Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 12 (2004) 5833–5843
Bioorganic & Medicinal Chemistry

In silico prediction of central nervous system activity of compounds. Identification of potential pharmacophores by the TOPS–MODE approach

Miguel Angel Cabrera Pérez[a,*] and Marival Bermejo Sanz[b]

[a]*Drug Design Department, Center of Chemical Bioactive, Central University of Las Villas, Santa Clara 54830, Villa Clara, Cuba*
[b]*Department of Pharmacy and Pharmaceutical Technology, University of Valencia, Burjassot 46100, Valencia, Spain*

Received 9 June 2004; revised 14 August 2004; accepted 26 August 2004
Available online 17 September 2004

Abstract—The central nervous system (CNS) activity has been investigated by using a topological substructural molecular approach (TOPS–MODE). A discriminant analysis to classify CNS and non-CNS drugs was developed on a data set (302 compounds) of great structural variability where more than 81% (247/302) were well classified. Randic's orthogonalization procedures was carried out to allow the interpretation of the model and to avoid the collinearity among descriptors.

The discriminant model was assessed by a *leave-n-out* (when *n* varies from 2 to 20) cross-validation procedure (79.94% of correct classification), an external prediction set composed by 78 CNS/non-CNS drugs (80.77% of correct classification) and a 5-fold *full* cross-validation (removing 78 compounds in each cycle, 80.00% of good classification). With this methodology was demonstrated that the hydrophobicity increase the CNS activity, while the dipole moment and the polar surface area decrease it; evidencing the capacity of the TOPS–MODE descriptors to estimate CNS activity for new drug candidates. The structural contributions to the CNS activity for two compounds are presented on the basis of fragment contributions. The model has also been able to identify potential structural pharmacophore, showing its possibilities in the lead generation and optimization processes.

1. Introduction

The drug capacity to penetrate the blood–brain barrier (BBB) has a great importance in the drug discovery process. Compounds with central nervous system (CNS) activity need a high blood–brain penetration, while negligible values are desirable for drugs with a peripheral site of action.[1] The BBB is the specialized system of capillary endothelial cells that protects the brain from harmful substances, limiting strictly the transport into the brain through both physical (tight junctions) and metabolic (enzymes) barriers. Thus, the BBB is often the rate-limiting factor in determining permeation of therapeutics drugs into the brain.[2,3]

Compounds with receptors located in the central nervous system, such as GABA[4] and NMDA receptor complex,[5] should readily permeate across the BBB. By contrast, compounds acting on peripheral receptors or enzymes should have low to zero CNS penetration in order to avoid potential side effects.[2]

In recent years, several experimental and theoretical approaches have been reported to measure[6] and/or estimate the BBB penetration.[2,3,7–22] Early prediction of log BB using various physicochemical parameters such as the octanol–water partition coefficient,[23] molecular size descriptors[17,24] and solvation parameters[9–11,19] have been carried out. Other approaches using lipophilicity, polarity, polarizability, polar surface area and hydrogen bonding parameters have also been used.[8,18,21]

Another way to predict the BBB penetration has been by classification methods that differentiate CNS from CNS inactive compounds.[25–27] Different theoretical models to predict the CNS activity, using topological indices, hydrogen bonding and hydrophobic descriptors have been reported.[2,14,18,28–30]

Keywords: TOPS–MODE approach; CNS prediction; In silico; Topological descriptors.
* Corresponding author. Tel.: +53 42 281473/281192; fax: +53 42 281130; e-mail addresses: macabrera@cbq.uclv.edu.cu; migue@gammu.com

0968-0896/$ - see front matter © 2004 Published by Elsevier Ltd.
doi:10.1016/j.bmc.2004.08.038

From the topological representation the graph-theoretical methods have become one of the most important tools for quantifying molecular structure. The TOPS–MODE approach[31] based on the calculation of the spectral moments of the bond matrix of molecular graph has been used to generate graph-theoretical descriptors, expressing physical and biological properties in terms of substructural features of molecules. This approach has been successfully applied to different QSPR and QSAR studies.[29,30,32–39]

Taking into consideration the above mentioned, the aims of the present paper were: to use the TOPS–MODE approach in the generation of a discriminant function by a Linear Discriminant Analysis (LDA) that permits the classification of drugs (CNS and non-CNS compounds) on data set drawn from literature; to assess the 'in silico' model by the use of different validation tests. Finally, to evaluate the positive and/or negative fragment contributions, on the CNS activity, in order to identify potential pharmacophores, as an important tool in the drug design field.

2. The TOPS–MODE approach

The TOPS–MODE approach is based on the calculation of the spectral moments of the so-called bond matrix,[39] whose theoretical basis has been described in previous reports.[29,30,32–34] Nevertheless, an overview of this approach will be given below.

The bond matrix is defined as a square and symmetric matrix whose entries are ones or zeros if the corresponding bonds are adjacent or not. The order of this matrix (m) is the number of bonds in the molecular graph, being two bonds adjacent if they are incident to a common atom. The spectral moments of the edge adjacency matrix are defined as the traces, that is, the sum of the main diagonal, of the different powers of such a matrix.

In order to apply the present approach to the structure–property relationship, the following steps should be followed. First, to select an adequate training set with great structural diversity. Second, to draw the hydrogen depleted molecular graphs for each molecule of the training set. The third step is to differentiate the molecular bonds with appropriate weights. The fourth, to compute the spectral moments of the bond matrix for each molecule of the data set. Fifth, to find a qualitative structure–property relationship by using a discriminant analysis:

$$P = a_0\mu_0 + a_1\mu_1 + a_2\mu_2 + \cdots + a_k\mu_k + b \qquad (1)$$

where P is the studied property, in our case, the CNS activity, μ_k is the kth spectral moment and the a'_ks are the coefficients obtained by the discriminant equation. Sixth, to test the predictive capacity of the model by cross-validation procedures and an external prediction set. Finally, to compute the contribution of the different bonds in order to determine their quantitative contribution to the CNS activity of the molecules studied.

3. Computational strategies and data set

The current selection of bond weight for calculating TOPS–MODE descriptors was carried out through accounting for hydrophobic/polarity, electronic and steric features of molecules. Thus, atomic contributions for partition coefficient (H),[40] polar surface area (PSA),[41] molar refraction (MR),[42] Gasteiger–Marsilli atomic charge (Ch),[43] polarizability (Pol),[44] van der Waals atomic radii (vdW)[45] and atomic mass (AM) are transformed into bond contributions. The way in which these atomic contributions were transformed into bond contributions have been described by Estrada et al.[46]

$$w(i,j) = \frac{w_i}{\delta_i} + \frac{w_j}{\delta_j} \qquad (2)$$

where w_i and δ_i are the atomic weight and vertex degree of the atom i. Also the bond dipole moment (Dp) was used as bond weight. The calculation of the TOPS–MODE descriptors was carried out with the computer software MODESLAB 1.0.[47] The input of the chemical structures into the MODESLAB software was as simplified molecular input line entry specification (SMILES).[48] We calculated the first 10 spectral moments (μ_1–μ_{10}) for each bond weight and the number of bonds in the molecules (μ_0).

A data set of 380 compounds, with great molecular diversity, was carefully selected from literature.[2,49] This data was composed by 189 drugs with CNS activity and the rest of compounds (191 drugs) with lack of central pharmacological activity. The structural diversity of this data set was evidenced through a hierarchical cluster analysis of the active and inactive sets of compounds. Also this procedure permit to select compounds for the training and test sets, in a representative way, in all level of the linking distance. Two dendrograms, using the Euclidean distance (X-axis) and the complete linkage (Y-axis), illustrating the results of the cluster analysis developed. See Figures 1 and 2.

As can be seen, there is a great number of different subsets, which prove the molecular variability of the selected compounds. The data set was randomly divided in two subsets, the training set containing 302 drugs (150 active and 152 inactive compounds) and the external prediction set containing 78 compounds (39 active and 39 inactive compounds). The compounds belonging to the external prediction set were never used in the development of the discriminant function and they were reserved to validate the discriminant model.

The discriminant function was obtained by using the stepwise Linear Discriminant Analysis (LDA) as implemented in the STATISTICA version 5.5.[50] The default parameters of this program were used in the development of the model. The variables to be included in the equation were selected using a forward stepwise procedure as a variable selection strategy.

The quality of the model was determined by examining the Wilks λ, the Mahalanobis distance (D^2), the Fisher

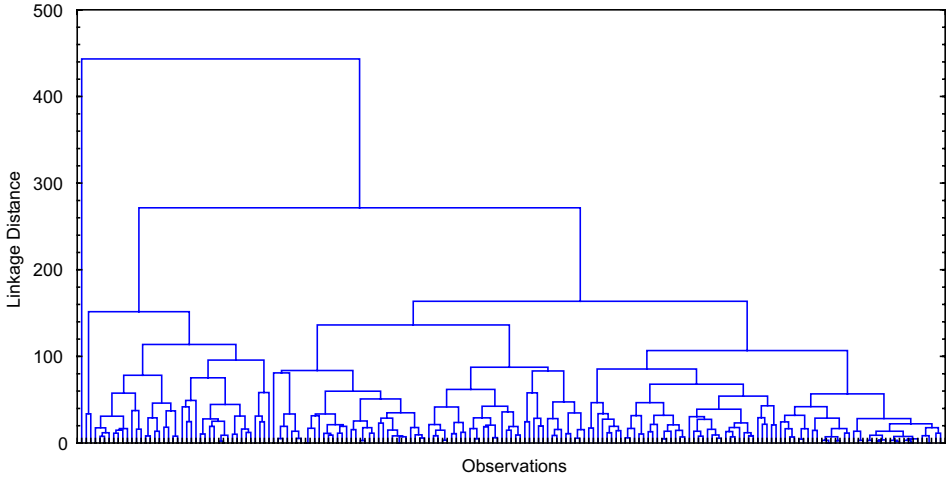

Figure 1. A dendrogram illustrating the results of the hierarchical cluster analysis of the set of active compounds used in the training and prediction set of the present work.

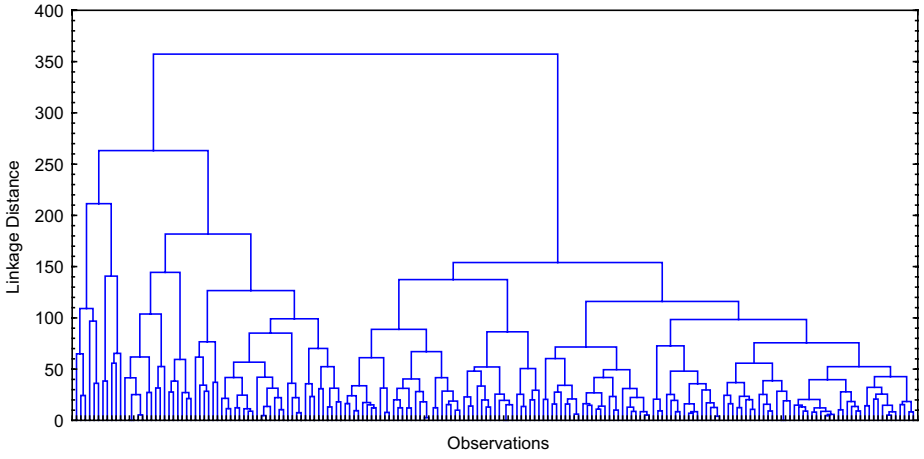

Figure 2. A dendrogram illustrating the results of the hierarchical cluster analysis of the set of inactive compounds used in the training and prediction set of the present work.

ratio (F) and the corresponding p-level ($p(F)$) as well as the percentage of good classification and the proportion between the cases and variables in the equation. The Wilks λ for the overall discrimination can take values in the range of 0 (perfect discrimination) to 1 (no discrimination). The D^2 statistics indicates the separation of the respective groups, showing if the model possesses an appropriate discriminatory power for differentiating between the respective two groups. Compounds were considered unclassified by the model (NC) when the differences in the percentage of classification between groups do not differ in more than 5%. The validation of the model was carried out by a *leave-n-out* cross-validation procedures.[50] Also a *full* cross-validation test of the model was investigated. From the general data set, groups of 78 observations were randomly selected five times. Each group was left out (*leave-group-out*, LGO) and that group predicted by the model developed from the remaining observations. In this way, every observation was left out once and its value predicted. The significant criterion for assessing model quality was the Wilks λ, the Fisher ratio (F) and the percentage of correct classification.

4. Computation of fragment contributions

The computation of fragment contributions to the toxicological property under study is probably the most important advance of the TOPS–MODE approach to the study of toxicological variables compared to the traditional QSAR and QSPR methods. This procedure can be useful for the identification of structural alerts in the study of carcinogenic prediction. The computation of fragment contribution has been described in more details in previous reports.[51–54] The procedure consists of calculating the spectral moments for all the fragments contained in a given substructure, and by the difference of these moments we obtain the contribution of the substructure. The general algorithm in this computational

approach is as follows. First, we select the substructure whose contribution to the moments we would like to determine. Then we generate all the fragments (subgraphs), which are contained in the corresponding substructure, and calculate the spectral moments for both, the substructure and all their fragments. The contribution of the substructure to the spectral moments is finally obtained as the difference between the spectral moments of the substructure and all those from their fragments. Having the contributions of the different structural fragments of interest, we only need to substitute these contributions into the quantitative model developed to describe the property studied, for example, model (Eq. 1) and we obtain the quantitative contribution of the different fragments to P.

5. Results and discussion

The classification function to discriminate CNS from CNS inactivity of the compounds belonging to the training set is given below together with the statistical parameters of the LDA:

$$\begin{aligned}\text{class} = &-0.319 + (1.920 \times \mu_1^{\mathrm{H}}) + (0.488 \times \mu_2^{\mathrm{H}}) \\ &- (0.267 \times \mu_3^{\mathrm{H}}) - (0.047 \times \mu_1^{\mathrm{PSA}}) \\ &- (1.381 \times \mu_1^{\mathrm{Dp}}) - (0.454 \times \mu_2^{\mathrm{Dp}}) + (0.242 \times \mu_3^{\mathrm{Dp}}), \\ &N = 302,\ \lambda = 0.577,\ D^2 = 2.92, \\ &F_{\exp}(7, 294) = 30.67,\ p < 0.001 \end{aligned} \quad (3)$$

The classification results for the training set are illustrated in Table 1.

The model classified correctly the 83.33% of drugs with CNS activity in the training set and the 80.26% of inactive compounds, for a global good classification of 81.79%. If the unclassified compounds are considered the percentages of good classification will be 86.81% and 84.14% for active and inactive CNS drugs. The percentages of *false positives* and *false negatives* in the training set were 8.61% (26/302) and 7.28% (22/302), respectively. *False positives* are those compounds without CNS activity that are classified as active, and the *false negatives* are those compounds with CNS activity and the model classified it as inactive (see Table 1). From a practical point of view, in the development of the classification model, is considered more important avoiding *false negatives* because those are compound that will be rejected for their wrong predicted activity and therefore they will never be evaluated experimentally, and their true activity would never be discovered. On the contrary, the *false positive* compounds eventually will be detected.

On the other hand, it has been largely known that spectral moments of higher order are mathematically co-linear. For this reason, in the present study, the Randic's orthogonalization procedure was carried out[55–58] in order to avoid the collinearity among different variables, which makes impossible to know the relative importance of each variable and can produce over-fitting results.

The Randic method of orthogonalization has been described in details in several publications.[55–58] Thus, we will give a general overview here. The first step in this procedure is to select the appropriated order of orthogonalization, which in this case is the order in which the variables were selected in the forward stepwise procedure of the LDA.

The new orthogonalized variable will have the following nomenclature ${}^{m}\mathrm{O}(\mu_k^w)$, where the symbol O means orthogonal, m is the degree of importance of the descriptor to explain the property determined by the order in which it is selected by forward stepwise analysis, μ_k is the kth spectral moment and w is the bond weight used.

In this sense, in Eq. 4 we used ${}^{1}\mathrm{O}(\mu_1^{\mathrm{H}})$ as the first orthogonal variable. Afterwards, the successive residuals of the step-by-step regressions between each variable selected in the model and the others in order of statistical significance were calculated.[59] All these residuals were used as the remnant orthogonal variables in the Eq. 4.[59] In this analysis the least squares method selected all orthogonal analogs of collinear variables. It ensured us that, in spite of variables collinearity, each variable have an amount of information not encoded in the others.[57,59,60]

The results of the Randic's orthogonalization procedure are depicted in Table 2.

As a result of the orthogonalization of the Eq. 3 we obtain the following equation:

$$\begin{aligned}\text{class} = &\ 0.091 + 1.196\times{}^{1}\mathrm{O}(\mu_1^{\mathrm{H}}) + 1.082\times{}^{6}\mathrm{O}(\mu_2^{\mathrm{H}}) \\ &- 1.482\times{}^{4}\mathrm{O}(\mu_3^{\mathrm{H}}) - 2.073\times{}^{3}\mathrm{O}(\mu_1^{\mathrm{PSA}}) \\ &- 3.852\times{}^{5}\mathrm{O}(\mu_1^{\mathrm{Dp}}) - 11.332\times{}^{7}\mathrm{O}(\mu_2^{\mathrm{Dp}}) \\ &+ 0.781\times{}^{2}\mathrm{O}(\mu_3^{\mathrm{Dp}}) \end{aligned} \quad (4)$$

It must be highlighted here that the orthogonal descriptor-based model coincides with the collinear spectral moment-based model in all statistical parameters, as well as in the classification matrix (see Table 2).

The variables in the model (Eq. 4) encoded specific structure information. The first three variables are related to the hydrophobicity or partition coefficient and have, in general way, a positive contribution to the CNS activity. It is known that the capacity to cross the biological membrane drops at both low and high $\log P$ values. The main reasons of these are due to (1) the inability of weakly lipophilic compounds to penetrate the lipid portion of the membrane and (2) the excessive partitioning of strongly lipophilic compounds into the lipid portion of the membrane and their subsequent inability to pass through the aqueous portion of the membrane.[61] However, until excessive partition values are not reached, increments of $\log P$ will increase the

Table 1. Results of the classification of the training set compounds

No	Compounds	Class	Prob.
Training active set			
1	Acecarbromal	+	0.96
2	Acetylpheneturide	+	0.93
3	Allobarbital	+	0.92
4	Amobarbital	+	0.89
5	Annansiol	+	0.68
6	Antiepilepsirium	+	0.90
7	Apomorphine	–	0.80
8	Aponal	+	0.70
9	Atolide	+	0.87
10	Atrolactamide	–	0.56
11	Baldrianol	–	0.61
12	Bason	+	0.72
13	Benzamyl	+	0.96
14	Benzobarbital	+	0.96
15	Biperidine	+	0.95
16	Brallobarbital	+	0.92
17	Bromisoval	NC	0.51
18	Bomoaprobarbital	+	0.88
19	Bomophenazone	+	0.78
20	Brosuximide	+	0.79
21	Buramate	–	0.78
22	Butabarbital	+	0.94
23	Butallylonal	+	0.93
24	Butobarbital	+	0.94
25	Canadine	+	0.66
26	Carbocloral	+	0.91
27	Carbromal	+	0.84
28	Carburazepam	+	0.95
29	Cenestil	+	0.54
30	Cetohexazine	+	0.59
31	Cetohexazine	+	0.58
32	Chloral-hydrate	NC	0.52
33	Chlordiazepoxide	+	0.95
34	Chlorprothixene	NC	0.51
35	Cinolazepam	+	0.76
36	Cipenam	+	0.98
37	Cloperidone	+	0.98
38	Clotiazepam	+	0.63
39	Cyclobarbital	+	0.90
40	Cyheptamide	NC	0.49
41	Doxenitoin	+	0.91
42	Doxylamine	+	0.85
43	Estazolam	+	0.58
44	Etaqualone	+	0.78
45	Ethadione	+	0.91
46	Ethanion	+	0.86
47	Ethyl-loflazepate	+	0.96
48	Etomidate	+	0.78
49	Femerazo	+	0.69
50	Fenadiazole	–	0.91
51	Fentanyl	+	0.99
52	Fepiron	+	0.74
53	Fletazepam	+	0.97
54	Flufenisal	–	0.64
55	Flupentixol	+	0.82
56	Flurazepam	+	0.99
57	Haloperidol	+	0.89
58	Haloxazolam	+	0.98
59	Heptobarbital	+	0.79
60	Hexetal	+	0.93
61	ICI45337	+	0.90
62	IL-40	+	0.97
63	Isopral	+	0.77
64	Ketamine	+	0.82
65	Lamotane-x	–	0.75
66	Lorazepam	+	0.76
67	L-Tryptophan	–	0.85
68	Mecap	+	0.77
69	Menthoval	+	0.87
70	Meprobamate	–	0.86
71	Meprophendiol	–	0.91
72	Mequitazine	+	0.91
73	Mesuximide	+	0.80
74	Methabarbital	+	0.94
75	Metalonine	–	0.64
76	Metharbital	+	0.94
77	Metetoin	+	0.98
78	Methadone	+	0.93
79	Methamphidon	+	0.93
80	Methaqualone	+	0.67
81	Methohexital	+	0.96
82	Methylphenitoine	+	0.97
83	Metindion	+	0.93
84	Metomidate	NC	0.50
85	Mexitilene	–	0.62
86	MK-801	+	0.82
87	Naltrexone	+	0.89
88	Neobonyval	+	0.58
89	Niaprazine	+	0.89
90	Nirvanol	+	0.95
91	Nitrazepam	–	0.57
92	Nonapyrimine	+	0.61
93	Norhexobarbital	+	0.77
94	Nortriptyline	–	0.61
95	Orphenadrine	+	0.87
96	Oxcarbazepine	+	0.72
97	Oxitriptiline	+	0.85
98	Paralcetaldehide	+	0.89
99	Paramethadione	+	0.93
100	Pentenal	+	0.88
101	Pentobarbital	+	0.95
102	Perlapine	+	0.82
103	Perphenazine	–	0.87
104	Pethrichloral	+	0.99
105	Petrichloral	–	0.55
106	Phenacemide	+	0.74
107	Pheneturide	+	0.90
108	Phenythilone	+	0.83
109	Phenytoin	+	0.95
110	Piperine	+	0.92
111	Prazepam	+	0.91
112	Primidone	+	0.81
113	Progabide	–	0.79
114	Promethazine	+	0.90
115	Propanolol	NC	0.48
116	Propylbarbital	+	0.94
117	Quisqualamine	–	0.70
118	Remacemide	+	0.61
119	RO5-5807	+	0.88
120	Roletamide	+	0.53
121	Roxindol	+	0.71
122	Secobarbital	+	0.96
123	SL75102	–	0.62
124	Somnamid	+	0.70
125	SQ-10996	+	0.71
126	Stiripentol	–	0.75
127	Sulpiride	+	0.66
128	Sumatriptan	–	0.79
129	Suriclone	+	0.84
130	Talbutal	+	0.96

(*continued on next page*)

Table 1 (*continued*)

No	Compounds	Class	Prob.
131	Tamitinol	–	0.92
132	Temazepam	+	0.85
133	Tetrantoin	+	0.87
134	Thalidomide	+	0.87
135	Thiazolam	+	0.58
136	Thiopental	+	0.95
137	Thiopental sodium	+	0.97
138	Toletamine	+	0.79
139	Triazolam	+	0.58
140	Tricetamide	+	0.59
141	Trichloroisobutylsalicylate	+	0.87
142	Valbormine	+	0.64
143	Valeric acid	+	0.54
144	Valerylphenetidine	+	0.94
145	Valnocta mide	+	0.75
146	Valperinol	+	0.97
147	Valpromide	+	0.77
148	Zapizolam	+	0.79
149	Zimelidine	+	0.55
150	Zopiclone	+	0.89
Training inactive set			
151	Abbot29590	–	0.67
152	Acedoben	–	0.54
153	Acrivastine	+	0.68
154	Alarmine	–	0.71
155	Albuterol (salbutamol)	–	0.99
156	Alprenolol	+	0.70
157	Amantadine	–	0.81
158	Aminopicoline	–	0.87
159	Aminothiazole	–	0.96
160	Ammoidin	–	0.77
161	Amrinone	–	0.88
162	Antienite	+	0.59
163	APPA	–	0.96
164	Atenolol	–	0.80
165	Azaconazole	+	0.85
166	BA1	–	0.63
167	Baclofen	–	0.91
168	Betaxolol	–	0.67
169	Bicordin	–	0.80
170	Bitoscanate	–	0.90
171	Bromebric acid	–	0.69
172	bromthymol	–	0.54
173	Bufurarol	–	0.77
174	Camfazolinum	–	0.86
175	Captopril	–	0.57
176	Carbodine	–	0.98
177	Carmoxirol	+	0.70
178	Cefetamet pivoxil	–	0.95
179	Cefotiam	–	0.98
180	Cephalotin	–	0.57
181	Certuna	–	0.74
182	Chloramphenicol	–	0.97
183	Chlorthenoxazine	+	0.72
184	Ciclobendazol	+	0.74
185	Clofibric acid	+	0.59
186	Coumarin	–	0.54
187	Cyacetacide	–	0.91
188	Demeclocycline	–	0.94
189	Depreton	–	0.99
190	Desmehtylmethalibur	–	0.95
191	Dexamethasone	–	0.78
192	Dihydralazine	–	0.99
193	Dimepranol	–	0.72
194	Dobutamine	–	0.91
195	Enoxinone	NC	0.49

Table 1 (*continued*)

No	Compounds	Class	Prob.
196	Ethinylestradiol	–	0.60
197	Etofilline	–	0.98
198	Felodipine	+	0.89
199	Fenamole	–	0.77
200	Fenoprofen	+	0.73
201	Fluoxidin	–	0.92
202	Formycin A	–	0.98
203	Fumaric acid	–	0.88
204	Fumigatin	–	0.76
205	Furidanone	–	0.73
206	Glutarienol-calcium	–	0.95
207	Glycarbylamide	–	0.94
208	Guanabenz	–	0.95
209	Guanamine	–	0.98
210	Guanfacine	–	0.86
211	Heptaminol	–	0.81
212	Hydralazine	–	0.96
213	IMPY	–	0.61
214	Isobromin	–	0.99
215	Isoverin	–	0.81
216	Keprotofen	+	0.40
217	Labetalol	–	0.91
218	Lofemizole	NC	0.52
219	Lofemizolehydrochloride	NC	0.52
220	Lysepsina	–	0.72
221	Mannitol	–	0.99
222	Manozodil	–	0.93
223	Mercaptosuccinic acid	–	0.97
224	Methyldopa	–	0.94
225	Metirosine	–	0.84
226	Metoprine	–	0.97
227	Metoprolol	–	0.79
228	Mexiletine	–	0.62
229	Mg ferulate	–	0.90
230	Mibefadril	+	0.04
231	Milrinone	–	0.70
232	MJ10459-2	NC	0.52
233	Morinamide	–	0.67
234	Moroxydine	–	0.97
235	Mycosid	–	0.67
236	Nadolol	–	0.95
237	Naproxen	–	0.74
238	Natrii fenbutyras	+	0.27
239	Nicotinic acid	–	0.82
240	Nitrendipine	NC	0.48
241	Norfenefrin	–	0.98
242	Novenbitol	–	0.57
243	Noxitiolin	–	0.98
244	Olsalazinehydrochloride	–	0.99
245	Omnastine	–	0.87
246	Orazamide	–	0.96
247	Oxdralazine	–	0.99
248	Oxprenolol	+	0.27
249	Pamabrom	–	0.94
250	Pathocidin	–	0.78
251	Patulin	–	0.63
252	Penbutolol	–	0.54
253	Phenacetin	+	0.19
254	Pheniranine	+	0.22
255	Phentermine	–	0.56
256	Piperazine	–	0.75
257	Planomycin	–	0.93
258	Praxadine	–	0.89
259	Prednisolone	–	0.92
260	Prednisone	–	0.65
261	Procodazole	+	0.32

Table 1 (*continued*)

No	Compounds	Class	Prob.
262	Progabide	–	0.65
263	Progalli-p	–	0.87
264	Proscilladirin	–	0.67
265	Protheobromine	–	0.93
266	Protiofate	–	0.72
267	Pyracinamide	–	0.91
268	Pyridostigmine	+	0.33
269	Quazodine	–	0.76
270	R8231	+	0.14
271	Redimyl	–	0.56
272	Regutensin	+	0.28
273	Remikiren	+	0.09
274	Ribavirin	–	0.97
275	Riodoxol	–	0.64
276	Salsolidine	–	0.83
277	Selectan	NC	0.48
278	Spinulosin	–	0.89
279	Sulfasalazine	–	0.91
280	Tenuazonic acid	+	0.29
281	Terbutaline	–	0.97
282	Tetramizolehydrochloride	+	0.13
283	Tetridamine	–	0.77
284	Thiaguanosine	–	0.99
285	Thienamicyn	–	0.93
286	Tiacrilast	–	0.90
287	Tianafac	–	0.92
288	Tioguanine	–	0.99
289	Tioxolone	–	0.96
290	Tizoprolic acid	–	0.86
291	Tocainide	–	0.53
292	Tolamolol	–	0.82
293	Toliprolol	–	0.56
294	Toxopyrimidine	–	0.99
295	Trapidil	+	0.34
296	Trichofytocid	–	0.67
297	Triclacetamol	+	0.14
298	Vitacampher	–	0.74
299	VUFB-7904	–	0.99
300	Warfarin	NC	0.49
301	Wormin	–	0.65
302	Xanthine	+	0.47

+: Positive values are for compounds with CNS activity. –: Negative values are for compounds with CNS inactivity. NC: non-classified drugs.

molecule capacity to cross the biological membrane. Other variable $^{3}O(\mu_1^{PSA})$, with negative contributions to the CNS activity, is related with the molecular polar surface area. This is a logical result due to when the hydrogen bonding capacity of a molecule is increased; the possibility to pass the biological membrane is lower because of the affinity of compound for the aqueous layer of the membrane.[62] On the other hand, the spectral moments weighted with the bond dipole moment (the last three variables of Eq. 4) have a negative contribution to the CNS activity. Once the dipole moment is increased, greater polarity of the molecules is evidenced; with a negative effect on membrane permeability.[63]

For a more exhaustive testing of the predictive power of the model, we carried out a *leave-n-out* cross-validation procedure. This validation technique is implemented in the module for classification trees training in STATISTICA.[50] In this module, the user can select discriminant-based linear combination splits as the split method, prune on misclassification error as the stopping rule and the same prior probabilities as in Eq. 4 and obtains this equation as the split rule. Once Eq. 4 is modeled in the classification trees' module the folding parameter of the cross-validation can be varied to carry out the *leave-n-out* procedure. This model shows a range between 76.5% and 81.5% of global good classification when n varies from 2 to 20 in leave-n-by time cross-validation procedures. The model seems to be stabilized at around 80% of good classification when n is >8.

The most important criterion for the quality of the discriminant model is based on the statistics for the external prediction set. Eq. 4 classified correctly the 87.18% and 74.36% of active and inactive compounds, respectively. The global classification was 80.77%. If unclassified compounds are considered, the percentage of good classification for the CNS active compounds is 91.89% and 80.56% for inactive compounds. The global classification was 86.30%. The percentage of *false negative* and *false positive* compounds was 6.41% (5/78) and 12.82% (10/78), respectively. In Table 3 the classification of compounds in the external prediction set is presented.

Finally, the reliability of the model was tested by a 5-fold *full* cross-validation test. For each group of observations left out (~20% of the whole data set, 78 compounds), a model was developed from the remaining 80% of the data (302 compounds). This process was carried out five times on five unique subsets. The statistical results are depicted in Table 4.

The result of predictions on the *full* cross-validation test evidenced the quality of the obtained model.

The more important advantage of the TOPS–MODE approach to molecular design, respect to other 'in silico' methodologies, is the possibility of obtaining the quantitative contribution of any fragment to the studied property. Several 'in silico' methods to predict the CNS activity have been developed but the specific effect of the fragments on the CNS activity has not been evaluated.[1–3,7–12,14–21] In order to demonstrate the importance of this methodology in the identification of potential pharmacophores, on the chemical structures, and their further role in the design of potent drugs with action on the central nervous system, several structural fragments were evaluated. As can be seen in Figure 3 and analyzing the fragments F1, F2 and F3 we can appreciated that the phenyl ring has the highest positive contribution to the CNS activity and it is explained due to in F2 there is an increment of the polarity of the fragment, decreasing the hydrophobicity and the capacity to cross the biological barriers; in the case of F3 the substituent increase the polar surface area and the polarity with a more deeper negative effect on the hydrophobicity. On the other hand when the phenyl ring is changed by the pyridine fragment (F6) there is a negative effect on the CNS active, being the main reason the change in the polarity and polar surface area of this fragment.

Table 2. Results of Randic's orthogonalization analysis

Non-orthogonal spectral moments

	μ_1^{Dp}	μ_2^{Dp}	μ_3^{Dp}	μ_1^{H}	μ_2^{H}	μ_3^{H}	μ_1^{PSA}
μ_1^{Dp}	1.00	0.65	0.92	−0.07	0.47	0.36	0.57
μ_2^{Dp}	0.65	1.00	0.83	0.33	0.97	0.84	0.37
μ_3^{Dp}	0.92	0.83	1.00	0.03	0.69	0.62	0.51
μ_1^{H}	−0.07	0.33	0.03	1.00	0.43	0.64	−0.51
μ_2^{H}	0.47	0.97	0.69	0.43	1.00	0.88	0.27
μ_3^{H}	0.36	0.84	0.62	0.64	0.88	1.00	0.00
μ_1^{PSA}	0.57	0.37	0.51	−0.51	0.27	0.00	1.00

Orthogonal spectral moments

	$^5O(\mu_1^{Dp})$	$^7O(\mu_2^{Dp})$	$^2O(\mu_3^{Dp})$	$^1O(\mu_1^{H})$	$^6O(\mu_2^{H})$	$^4O(\mu_3^{H})$	$^3O(\mu_1^{PSA})$
$^5O(\mu_1^{Dp})$	1.00	0.00	0.00	0.00	0.00	0.00	0.00
$^7O(\mu_2^{Dp})$	0.00	1.00	0.00	0.00	0.00	0.00	0.00
$^2O(\mu_3^{Dp})$	0.00	0.00	1.00	0.00	0.00	0.00	0.00
$^1O(\mu_1^{H})$	0.00	0.00	0.00	1.00	0.00	0.00	0.00
$^6O(\mu_2^{H})$	0.00	0.00	0.00	0.00	1.00	0.00	0.00
$^4O(\mu_3^{H})$	0.00	0.00	0.00	0.00	0.00	1.00	0.00
$^3O(\mu_1^{PSA})$	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Forward stepwise LDA with orthogonal spectral moments

$^3O(\mu_1^{PSA})$	$^1O(\mu_1^{H})$	$^2O(\mu_3^{Dp})$	$^5O(\mu_1^{Dp})$	$^7O(\mu_2^{Dp})$	$^4O(\mu_3^{H})$	$^6O(\mu_2^{H})$	Const	λ	P	% T	% +	% −
−2.073	1.196	0.781	−3.852	−11.333	−1.482	1.082	0.091	0.577	0.00	81.79	83.33	80.26
−2.052	1.170	0.738	−3.806	−11.137	−1.433		0.098	0.590	0.00	79.80	80.00	79.60
−1.962	1.116	0.664	−3.676	−10.342			0.075	0.627	0.00	76.82	81.33	72.37
−1.854	1.068	0.647	−3.401				0.080	0.664	0.00	76.49	81.33	71.71
−1.746	0.993	0.601					0.073	0.707	0.00	75.83	80.67	71.05
−1.644	0.938						0.059	0.754	0.00	71.85	82.67	61.18
−1.403							−0.008	0.858	0.00	65.56	74.00	57.24

Also if the electronic charge of the fragment is decreasing, the fragment will have a better contribution to the CNS activity (see fragment F6 and F13).

The effect of this fragment on real drugs can be appreciated in two compounds not used in the present study to obtain the discriminant model.

For example the diazepam, a benzodiazepine drug, has shown to enhance the actions of GABA at the $GABA_A$ receptor by increasing the frequency of Cl-channel opening with little effect on either the opening time or its conductance.[64] Some fragments belongs to this drug appears in Figure 3.

As can be seen in this figure, the seven member imino ring (F8) has a positive contribution to the CNS activity. This situation coincides with the affinity towards the BZ-binding site reported for this drug.[65] Also the phenyl ring has a positive contribution to the CNS activity. This result is agreed with reported by other authors where these positions are effective location in these molecules for enhancing the affinity of the compound for the BZ-binding site.[66–68] The phenyl ring attached to the imino ring increase the CNS activity, due to this fragment increasing the lipophilicity and the electronic charge at this ring and directly increase the affinity of the ligand for the binding site (see also the contribution of fragment F4).

Additionally, the carbonyl group for diazepam has also been shown to substantially contribute to the binding affinity of these compounds (see the contribution of F7 and F8 fragments). This fact could be explained if we consider the effect of the polarity and the dipole moment on this bond, which could produce the high affinity of the carbonyl group of this compound for the BZ-binding site.

The other compound is the Aripiprazole, a dopamine D2 receptor partial agonist, with partial agonist activity at serotonin 5HT1A receptors and antagonist activity at 5HT2A receptors.[69] This drug has four fragments with positive contribution to the CNS activity, F1, F2, F9 and F10.

The presence of different fragments with positive contribution does not presuppose the development of CNS activity per se, because it is well known that the activity is the consequence of the sum of the contribution of all fragments in the molecule. However, the identification of fragment contributions can be of interest to explain the action mechanism of different drugs with potent CNS activity.

As may be seen, the TOPS–MODE approach not only permits the correct classification of heterogeneous data set according to the CNS activity, but it also facilitates the determination of fragment contributions, which are important to obtain adequate values of this biological

Table 3. Results of the classification of the test set compounds

No	Compounds	Class	Prob.
Test active set			
1	Albutoin	+	0.78
2	Alozafone	+	0.87
3	Amilene hydrate	NC	
4	Amylurea	–	0.57
5	Aphendion	+	0.96
6	Aprobarbital	+	0.92
7	Bagrosin	+	0.78
8	Beclamide	+	0.60
9	Benzodihinehydrochlorid	+	0.88
10	Bromazepam	+	0.76
11	Bromobutanol	+	0.76
12	Brotizolam	+	0.76
13	Butalbital	+	0.90
14	Butoctamide	+	0.79
15	Carbamazepine	+	0.71
16	Carbromide	+	0.80
17	Centazolone	–	0.75
18	Chinoin-1045	+	0.86
19	Chlorphenacemid	NC	0.52
20	Chlozepate	+	0.84
21	Crotarbital	+	0.94
22	Denzimol Hydrochloride	–	0.58
23	Dioxamate	+	0.93
24	Diphoxazide	+	0.86
25	Ectylurea	+	0.58
26	Etchlorovynol	+	0.74
27	Ethotoin	+	0.96
28	Felbamate	–	0.83
29	Fenobam	+	0.92
30	Febarbamate	+	0.96
31	Fluoxetine	+	0.92
32	Glutethimide	+	0.86
33	Hedonal	+	0.54
34	Hexobarbital	+	0.84
35	Ibrotamide	+	0.79
36	Isoladol	–	0.90
37	Ketanime	+	0.82
38	Lormetazepam	+	0.84
39	Mefobarbital	+	0.91
Test inactive set			
40	Acedoben	–	0.63
41	Aciclovir	–	0.98
42	Ag-307	–	0.99
43	Alclofenac	+	0.57
44	Aminodal	–	0.91
45	Amiquinsin	–	0.92
46	Antibrucellin	–	0.71
47	Astemizole	+	0.86
48	Azaserine	–	0.95
49	Benserazide	–	0.99
50	Bijosal	–	0.78
51	Bromotheamin	NC	0.52
52	BVDU	–	0.61
53	Carbenicillin	+	0.91
54	Cefamandole	–	0.81
55	Cephalexin	+	0.64
56	Chloranil	+	0.75
57	Citenazon	–	0.99
58	Colestipol	–	0.99
59	Cutison	–	0.94
60	Debrisoquin	–	0.82
61	Desderman	–	0.59
62	Diamide	+	0.69
63	Diltiazem	+	0.91
64	Drazidox	–	0.97
65	Estradiol valerate	+	0.65
66	Eucalyptol	+	0.69
67	Fenchone	NC	0.51
68	Flucytosine	–	0.65
69	Fuberidazole	+	0.79
70	Furanomycin	–	0.69
71	Geraniol	–	0.83
72	Gobab	–	0.98
73	Guanazol	–	0.84
74	Hidrocortison	–	0.89
75	Iodothymol	–	0.59
76	Isotretinoim	NC	0.51
77	Ketoxal	–	0.91
78	LK274	–	0.97

+: Positive values are for compounds with CNS activity. –: Negative values are for compounds with CNS inactivity. NC: non-classified drugs.

Table 4. Results of the 5-fold *full* cross-validation procedure

	% Class (training set)	λ	F	% Class (test set)
1	82.11	0.558	33.26	71.79
2	80.46	0.597	28.37	79.48
3	80.79	0.594	28.62	83.33
4	82.12	0.578	30.62	78.20
5	79.07	0.616	26.08	87.18
Mean	**80.91**	**0.588**	**29.39**	**80.00**

activity. In this sense, the present work represents a step forward in the development of a theory that permits not only the prediction of biological activities but also the easy interpretation of the results in terms of structural concepts.

6. Conclusions

The prediction of central nervous system activity has been a goal of pharmaceutical companies. For this reason, several in silico methods have been applied in order to predict this property in the early stage of drug development and some of them have become in important tools to select new drug candidates. In this study, the TOPS–MODE approach has been a successful methodology to classify CNS active/inactive compounds. The procedure has showed that a good linear discriminant equation can be obtained using the polar surface, hydrophobicity and dipole moment as weights in the diagonal entries of the bond matrix. The model developed in the current work is easily calculated and suitable for the rapid prediction of CNS activity, and the validation and cross-validation of the QSAR model support this claim. This suggests that the present method should be regarded as a choice for lead optimization programs and the high throughput screening of large databases.

On the other hand, the present methodology permits the identification and quantification of fragment contributions that are responsible of the CNS activity. This possibility plays an important role in the interpretation of

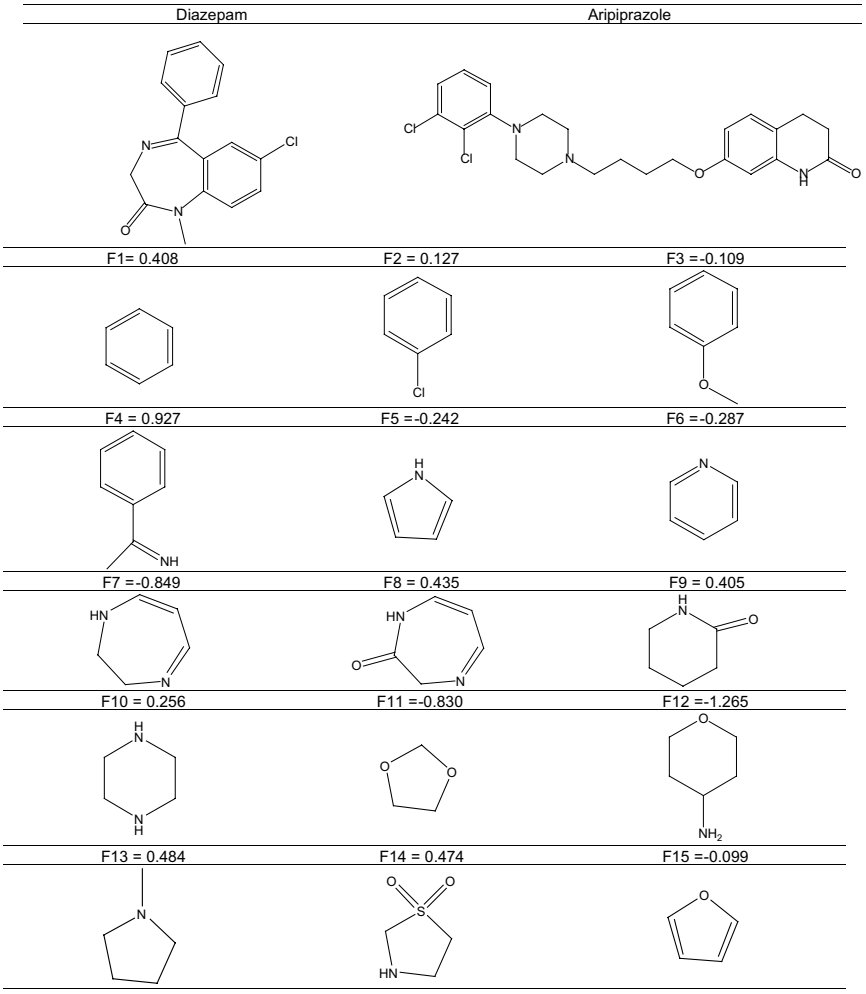

Figure 3. Results of some fragment contribution to CNS activity, according to Eq. 4.

the action mechanism for CNS drugs. Refinement of the model is possible with the incorporation of more CNS active/inactive compounds, and their power in the design of new drugs with good CNS activity could be improved with the construction and evaluation of large database.

Acknowledgements

The authors acknowledge to the owners of the Software MODESLAB 1.0 for their free contribution to the development of this work. Also, to the anonymous referees for their valuable comments in order to increase the quality of the manuscript.

References and notes

1. Norinder, U.; Haeberlein, M. *Adv. Drug. Deliver. Rev.* **2002**, *54*, 291–313.
2. van de Waterbeemd, H.; Camenish, G.; Folkers, G.; Chretien, J. R.; Raevsky, O. A. *J. Drug. Target.* **1998**, *6*, 151–165.
3. Rose, K.; Hall, L. H.; Kier, L. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 651–666.
4. N'Goka, V.; Schlewer, G.; Linget, J. M.; Chambon, J. P. Wermuth, C. G. *J. Med. Chem.* **1991**, *34*, 2547–2557.
5. Cai, S. X.; Kher, S. M.; Zhou, Z. L.; Llyin, V.; Espitia, S. A.; Tran, M.; Hawkinson, J. E.; Woodward, R. M.; Weber, E.; Keana, J. F. W. *J. Med. Chem.* **1997**, *40*, 730–738.
6. Bonate, P. L. *J. Neurosci. Methods* **1995**, *56*, 1–15.
7. Young, R. C.; Mitchell, R. C.; Brown, T. H.; Ganellin, C. R.; Griffiths, R.; Jones, M.; Rana, K. K.; Saunders, D. Smith, I. R.; Sore, N. E.; Wilks, T. J. *J. Med. Chem.* **1988**, *9*, 656–671.
8. Clark, D. E. *J. Pharm. Sci.* **1999**, *88*, 815–821.
9. Keserü, G. M.; Molnár, L. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 120–128.
10. Abraham, M. H.; Chadha, H. S.; Mitchell, R. C. *J. Pharm. Sci.* **1994**, *83*, 1257–1268.
11. Abraham, M. H.; Chadha, H. S.; Mitchell, R. C. *Drug Des. Discovery* **1995**, *13*, 123–131.
12. Abraham, M. H.; Takács-Novák, K.; Mitchell, R. C. *J. Pharm. Sci.* **1997**, *86*, 310–315.
13. Calder, J. A. D.; Ganellin, C. R. *Drug Des. Discovery* **1994**, *11*, 259–268.
14. Crivori, P.; Cruciani, G.; Carrupt, P. A.; Testa, B. *J. Med. Chem.* **2000**, *43*, 2204–2216.
15. Feher, M.; Sourial, E.; Schimidt, J. M. *Int. J. Pharm.* **2000**, *201*, 239–247.
16. Iyer, M.; Mishra, R.; Han, Y.; Hopfinger, A. *J. Pharm. Res.* **2002**, *19*, 1611–1621.

17. Kaliszan, R.; Markuszewski, M. *Int. J. Pharm.* **1996**, *145*, 9–16.
18. Kelder, J.; Grootenhuis, P. D. J.; Bayada, D. M.; Delbressine, L. P. C.; Ploemen, J. P. *Pharm. Res.* **1999**, *16*, 1514–1519.
19. Lombardo, F.; Blake, J. F.; Curatolo, W. J. *J. Med. Chem.* **1996**, *39*, 4750–4755.
20. Luco, J. M. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 396–404.
21. Norinder, U.; Sjoberg, P.; Osterberg, T. *J. Pharm. Sci.* **1998**, *87*, 952–959.
22. Lobell, M.; Molnár, L.; Keserü, G. M. *J. Pharm. Sci.* **2003**, *92*, 360–370.
23. Hansch, C.; Bjorkroth, J. P.; Leo, A. *J. Pharm. Sci.* **1987**, *76*, 387–663.
24. van Bree, J. B. M. M.; de Boer, A. G.; Danhof, M.; Ginsel, L. A.; Breimer, D. D. *J. Pharmacol. Exp. Ther.* **1988**, *247*, 1233–1239.
25. Clark, D. E. *Chem. High T. Scr.* **2001**, *4*, 477–496.
26. Engkvist, O.; Drede, P.; Rester, U. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 155–160.
27. Doniger, S.; Hofmann, T.; Yeh, J. *J. Comput. Biol.* **2002**, *9*, 849–864.
28. Basak, S. C.; Gute, B. D.; Drewes, L. R. *Pharm. Res.* **1996**, *13*, 775–778.
29. Estrada, E.; Peña, A.; García-Domenech, R. *J. Comput. Aided Mol. Des.* **1998**, *12*, 583–595.
30. Estrada, E.; Peña, A. *Bioorg. Med. Chem.* **2000**, *8*, 2755–2770.
31. Estrada, E. *SAR QSAR Environ. Res.* **2000**, *11*, 55–73.
32. Pérez, M.; Terán, M. C. *Bioorg. Med. Chem.* **2004**, *12*, 2985–2993.
33. Pérez, M.; Terán, González, H.; Cabrera, M. A.; Molina, R. *Bioorg. Med. Chem.* **2004**, *12*, 735–744.
34. Pérez, M.; Terán, M. C. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3077–3079.
35. Pérez, M.; González, H.; Molina, R.; Cabrera, M. A.; Ramos, R. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1192–1199.
36. Cabrera, M. A.; Bermejo, M.; Pérez, M.; Ramos, R. *J. Pharm. Sci.* **2004**, *93*, 1701–1717.
37. Cabrera, M. A.; González, H.; Fernández, C.; Plá-Delfina, J. M.; Bermejo, M. *Eur. J. Pharm. Biopharm.* **2002**, *53*, 317–325.
38. Cabrera, M. A.; Ruiz, A.; Fernández, C.; González, I.; Bermejo, M. *Eur. J. Pharm. Biopharm.* **2003**, *56*, 197–206.
39. Pérez, M.; Helguera, A. *J. Comput. Aided Mol. Des.* **2003**, *17*, 665–672.
40. Wang, R.; Gao, Y.; Lai, L. *Perspect. Drug Discovery Des.* **2000**, *19*, 47–66.
41. Ertl, P.; Rohde, B.; Selzer, P. *J. Med. Chem.* **2000**, *43*, 3714–3717.
42. Ghose, A. K.; Crippen, G. M. *J. Chem. Inf. Comput. Sci.* **1987**, *27*, 21–35.
43. Gasteiger, J.; Marsilli, M. *Tetrahedron Lett.* **1978**, *34*, 3181–3184.
44. Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8533–8542.
45. Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441–451.
46. Estrada, E.; Uriarte, E.; Gutierrez, Y.; González, H. *SAR QSAR Environ. Res.* **2003**, *14*, 145–163.
47. Gutiérrez, Y.; Estrada, E. 2002. *MODESLAB 1.0 (Molecular DEScriptors LABoratory) for Windows*. Universidad de Santiago de Compostela, España.
48. Weininger, D. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36.
49. Negwer, M. *Organic-Chemical Drugs and their Synonyms*; Akademie: Berlin, 1987.
50. StatSoft. *STATISTICA version 5.5 for Windows*, StatSoft, **1999**.
51. Estrada, E. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 320–328.
52. Estrada, E.; Uriarte, E.; Montero, A.; Teijeira, M.; Santana, L.; De Clercq, E. *J. Med. Chem.* **2000**, *43*, 1975–1985.
53. Estrada, E.; Uriarte, E. *SAR QSAR Environ. Res.* **2001**, *12*, 309–324.
54. Estrada, E.; Molina, E.; Uriarte, E. *SAR QSAR Environ. Res.* **2001**, *12*, 445–459.
55. Randic, M. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 311–320.
56. Randic, M. *New J. Chem.* **1991**, *15*, 517–525.
57. Randic, M. *J. Mol. Struct. (Teochem)* **1991**, *233*, 45–59.
58. Lucic, B.; Nikolic, S.; Trinajstic, N.; Juric, D. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 532–538.
59. Estrada, E.; Perdomo, I.; Torres-Labandeira, J. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1561–1568.
60. Randic, M. *J. Comput. Chem.* **1993**, *4*, 363–370.
61. Xu, J.; Hagler, A. *Molecules* **2002**, *7*, 566–600.
62. Conradi, R. A.; Burton, P. S.; Borchardt, R. T. In *Lipophilicity in Drug Action and Toxicology*; Pliška, V., Testa, B., van de Waterbeemd, H., Eds.; VCH: Weinheim, 1996; pp 233–252.
63. Platts, J. A.; Abraham, M. H.; Zhao, Y. H.; Hersey, A.; Ijaz, L.; Butina, D. *Eur. J. Med. Chem.* **2001**, *36*, 719–730.
64. Sieghart, W. *Pharmacol. Rev.* **1995**, *47*, 181–234.
65. Fryer, R. I.; Leimgruber, W.; Trybulski, K. *J. Med. Chem.* **1982**, *25*, 1050–1055.
66. Maddalena, D. J.; Johnston, G. A. R. *J. Med. Chem.* **1995**, *38*, 715–724.
67. So, S. S.; Karplus, M. *J. Med. Chem.* **1996**, *39*, 5246–5256.
68. Hasegawa, K.; Funatsu, K. *J. Mol. Struct. (Theochem)* **1998**, *425*, 255–262.
69. Potkin, S. G.; Saha, A. R.; Kujawa, M. J.; Carson, W. H.; Ali, M.; Stock, E.; Stringfellow, J.; Ingenito, G.; Marder, S. R. *Arch. Gen. Psychiat.* **2003**, *60*, 681–690.