that autocorrelations of electronegativity and polarizabilities at substructural fragments of size 3 and 4 in the CEB molecule are the most relevant features governing the nonlinear modeling of the negative inotropic potency.

The present work demonstrates the wide applicability of BRGNN approach like a robust and sensitive method to quantify structure–activity relationships. In general, ANNs are more appropriate than other approaches for particular problems at hand. Their general ability to model nonlinear effects makes them useful tools to solve relationships when traditional linear models fail. Moreover, Bayesian regularization is a way to avoid the well-known shortcomings of neural networks; in this sense, the combination of BRANNs with GA (BRGNNs) led to a powerful method for the scientific community interested in handling complex biological activities.

4. Material and methods

4.1. 2D Spatial autocorrelation approach

The binding of a substrate to its receptor is dependent on the shape of the substrate and on a variety of effects such as the molecular electrostatic potential, polarizability, hydrophilicity, and lipophilicity. Therefore, in a QSAR study the strategy for encoding molecular information must, in some way, either explicitly or implicitly, account for these physicochemical effects. Furthermore, usually data sets include molecules of different sizes with different number of atoms, so the structural encoding structures must allow comparing such molecules. Thus, we were faced with the problem of having to compare

molecules with different number of atoms. Information having variable length can be transformed into fixed-length information by autocorrelation.³⁰

Autocorrelation vectors have several useful properties. First, a substantial reduction in data can be achieved by limiting the topological distance, l . Second, the autocorrelation coefficients are independent of the original atom numberings, so they are canonical. And third, the length of the correlation vector is independent of the size of the molecule.³⁰

For the autocorrelation vectors, H-depleted molecular structure is represented as a graph G and physicochemical properties of atoms (i.e., atomic masses, atomic van der Waals volumes, atomic Sanderson electronegativities, and atomic polarizabilities) as real values assigned to the vertices of G (Table 6).

These descriptors can be obtained by summing up the products of certain properties of two atoms, located at given topological distances or spatial lag in G . Three spatial autocorrelation vectors were employed for modeling the inotropic activity.

- Moran's index:²⁴

$$MATSp_k l = \frac{N}{2L} \frac{\sum_{ij} \delta_{ij} (p_{ki} - \bar{p}_k)(p_{kj} - \bar{p}_k)}{\sum_i (p_{ki} - \bar{p}_k)} \quad (2)$$

- Geary's coefficient:²⁵

$$GATSp_k l = \frac{(N-1)}{4L} \frac{\sum_{ij} \delta_{ij} (p_{ki} - \bar{p}_k)(p_{kj} - \bar{p}_k)}{\sum_i (p_{ki} - \bar{p}_k)} \quad (3)$$

- Broto-Moreau's autocorrelation coefficient:²⁶

$$ATSp_k l = \sum_i \delta_{ij} p_{ki} p_{kj} \quad (4)$$

where $MATSp_k l$, $GATSp_k l$, and $ATSp_k l$ are Moran's index, Geary's coefficient, and Broto-Moreau's autocorrelation coefficient at spatial lag l , respectively; p_{ki} and p_{kj} are the values of property k of atoms i and j , respectively; \bar{p}_k is the average value of property k ; L is the number of non-zero elements in the sum and $\delta(l, d_{ij})$ is a Dirac-delta function defined as

$$\delta(l, d_{ij}) = \begin{cases} 1 & \text{if } d_{ij} = l, \\ 0 & \text{if } d_{ij} \neq l, \end{cases} \quad (5)$$

where d_{ij} is the topological distance or spatial lag between atoms i and j .

Spatial autocorrelation measures the level of interdependence between properties, and the nature and strength of that interdependence. It may be classified as either positive or negative. In a positive case, all similar values appear together, while a negative spatial autocorrelation has dissimilar values appearing in close association.^{24,25} In a molecule, Moran's and Geary's spatial autocorrelation analysis tests whether the value of an atomic property at one atom in the molecular structure is independent of the values of the property at neighboring atoms. If dependence exists, the property is said to exhibit spatial autocorrelation. Moreau and Broto first applied autocorrelation function to the topology of molecular structures.^{24,25} The autocorrelation vectors represent the degree of similarity between molecules. In addition, 2D spatial autocorrelation code has been successfully applied in nonlinear QSAR studies, proving to contain relevant nonlinear information concerning different biological phenomena.^{20,23,29,30}




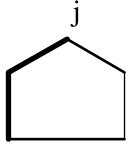
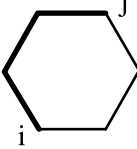

A data matrix is generated with the spatial autocorrelation vectors calculated for each compound. Afterwards, BRGNN was employed for building an optimum QSAR model (Fig. 2).

4.2. Data set and molecular descriptors

In the present study, a data set of 60 Diltiazem-like CEBs for which their negative inotropic activities are reported in the literature^{8–10} was used. In order to assure the robustness of the models, data set was several times (50) randomly divided into training sets (48 compounds, 80% of the data set) used for model building and test sets (12 compounds, 20% of the data set) used for external validation of the models. In this regard, we reported averages of square correlation coefficient for training (R^2_{train}) and test (R^2_{test}) sets fitting. Molecular structures, numbering of the substituents, and activities of the CEBs are summarized in Table 1.

By means of a geometry optimization step chemical structures with optimized Cartesian Coordinates were obtained from which the connectivity and topological distance matrixes can be derived for topological descriptors' calculation. In this way, we used the quantum chemical semi-empirical method PM3³⁵ included in Mopac 6.0³⁶ computer software. Afterwards, Dragon³⁷ computer software was employed to calculate the 2D autocorrelation vectors at spatial lags ranging from 1 to 8. As weighting properties we tried all available prop-

Table 6. Representation of different molecular graphs G and topological distances or spatial lags d_{ij}

Molecular graphs G						
d_{ij}	1	1	2	2	3	4

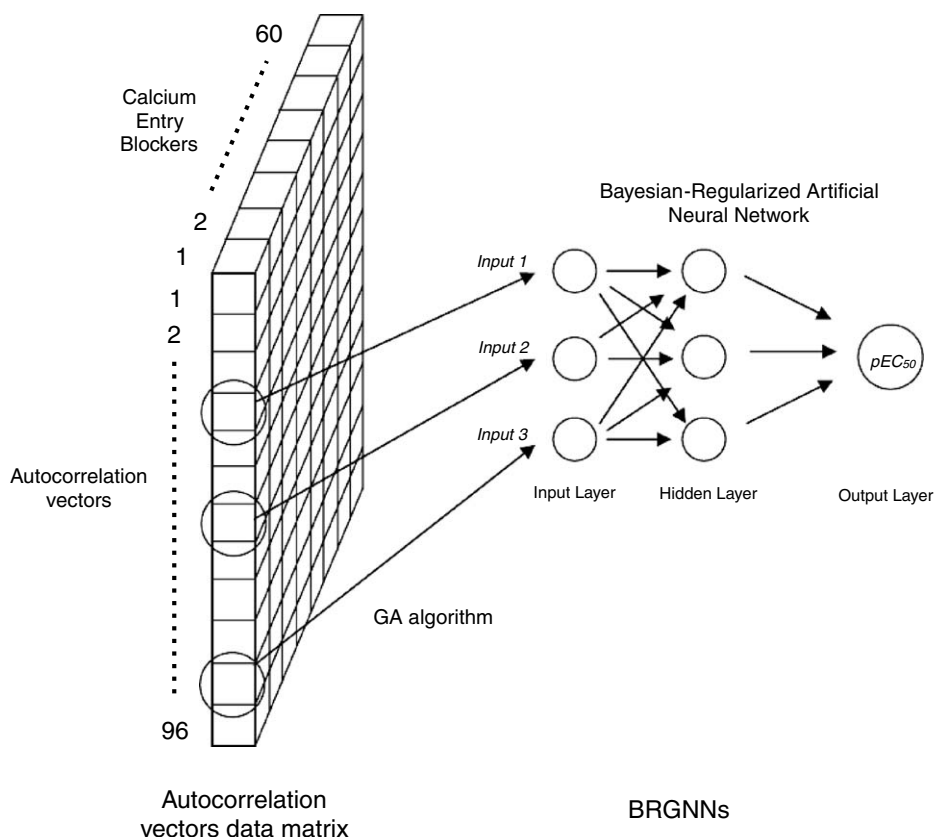


Figure 2. Schematic representation of Bayesian-regularized genetic neural network (BRGNN) technique with a prototype back-propagation neural network with 3–3–1 architecture. Autocorrelation vectors chosen by genetic algorithm constitute inputs and network is trained against negative inotropic activity (pEC_{50}) of the CEBs.

erties in Dragon software (atomic masses, atomic van der Waals volumes, atomic Sanderson electronegativities, and atomic polarizabilities). Since we calculated three types of autocorrelation descriptors (Moran's index, Geary's coefficient, and Broto-Moreau's autocorrelation coefficient) weighted by 4 atomic properties at 8 lags, a total of 96 ($3 \times 4 \times 8$) descriptors were computed.

4.3. Bayesian-regularized genetic neural networks (BRGNNs)

4.3.1. Artificial neural networks (ANNs). Artificial Neural Networks (ANNs) are computer-based models in which a number of processing elements, also called neurons, units, or nodes, are interconnected by links in a net-like structure forming 'layers'.³⁸ A variable value is assigned to every neuron. The neurons can be one of three different kinds. The input neurons receive their values from independent variables, input layer. The hidden neurons collect values from other neurons, giving a result that is passed on to a successor neuron. The output neurons take values from other units and correspond to different dependent variables, forming the output layer. In this sense, network architecture is commonly represented as $I-H-O$, where I , H , and O are the number of neurons in the input, hidden, and output layers, respectively (Fig. 2).

The links between units have associated values, named weights, that condition the values assigned to the neu-

rons. There exist additional weights assigned to bias values that act as neuron value offsets. The weights are adjusted through a training process in order to minimize network error. Commonly neural networks are adjusted, or trained, so that a particular input leads to a specific target output.

In contrast to common statistical methods, ANNs are not restricted to linear correlations or linear subspaces.³⁸ They can take into account nonlinear structures and structures of arbitrarily shaped clusters or curved manifolds. The characteristics of the ANNs have been found to be suitable for data processing, in which the functional relationship between the input and the output is not previously defined. This is due to the fact that structure-activity relationships are often nonlinear and very complex, and neural networks are able to approximate any kind of analytical continuous function, according to Kolmogorov's theorem.³⁹ As biological phenomena are often nonlinear by nature, ANN technique has been successfully applied in order to discover the possible existence of nonlinear relationships between biological activity and molecular descriptors that are ignored for the linear approach.⁴⁰

4.3.2. Bayesian-regularized artificial neural networks (BRANNs). Typically, neural network training aims to reduce the mean square errors of the network $F = MSE$. Regularization involves modifying the perfor-

mance function, usually known as cost function (F). It is possible to improve generalization if an additional term is added.

$$F = \beta \times MSE + \alpha \times MSW, \quad (6)$$

where MSW is the sum of squares of the network weights and biases, and α and β are objective function parameters. The relative size of the objective function parameters dictates the emphasis for training, getting a smoother network response. MacKay's Bayesian regularization automatically sets the correct values for the objective function parameters,⁴¹ in this sense the regularization is optimized.

Bayesian regularization overcomes the remaining deficiencies of neural networks.⁴² By using Bayesian regularization robust models, well matched to the data and able to make accurate predictions can be obtained. Since the algorithm automatically regularizes the training process usually nonvalidation set occurs so all data set can be devoted to train the network.¹⁵ The Bayesian neural nets have the potential to give models which are relatively independent of neural network architecture, above a minimum architecture, and the Bayesian regularization method estimates the number of effective parameters. Bayesian Regularized ANNs (BRANNs) have been successfully used on QSAR studies of biological activities and specifically in drug discovery.^{15,16,18,19,21–23}

Our BRANNs are classical back-propagation neural nets that incorporate the Bayesian regularization algorithm for finding the optimum weights. The Bayesian regularization takes place within the Levenberg–Marquardt algorithm⁴¹ implemented in Matlab environment.⁴² The input and output values were normalized prior network training.

4.3.3. Bayesian-regularized genetic neural networks (BRGNNs). In other respects, choosing the adequate descriptors for QSAR studies is difficult because there are no absolute rules that govern this choice. Evolutionary algorithms and specifically genetic algorithms have been used for variable selection problems combined to ANNs.^{14,17,20–23} Since 96 2D autocorrelation vectors were available for QSAR analysis and only a subset of them is statistically significant in terms of correlation with the modeled activity, GA searches were carried out for building optimum linear and nonlinear models. GAs are stochastic optimization methods that have been inspired by evolutionary principles. The distinctive aspect of a GA is that it investigates many possible solutions simultaneously, each of which explores different regions in parameter space.¹⁴ The first step is to create a population of models. These models mate with each other, mutate, crossover, reproduce, and then evolve through successive generations toward an optimum solution.

By combining the concepts of BRANN and GAs, BRGNNs are implemented in such a way that BRANN inputs are selected inside the GA framework (Fig. 2). The BRGNN approach is a version of the So and Karplus report¹⁴ incorporating Bayesian regularization that

was recently reported by us²¹ and was programmed within the Matlab environment using Genetic Algorithm⁴³ and Neural Networks Toolboxes.⁴¹ BRGNN has been successfully applied by our group for modeling the inhibitory activity of several therapeutic target enzymes.^{21–23}

The implemented BRGNN searches for the best fitted BRANN, in such a way that from one generation to another the GA tried to minimize the MSE of the networks (fitness function). By employing this approach, instead of a more complicated and time-consuming cross-validation-based fitness function, we gain in CPU time and simplicity of the routine. Furthermore, we can devote whole data set to train the networks. However, the use of the MSE fitness function could lead to undesirable well-fitted but poor generalized networks as algorithm solutions. In this connection, BRGNN approach avoids such results by two aspects: (1) keeping network architectures as simple as possible inside the GA framework, 1–2–1 (Section 4.3.1), less hidden neurons in the network means less parameters in the model and greater generalization abilities and (2) implementing Bayesian regularization in the network training function (Section 4.3.2) that assures also well-generalized networks.

References and notes

- Dustan, H. P. *Hypertension* **1989**, *13*, 1137.
- Wilson, S.; Colucci, M. D. *Am. J. Cardiol.* **1987**, *59*, 52B.
- Schwinger, R. H. G.; Böhm, M.; Erdmann, E. *Klin. Wochenschr.* **1990**, *68*, 797.
- Schleifer, K. J. *J. Med. Chem.* **1999**, *42*, 2204.
- Triggle, D. J. *Cell Mol. Neurobiol.* **2003**, *23*, 293.
- Triulzi, M. O.; Mattioli, R.; Signorini, G.; Cirino, D.; Esposti, D.; Aguggini, G.; Maggi, G. C. *Ital. Cardiol.* **1990**, *20*, 1137.
- (a) Inoue, H.; Konda, M.; Hashiyama, T.; Otsuka, H.; Takahashi, K.; Gaino, M.; Date, T.; Aoe, K.; Takeda, M.; Murata, S.; Narita, H.; Nagao, T. *J. Med. Chem.* **1991**, *34*, 675; (b) Floyd, D. M.; Kimball, S. D.; Krapcho, J.; Das, J.; Turk, C. F.; Moquin, R. V.; Lago, M. W.; Duff, K. J.; Lee, V. G.; White, R. E.; Ridgewell, R. E.; Moreland, S.; Brittain, R. J.; Normandin, D. E.; Hedberg, S. A.; Cucinotta, G. G. *J. Med. Chem.* **1992**, *35*, 756; (c) Das, J.; Floyd, D. M.; Kimball, S. D.; Duff, K. J.; Vu, T. C.; Lago, M. W.; Moquin, R. V.; Lee, V. G.; Gougoutas, J. Z.; Malley, M. F.; Moreland, S.; Brittain, R. J.; Hedberg, S. A.; Cucinotta, G. G. *J. Med. Chem.* **1992**, *35*, 773; (d) Kimball, S. D.; Floyd, D. M.; Das, J.; Hunt, J. T.; Krapcho, J.; Rovnyak, G.; Duff, K. J.; Lee, V. G.; Moquin, R. V.; Turk, C. F.; Hedberg, S. A.; Moreland, S.; Brittain, R. J.; McMullen, D. M.; Normandin, D. E.; Cucinotta, G. G. *J. Med. Chem.* **1992**, *35*, 780.
- Campiani, G.; Garofalo, A.; Fiorini, I.; Botta, M.; Nacci, V.; Tafi, A.; Chiarini, A.; Budriesi, R.; Bruni, G.; Romeo, M. R. *J. Med. Chem.* **1995**, *38*, 4393.
- Budriesi, R.; Cosimelli, B.; Ioan, P.; Lanza, C. Z.; Spinelli, D.; Chiarini, A. *J. Med. Chem.* **2002**, *45*, 3475.
- Budriesi, R.; Carosati, E.; Chiarini, A.; Cosimelli, B.; Cruciani, G.; Ioan, P.; Spinelli, D.; Spisani, R. *J. Med. Chem.* **2005**, *48*, 2445.
- (a) Rüker, G.; Rüker, C. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 683; (b) Gutman, I.; Rüker, C.; Rüker, G. *J. Chem.*

- Inf. Comput. Sci.* **2001**, 41, 739; (c) Pearlman, R. S.; Smith, K. M. *J. Chem. Inf. Comput. Sci.* **1999**, 39, 28; (d) Randic, M. *New J. Chem.* **1995**, 19, 781; (e) Hemmer, M. C.; Steinhauer, V.; Gasteiger, J. *Vibrat. Spectrosc.* **1999**, 19, 151; (f) Hemmer, M. C.; Gasteiger, J. *Anal. Chim. Acta* **2000**, 420, 145; (g) Schuur, J. H.; Setzer, P.; Gasteiger, J. *J. Chem. Inf. Comput. Sci.* **1996**, 36, 334; (h) Todeschini, R.; Vighi, M.; Provenzani, R.; Finizio, A.; Gramatica, P. *Chemosphere* **1996**, 32, 1527; (i) Bravi, G.; Gancia, E.; Mascagni, P.; Pegna, M.; Todeschini, R.; Zaliani, A. *J. Comput.-Aided Mol. Des.* **1997**, 11, 79; (j) Consonni, V.; Todeschini, P.; Pavan, M. *J. Chem. Inf. Comput. Sci.* **2002**, 42, 682.
12. González, M. P.; Morales, A. H. *J. Comput. Aided Mol. Des.* **2003**, 17, 665.
 13. Fernández, M.; Caballero, J.; Morales, A. H.; Castro, E. A.; González, M. P. *Bioorg. Med. Chem.* **2005**, 13, 3269.
 14. So, S. S.; Karplus, M. *J. Med. Chem.* **1996**, 39, 1521.
 15. Burden, F. R.; Winkler, D. A. *J. Med. Chem.* **1999**, 42, 3183.
 16. Burden, F. R.; Winkler, D. A. *Chem. Res. Toxicol.* **2000**, 13, 436.
 17. Hemmateenejad, B.; Akhond, M.; Miri, R.; Shamsipur, M. *J. Chem. Inf. Comput. Sci.* **2003**, 43, 1328.
 18. Winkler, D. A.; Burden, F. R. *Biosilico* **2004**, 2, 104.
 19. Winkler, D. A. *Mol. Biotech.* **2004**, 27, 139.
 20. Fernández, M.; Tundidor-Camba, A.; Caballero, J. *Mol. Simulat.* **2005**, 31, 575.
 21. Caballero, J.; Fernández, M. *J. Mol. Model.* **2006**, 12, 168.
 22. González, M. P.; Caballero, J.; Tundidor-Camba, A.; Helguera, A. M.; Fernández, M. *Bioorg. Med. Chem.* **2006**, 14, 200.
 23. Fernández, M.; Caballero, J. *Bioorg. Med. Chem.* **2006**, 14, 280.
 24. Moran, P. A. P. *Biometrika* **1950**, 37, 17.
 25. Geary, R. F. *The Incorporated Statistician* **1954**, 5, 115.
 26. Moreau, G.; Broto, P. *Nouv. J. Chim.* **1980**, 4, 359.
 27. Moreau, G.; Broto, P. *Nouv. J. Chim.* **1980**, 4, 757.
 28. Devillers, J.; Domine, D. *SAR QSAR Environ. Res.* **1997**, 7, 195.
 29. Wagener, M.; Sadowski, J.; Gasteiger, J. *J. Am. Chem. Soc.* **1995**, 117, 7769.
 30. Bauknecht, H.; Zell, A.; Bayer, H.; Levi, P.; Wagener, M.; Sadowski, J.; Gasteiger, J. *J. Chem. Inf. Comput. Sci.* **1996**, 36, 1205.
 31. Devillers, J. Autocorrelation descriptors for modelling (eco)toxicological endpoints. In *Topological Indices and Related Descriptors in QSAR and QSPR*; Devillers, J., Balaban, A. T., Eds.; Gordon and Breach Science Publishers, 1999; pp 595–612.
 32. Cherquaoui, D.; Esseffar, M.; Villemin, D.; Cence, J. M.; Chastrette, M.; Zakarya, D. *New J. Chem.* **1998**, 839–843.
 33. Guha, R.; Stanton, D. T.; Jurs, P. C. *J. Chem. Inf. Model.* **2005**, 45, 1109.
 34. Stanton, D. T. *J. Chem. Inf. Comput. Sci.* **2003**, 43, 1423.
 35. Stewart, J. J. P. *J. Comput. Chem.* **1989**, 10, 210.
 36. MOPAC version 6.0. Frank J. Seiler Research Laboratory, US Air Force Academy, Colorado Springs, CO, 1993.
 37. Todeschini, R.; Consonni, V.; Pavan, M. (2002) Dragon Software version 2.1.
 38. Sumpter, B. G.; Getino, C.; Noid, D. W. *Annu. Rev. Phys. Chem.* **1994**, 45, 439.
 39. Kolmogorov, A. N. *Dokl. Akad. Nauk. SSSR* **1957**, 114, 953.
 40. Mackay, D. J. C. *Neural Comput.* **1992**, 4, 415.
 41. The MathWorks Inc. Neural network toolbox user's guide for use with MATLAB, The Mathworks Inc., Massachusetts, 2004.
 42. Matlab version 7.0. The MathWorks, Inc., 2004.
 43. The MathWorks Inc. Genetic algorithm and direct search toolbox user's guide for use with MATLAB, The Mathworks Inc., Massachusetts, 2004.
 44. Takahata, Y.; Costa, M. C. A.; Gaudio, A. C. *J. Chem. Inf. Comput. Sci.* **2003**, 43, 540.