Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 14 (2006) 1108–1114
Bioorganic & Medicinal Chemistry

QSAR study on *para*-substituted aromatic sulfonamides as carbonic anhydrase II inhibitors using topological information indices

Georgia Melagraki,[a] Antreas Afantitis,[a,d] Haralambos Sarimveis,[b,*] Olga Igglessi-Markopoulou[a] and Claudiu T. Supuran[c]

[a]*Laboratory of Organic Chemistry, School of Chemical Engineering, National Technical University of Athens, 9, Heroon Polytechniou Str., Zografou Campus, Athens 15780, Greece*
[b]*Laboratory of Process Control and Informatics, School of Chemical Engineering, National Technical University of Athens, 9, Heroon Polytechniou Str., Zografou Campus, Athens 15780, Greece*
[c]*Universita degli Studi di Firenze, Polo Scientifico, Laboratorio di Chimica Bioinorganica, Room 188, Via della Lastruccia 3, 50019 Sesto Fiorentino (Florence), Italy*
[d]*Department of Chemo-informatics, Nova Mechanics Ltd, Larnaca, Cyprus*

Received 8 June 2005; revised 5 September 2005; accepted 13 September 2005
Available online 5 October 2005

Abstract—A linear quantitative structure–activity relationship has been developed for a series of *para*-substituted aromatic sulfonamides by using topological index methodologies. The compounds were studied for their carbonic anhydrase II (CAII) inhibitory activity. A large series of topological indices were calculated and the stepwise regression method was used to derive the most significant model. Very good results were obtained using multi-parametric regressions and showed that the information approach used in the present work is quite useful for modeling carbonic anhydrase inhibition.

1. Introduction

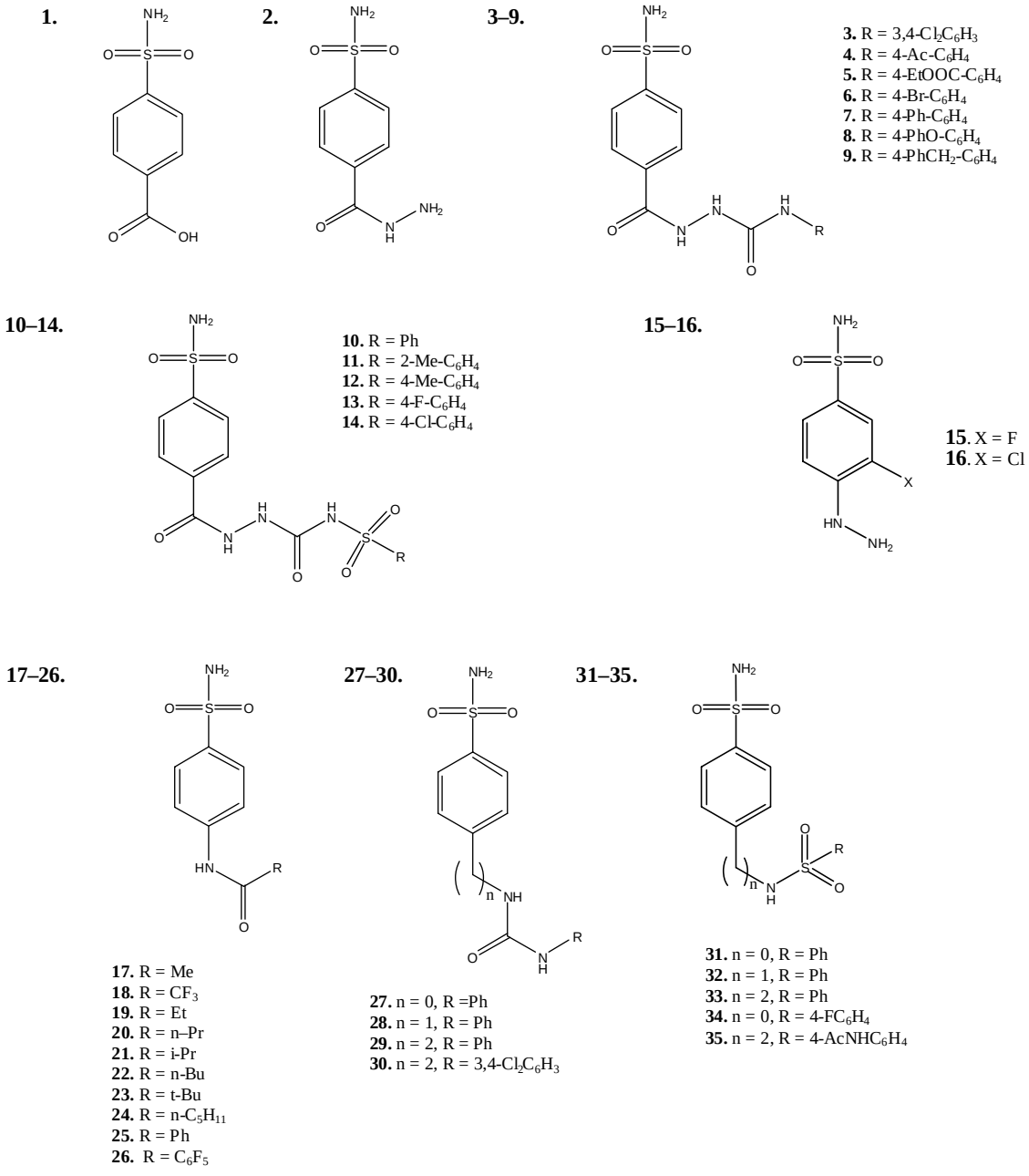
Carbonic anhydrases (CAs, EC 4.2.1.1) occupy a special place among the metallo-enzymes extensively studied in the last decade. These enzymes are ubiquitous in all kingdoms such as *Archaea*, *Bacteria*, algae, green plants as well as superior animals including vertebrates, and are encoded by five distinct, evolutionarily unrelated gene families: the α-CAs (present in vertebrates), the β-CAs (mainly present in *Bacteria* and plants), the γ-CAs (mainly in *Archaea*), and the recently isolated δ- and ε-classes of CAs (present in marine diatoms and chemolithoautotrophic bacteria, respectively). In higher vertebrates, including humans, 15 different isozymes were described among which the cytosolic CA II is physiologically one of the most important isoforms.[1,2]

CAs were proved to be very important as they are involved in crucial physiological processes, connected with the catalysis of the reversible hydration of carbonic dioxide to bicarbonate and a proton, as these chemical species are important in many physiological processes. This is a crucial reaction for respiration and transport of CO_2/bicarbonate between metabolizing tissues and excretion sites, secretion of electrolytes in a variety of tissues and organs, pH regulation and homeostasis, biosynthetic reactions (gluconeogenesis, lipogenesis, and ureagenesis), bone resorption, calcification, tumorigenicity, and many other physiologic or pathologic processes. Due to their important role, inhibition of these enzymes by carbonic anhydrase inhibitors (CAIs) may be exploited for the design of therapeutic agents useful in the management and prevention of many diseases.[3–6] Sulfonamides represent an important class of biologically active compounds. With the sulfanilamides as the lead structure, different classes of pharmacological agents have been obtained such as antibacterial sulfanamides, sulfonamides that inhibit the zinc enzyme carbonic anhydrase, the hypoglycemic sulfonamides extensively used in the treatment of some forms of diabetes, antithyroid drugs, and others.[7]

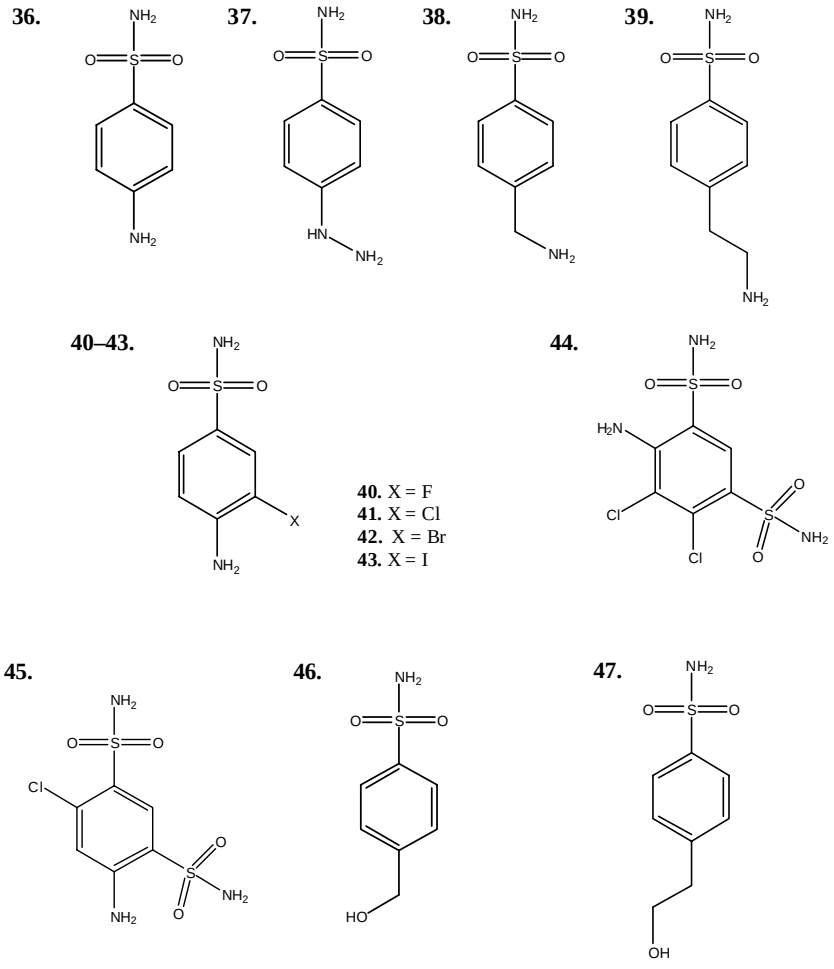
Keywords: Carbonic anhydrase; QSAR; Sulfonamides; Topological indices.
* Corresponding author. Tel.: +30 210 7723237; fax: +30 210 7723138; e-mail: hsarimv@central.ntua.gr

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.09.038

A great number of sulfanilamide derivatives were synthesized, characterized, tested, and are widely used in clinical medicine as pharmacological agents with a wide variety of biological actions.[8,9] More specifically, sulfonamide inhibitors of CA are extensively used in clinical medicine and as diagnostic tools, their main applications being in the treatment of glaucoma, macular edema, epilepsy, and other neurological disorders. CA inhibition with sulfanilamide discovered by Mann and Keilin has led to important drugs such as the sulfamides with CA inhibitory properties. Several such drugs are presently available, such as the recently introduced topical sulfonamides dorzolamide and brinzolamide, in addition to the classical, systemically acting inhibitors acetazolamide, methazolamide, ethoxzolamide, and dichlorophenamide which have been employed clinically for more than 45 years.[10] Sulfonamide CAIs derived from simple aromatic or heterocyclic sulfonamides have already shown excellent CA inhibitory properties against many CA isozymes isolated so far in diverse organisms. Many derivatives belonging to the heterocyclic and aromatic classes of sulfonamides have been synthesized and investigated for their biological activity.[11] The aromatic/heterocyclic sulfonamides act as carbonic anhydrase inhibitors and other types of derivatives show diuretic activity, hypoglycemic activity, anticancer properties or may act as inhibitors of the aspartic HIV protease being used for the treatment of AIDS and HIV infection among others.[12]



Scheme 1. The molecules used for this study are shown.



Scheme 1. *(continued)*

During the last few years, Supuran and his group has extensively studied different aromatic sulfonamides as potent carbonic anhydrase II inhibitors.[13–15] Although weaker CAIs as compared to heterocyclicones several aromatic sulfonamides were proved to be strong inhibitors with low K_I values within the nanomolar range. Special attention was paid to aromatic sulfonamides substituted at the *para*-position as they exhibit higher affinity with the zinc enzyme compared to an *ortho*-substituted aromatic sulfonamide. This may be due to the steric impairment of the *ortho*-substituent for the binding of such compounds to the Zn(II) ion within the enzyme active site.

Due to the biological importance of sulfonamides as potent CAIs, quantitative structure–activity relationship (QSAR) models have been proposed for the prediction of CA inhibitory activity of different aromatic and heterocyclic sulfonamides using different molecular descriptors. QSARs are mathematical relationships between a set of descriptors and the biological activity of the system being studied. QSAR models can be used as a useful tool in drug design, as they have the potential to decrease the time and effort required to develop new molecules by reducing the need for costly and time-consuming trial-and-error experiments.[16]

Clare, Supuran,[17,18] and Khadikar[19,20] have published some very interesting models for the prediction of CA II. In this work, a QSAR model is investigated in order to predict the CA II inhibitory activity of *para*-substituted aromatic sulfonamides using topological information indices and more specifically topological information indices. Our objective is to develop a rapid and reliable method to predict the CA inhibition activity.

The sulfonamides used incorporate hydrazine moieties, ureas, sulfureas or a simple aliphatic tail. The proposed model can be used as a first step for the formulation of an optimization problem from which the best *para*-substituent will be derived.

2. Results and discussion

2.1. Data set

The first step in developing QSAR equations is to compile a list of compounds for which the experimentally determined inhibitory activity is known. This list consists of 47 *para*-substituted aromatic sulfonamides collected from the literature.[21–23] The aromatic sulfonamides are presented in Scheme 1. The inhibition data are expressed in

Table 1. Experimental-predicted values

	Descriptors				Exp. activity K_I (nM)	MLR (Eq. 3) predicted activity	LOO predicted activity
	$^1\chi_{inf}$	$^0\chi^v_{inf}$	$^1\chi^v_{inf}$	N-rings			
1	2.0382	2.0759	2.7774	1	2.4116	2.1081	2.0676
2	2.1488	1.7885	2.5495	1	2.0934	2.1315	2.1408
3	1.9054	1.9551	2.6079	2	1.1139	1.3877	1.4000
4	1.9023	1.879	2.4897	2	1.1761	1.3834	1.3925
5	1.9825	2.0481	2.5069	2	0.9542	1.1510	1.1679
6	1.7729	1.8742	2.5835	2	0.8633	1.5756	1.6069
7	1.718	1.5986	2.3455	3	1.0414	0.8889	0.8591
8	1.5931	1.6358	2.4147	3	1.2553	1.0175	0.9755
9	1.5931	1.7561	2.5397	3	1.1761	0.9852	0.9532
10	1.9568	1.8262	2.6296	2	1.8261	1.4839	1.4519
11	2.1167	1.9931	2.629	2	1.7324	1.1665	1.1124
12	2.0206	1.9931	2.594	2	0.9912	1.2318	1.2475
13	2.0206	1.9931	2.7229	2	0.9777	1.3228	1.3554
14	2.0206	1.9931	2.7229	2	0.9590	1.3228	1.3571
15	2.192	2.0382	2.4997	1	1.7076	1.8058	1.8151
16	2.192	2.0382	2.4997	1	1.8808	1.8058	1.7986
17	1.7274	2.067	2.6106	1	2.3909	2.2905	2.2844
18	1.6457	2.1622	2.8222	1	2.1239	2.4210	2.4673
19	1.8228	2.2826	2.8402	1	2.3655	2.1472	2.1233
20	1.8024	2.3522	2.8278	1	2.3560	2.0879	2.0571
21	2.0079	2.1085	2.7744	1	2.4116	2.1017	2.0675
22	1.7755	2.3746	2.7933	1	2.3304	2.0663	2.0355
23	1.6457	2.1622	2.5366	1	2.3617	2.2195	2.2080
24	1.7448	2.3719	2.7472	1	1.7993	2.0652	2.0960
25	1.6955	1.6383	2.4332	2	1.5682	1.7784	1.8025
26	2.1526	2.0352	2.5583	2	1.2304	1.0408	1.0185
27	1.4519	1.6477	2.439	2	2.3802	2.0013	1.9406
28	1.4395	1.8455	2.6515	2	2.0212	1.9648	1.9568
29	1.4247	1.9374	2.6745	2	1.8751	1.9028	1.9063
30	1.7729	2.2202	2.8235	2	1.1139	1.3984	1.4350
31	1.5492	1.761	2.3078	2	1.6902	1.7041	1.7055
32	1.5198	1.9533	2.3788	2	1.6021	1.5891	1.5876
33	1.4906	2.0404	2.3764	2	1.4472	1.5275	1.5426
34	1.7749	2.0047	2.4251	2	0.9542	1.3313	1.3623
35	1.7879	2.247	2.671	2	1.8751	1.2499	1.1646
36	1.6767	1.6858	2.0404	1	2.4771	2.3174	2.2528
37	1.65	1.7296	2.2842	1	2.5051	2.4706	2.4635
38	1.65	1.9591	2.585	1	2.2304	2.4530	2.4765
39	1.6143	2.0458	2.5654	1	2.2041	2.3858	2.4008
40	2.085	2.0546	2.4508	1	1.7782	1.8551	1.8610
41	2.085	2.0546	2.4508	1	2.0414	1.8551	1.8409
42	2.085	2.0546	2.4508	1	1.6021	1.8551	1.8744
43	2.085	2.0546	2.4508	1	1.8451	1.8551	1.8558
44	2.1895	2.2281	2.5136	1	1.4472	1.6277	1.6507
45	2.1556	2.1494	2.3994	1	1.8751	1.6577	1.6257
46	1.65	2.2842	2.585	1	2.0969	2.1274	2.1310
47	1.6143	2.3459	2.5654	1	2.0414	2.0853	2.0932

terms of nanomolar affinity (K_I) for the investigated isozyme and are presented in the sixth column of Table 1.

2.2. Descriptors

First, the chemical structures were designed using MDL ISIS/DRAW 2.5 and were saved as .mol files.[24] The TOPIX[25] program was used to calculate the values of 29 topological descriptors shown in Table 2. Topological indices are 2D descriptors which take into account the internal atomic arrangement of compounds and encode in numerical form information about molecular size, shape, branching, presence of heteroatoms, and multiple bonds. Topological indices are a very useful tool for QSAR taking into account their simplicity and rapidity of computation.[26] This is particularly valuable now as one can analyze structures used for QSAR studies prior to any high throughput synthesis and testing.

2.3. Statistical analysis

Once the descriptors have been computed, it is necessary to decide which ones will be used. Among the aforementioned indices the selection of the best combinations was made with the use of an Elimination Selection Stepwise Regression (ES-SWR) algorithm that was developed in-house. The aim of variable subset selection is to reach optimal model complexity in predicting a response variable by a reduced set of descriptors that are not highly intercorrelated.

Table 2. Calculated descriptors

	Descriptor
1	Kappa^{1}
2	Kappa^{2}
3	Kappa^{3}
4	Mean Wiener
5	Wiener information index
6	Polarity
7	Gordon
8	Balaban
9	Schultz
10	Quadratic index
11	Zagreb^{1}
12	Zagreb^{2}
13	Wiener
14	Number of rings
15	Number of branches
16	Topological diameter
17	Topological radius
18	Xu^{1}
19	Xu^{2}
20	Xu^{3}
21	Kier-Hall^{0} ($^{0}\chi^{v}$)
22	Kier-Hall^{1} ($^{1}\chi^{v}$)
23	Kier-Hall^{inf,0} ($^{0}\chi^{v}_{inf}$)
24	Kier-Hall^{inf,1} ($^{1}\chi^{v}_{inf}$)
25	Randic^{0} ($^{0}\chi$)
26	Randic^{1} ($^{1}\chi$)
27	Randic^{inf,0} ($^{0}\chi_{inf}$)
28	Randic^{inf,1} ($^{1}\chi_{inf}$)
29	Modified Randic index

The best models that were produced are shown in Table 3. The descriptors that were used by the models are the mean information content based on the vertex degree equality and the edge equality both for the δ_i and δ^{v}_{i} values and the indicator parameter accounting for the number of rings in the molecule. The details concerning the information indices are given in Appendix A.

The correlation matrix for the aforementioned indicators is presented in Table 4 and shows that there is no significant correlation among the descriptors. From the correlation matrix we can also conclude that none of the aforementioned indices is highly correlated with the activity. This means that it is not possible to obtain a statistically significant mono-parametric model. Based on the correlation matrix we conclude that only multi-parametric regressions involving combinations of the indices mentioned before may result in a statistically significant regression expression.

Among the proposed models, the best multi-parametric models were found to be the following:

The best bi-parametric model was:

$$\log K_I = -0.96(\pm0.40)^{1}\chi_{inf} - 0.6852(\pm0.15)\text{N-rings} + 4.57(\pm0.85) \quad n = 47,\ R = 0.8149,$$
$$R^2 = 0.6642,\ R^2_{adj} = 0.6489,\ \text{RMS} = 0.2926,\ F = 43.52. \tag{1}$$

Among the tri-parametric models the best one was found to be the following:

$$\log K_I = -0.93(\pm0.39)^{1}\chi_{inf} - 0.56(\pm0.51)^{0}\chi^{v}_{inf} - 0.79(\pm0.17)\text{N-rings} + 5.79(\pm1.38) \quad n = 47,$$
$$R = 0.8356,\ R^2 = 0.6984,\ R^2_{adj} = 0.6773,$$
$$\text{RMS} = 0.2773,\ F = 33.18. \tag{2}$$

The best tetra-parametric model was the following:

$$\log K_I = -0.94(\pm0.37)^{1}\chi_{inf} - 1.00(\pm0.64)^{0}\chi^{v}_{inf} + 0.71(\pm0.66)^{1}\chi^{v}_{inf} - 0.85(\pm0.17)\text{N-rings} + 4.98(\pm1.52) \quad n = 47,\ R = 0.8534,$$
$$R^2 = 0.7283,\ R^2_{adj} = 0.7024,\ RMS = 0.2632,$$
$$F = 28.14. \tag{3}$$

Table 3. Produced models

	R^2	RMS	R^2_{cv}	RMS_{cv}
Biparametric				
$^{0}\chi_{inf}$ · N-rings	0.5272	0.3472	0.4379	0.3785
$^{1}\chi_{inf}$ · N-rings	0.6642	0.2926	0.6217	0.3106
$^{0}\chi^{v}_{inf}$ · N-rings	0.5375	0.3434	0.4774	0.3650
$^{1}\chi^{v}_{inf}$ · N-rings	0.4909	0.3603	0.4216	0.3840
Triparametric				
$^{0}\chi_{inf}$·$^{1}\chi_{inf}$·N-rings	0.6645	0.2924	0.5938	0.3218
$^{0}\chi_{inf}$ · $^{0}\chi^{v}_{inf}$ · N-rings	0.5777	0.3281	0.4763	0.3654
$^{0}\chi_{inf}$ · $^{1}\chi^{v}_{inf}$ · N-rings	0.5275	0.3470	0.4090	0.3881
$^{1}\chi_{inf}$ · $^{0}\chi^{v}_{inf}$ · N-rings	0.6984	0.2773	0.6480	0.2996
$^{1}\chi_{inf}$·$^{1}\chi^{v}_{inf}$ · N-rings	0.6644	0.2925	0.6059	0.3169
$^{0}\chi^{n}_{inf}$ · $^{1}\chi^{v}_{inf}$ · N-rings	0.5635	0.3336	0.4819	0.3634
Tetraparametric				
$^{0}\chi_{inf}$ · $^{1}\chi_{inf}$ · $^{0}\chi^{v}_{inf}$ · N-rings	0.6995	0.2768	0.6233	0.3099
$^{0}\chi_{inf}$ · $^{1}\chi_{inf}$ · $^{1}\chi^{v}_{inf}$ · N-rings	0.6647	0.2924	0.5767	0.3285
$^{0}\chi_{inf}$·$^{0}\chi^{v}_{inf}$ · $^{1}\chi^{v}_{inf}$ · N-rings	0.6055	0.3171	0.4875	0.3614
$^{1}\chi_{inf}$ · $^{0}\chi^{v}_{inf}$ · $^{1}\chi^{v}_{inf}$ · N-rings	0.7283	0.2632	0.6682	0.2908
Pentaparametric				
$^{0}\chi_{inf}$·$^{1}\chi_{inf}$·$^{0}\chi^{v}_{inf}$ · $^{1}\chi^{v}_{inf}$ · N-rings	0.7296	0.2625	0.6481	0.2995

Table 4. Correlation matrix

	$^0\chi_{inf}$	$^1\chi_{inf}$	$^0\chi^v_{inf}$	$^1\chi^v_{inf}$	N-rings	Activity
$^0\chi_{inf}$	1					
$^1\chi_{inf}$	0.499582	1				
$^0\chi^v_{inf}$	0.333857	0.231057	1			
$^1\chi^v_{inf}$	0.082052	0.109194	0.603135	1		
N-rings	−0.63249	−0.29885	−0.56893	−0.13553	1	
Activity	0.294774	−0.18833	0.220363	0.078274	−0.70042	1

Finally, the penta-parametric model is shown below:

$$\begin{aligned}\log K_I = {} & 0.71(\pm 0.67)^0\chi_{inf} - 0.30(\pm 1.37)^1\chi_{inf} \\ & - 0.90(\pm 0.41)^0\chi^v_{inf} - 1.01(\pm 0.65)^1\chi^v_{inf} \\ & - 0.87(\pm 0.21)\text{N-rings} + 5.49(\pm 2.77) \\ n = 47,\ & R = 08541,\ R^2 = 0.7296, \\ R^2_{adj} = 0.6966,\ & \text{RMS} = 0.2625,\ F = 22.12. \end{aligned} \quad (4)$$

The cross-validation statistical technique has been applied to estimate the quality with regard to predictive ability of the models. This is the most common validation technique, where a number of modified data sets are created by deleting, in each case, one or a smaller group of objects from the data in such a way that each object is taken away once and only once. For each reduced data set, the model is calculated, and responses for the deleted objects are predicted from the model.[27,28] The simplest and most general cross-validation procedure is the leave-one-out technique (LOO technique), where each object of the data set is taken away, one at a time. In this case, given n objects, n reduced models are developed. This technique is particularly important as the deletion scheme is unique and the results of different methods are easily compared.

PRESS is the prediction error sum of squares, derived from the LOO procedure. From the PRESS statistic the SSY (sum of squares of deviations of the experimental values from their mean), R^2_{CV} and S_{PRESS} statistics can be easily calculated (Eqs. 5 and 6).

$$R^2_{CV} = 1 - \frac{\text{PRESS}}{\text{SSY}} = 1 - \frac{\sum_{i=1}^{n}(y_{exp} - y_{pred})^2}{\sum_{i=1}^{n}(y_{exp} - \bar{y})^2}, \quad (5)$$

$$S_{PRESS} = \sqrt{\frac{\text{PRESS}}{n}}. \quad (6)$$

In Table 3, the statistical values R^2, RMS for the prediction and R^2_{cv}, RMS_{cv} for the cross-validation method, are presented for each model.

In terms of the R^2_{cv}, RMS_{cv} statistics, the tetra-parametric model that uses the descriptors $^1\chi_{inf}$, $^0\chi^v_{inf}$, $^1\chi^v_{inf}$, and N-rings was found to be the most accurate. Since these statistics are the most important as far as the predicting abilities of the produced models are concerned, the results of the aforementioned tetra-parametric model are shown in detail in Table 1. More specifically, Table 1 depicts the values of the four topological descriptors that were utilized by the model, the predictions produced by Eq. 3, and the results that were obtained using the LOO cross-validation method for this specific set of topological descriptors.

3. Conclusions

The inhibition of the physiologically relevant isoform CA II can be successfully modeled using topological information indices. Different multi-parametric models were established among which, a tetra-parametric model with $^1\chi_{inf}$, $^0\chi^v_{inf}$, $^1\chi^v_{inf}$, and N-rings was found to be statistically most significant. To the best of our knowledge, information indices have never been used before for this purpose.

These indices can be considered as a quantitative measure of the lack of structural homogeneity or the diversity of a graph, in this way being related to symmetry associated with structure. They are graph-theoretical indices that view the molecular graph as a source of different probability distributions to which information theory definitions can be applied. The information content of a graph is not unique, depending on the equivalence relation defined on the graph.

The mean information content used in this study is based on partitioning graph elements in equivalence. Elements are considered to be equivalent if their values are equivalent. Specifically the mean information content on the edge equality is based on the equivalence of the edge connectivity (calculated from δ_i and δ^v_i values) and the mean information content on the vertex degree equality is based on the equivalence of the vertex degree as shown from δ_i and δ^v_i values.

As a second step, we will further proceed with the optimization of the *para*-substituent. Having solved the forward problem, which is finding a statistically significant equation using topological descriptors, our aim is, within an optimization framework, to identify the structure with activity closest to a given value.

Acknowledgment

G.M. thanks the Greek State Scholarship Foundation for a doctoral assistantship.

Appendix A. Mean information content on the edge equality

This index is based on the partition of the edges in the graph according to equivalence of their edge connectivity values

$$^{\mathrm{E}}\bar{I}_x^{\mathrm{E}} = -\sum_{g=1}^{G} \frac{B_g}{B} \cdot \log_2 \frac{B_g}{B},$$

where B_g is the number of edges having the same edge connectivity, G is the number of different connectivity values, and B is the bond number.

In this work, the above formula was used to calculate descriptors $^{1}\chi_{\mathrm{inf}}$ and $^{1}\chi^{v}_{\mathrm{inf}}$ based on the edge connectivity derived from δ_i and δ^{v}_i values.

Mean information content on the vertex degree equality.

This index is derived from the adjacency matrix A and based on the partition of vertices according to the vertex degree equality.

$$^{\mathrm{V}}\bar{I}^{\mathrm{E}}_{\mathrm{adj,deg}} = -\sum_{g=1}^{G} \frac{^{g}F}{A} \cdot \log_2 \frac{^{g}F}{A},$$

where ^{g}F is the vertex degree count, that is, the number of vertices with the same vertex degree, A is the atom number, and G is the maximum vertex degree value.

In this work, the above formula was used to calculate descriptors $^{0}\chi_{\mathrm{inf}}$ and $^{0}\chi^{v}_{\mathrm{inf}}$ based on the vertex degree derived from δ_i and δ^{v}_i values.

References and notes

1. Pastorekova, S.; Parkkila, S.; Pastorek, J.; Supuran, C. T. *J. Enz. Inh. Med. Chem.* **2004**, *19*, 199.
2. Scozzafava, A.; Mastrolorenzo, A.; Supuran, C. T. *Expert Opin. Ther. Pat.* **2004**, *14*, 667.
3. Supuran, C. T.; Scozzafava, A.; Casini, A. *Med. Res. Rev.* **2003**, *23*, 146.
4. Casini, A.; Scozzafava, A.; Mincione, F.; Menabuoni, L.; Supuran, C. T. *J. Enz. Inh. Med. Chem.* **2002**, *17*, 333.
5. Casini, A.; Mincione, F.; Vullo, D.; Menabuoni, L.; Scozzafava, A.; Supuran, C. T. *J. Enz. Inh. Med. Chem.* **2002**, *17*, 9.
6. Kim, C.-Y.; Whittington, D. A.; Chang, J. S.; Liao, J.; May, J. A.; Christianson, D. W. *J. Med. Chem.* **2002**, *45*, 888.
7. Winum, J.-Y.; Scozzafava, A.; Montero, J.-L. ;.; Supuran, C. T. *Med. Res. Rev.* **2005**, *25*, 186.
8. Cecchi, A.; Winum, J.-Y.; Innocenti, A.; Vullo, D.; Montero, J.-L.; Scozzafava, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5775.
9. Franchi, M.; Vullo, D.; Gallori, E.; Pastorek, J.; Russo, A.; Scozzafava, A.; Pastorekova, S.; Supuran, C. T. *J. Enz. Inh. Med. Chem.* **2003**, *18*, 333.
10. Zimmerman, S.; Innocenti, A.; Casini, A.; Ferry, J. G.; Scozzafava, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 6001.
11. Chazalette, C.; Masereel, B.; Rolin, S.; Thiry, A.; Scozzafava, A.; Innocenti, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5781.
12. Hassan, M.-U.; Chohan, Z. H.; Scozzafava, A.; Supuran, C. T. *J. Enz. Inh. Med. Chem.* **2004**, *19*, 263.
13. Rusconi, S.; Innocenti, A.; Vullo, D.; Mastrolorenzo, A.; Scozzafava, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5763.
14. Scozzafava, A.; Menabuoni, L.; Mincione, F.; Briganti, F.; Mincione, G.; Supuran, C. T. *J. Med. Chem.* **2000**, *43*, 4542.
15. Scozzafava, A.; Menabuoni, L.; Mincione, F.; Supuran, C. T. *J. Med. Chem.* **2002**, *45*, 1466.
16. Debnath, A. K. In *Combinatorial Library Design and Evaluation: Principles Software Tools and Applications in Drug Discovery*; Ghose, A. K., Viswanadhan, V. N., Eds.; Marcel Dekker: New York, 2001, pp 73–129.
17. Supuran, C. T.; Clare, B. W. *J. Enz. Inh. Med. Chem.* **2004**, *19*, 237.
18. Supuran, C. T.; Clare, B. W. *Eur. J. Med. Chem.* **1998**, *33*, 489.
19. Agrawal, V. K.; Bano, S.; Supuran, C. T.; Khadikar, P. V. *Eur. J. Med. Chem.* **2004**, *39*, 593.
20. Jaiswal, M.; Khadikar, P. V.; Scozzafava, A.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3283.
21. Borras, J.; Scozzafava, A.; Menabuoni, L.; Mincione, F.; Briganti, F.; Minicione, G.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **1999**, *7*, 2397.
22. Vullo, D.; Franchi, M.; Gallori, E.; Antel, J.; Scozzafava, A.; Supuran, C. T. *J. Med. Chem.* **2004**, *47*, 1272.
23. Winum, J.-Y.; Dogne, J.-M.; Casini, A.; De Leval, X.; Montero, J.-L.; Scozzafava, A.; Vullo, D.; Innocenti, A.; Supuran, C. T. *J. Med. Chem.* **2005**, *48*, 2121.
24. Dalby, A.; Nourse, J.; Hounshell, D.; Gushurst, A.; Grier, D.; Leland, B.; Laufer, J. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 244.
25. http://www.lohninger.com/topix.
26. Todeschini, R.; Consonni, V. In *Handbook of Molecular Descriptors Methods and Principles in Medicinal Chemistry in Series of Methods and Principles of Medicinal Chemistry*; Wiley-VCH: Weinheim Germany, 2000; Vol. 11.
27. Efron, B. *Journal of American Statistical Association* **1983**, *78*, 316.
28. Osten, D. W. *J. Chemom.* **1988**, *2*, 39.