Journal of Enzyme Inhibition and Medicinal Chemistry, April 2009; 24(2): 337–349

QSAR studies for the inhibition of the transmembrane carbonic anhydrase isozyme XIV with sulfonamides using PRECLAV software

SHALINI SINGH[1], PADMAKAR V. KHADIKAR[2], ANDREA SCOZZAFAVA[3], & CLAUDIU T. SUPURAN[3]

[1]*Department of Chemistry, Bareilly College, Bareilly (U.P) 243001, India,* [2]*Research Division, Laxmi Fumigation and Pest Control, PVT. Ltd., 3, Khatipura, Indore 452 007, India, and* [3]*Laboratorio di Chimica Bioinorganica, Dipartimento di Chimica, University of Florence, via della Lastruccia, 3, Rm.188, Polo Scientifico, 50019 Sesto Fiorentino, Firenze, Italy*

(*Received 14 December 2007; accepted 5 March 2008*)

Abstract

QSAR studies for the inhibition of isozyme XIV of human carbonic anhydrase (CA, EC 4.2.1.1) by a series of sulfonamides including clinically used derivatives (acetazolamide, methazolamide, ethoxzolamide, dichlorophenamide, dorzolamide, brinzolamide, benzolamide, and zonisamide) are presented. Statistical calculations done using the PRECLAV program, for the correlation between the observed inhibition values and the calculated ones were good ($s = 0.2416$, $r^2 = 0.9259$, $F = 75.0196$, $r^2_{CV} = 0.8985$). The obtained results by using PRECLAV descriptors have been compared with those where the descriptors have been calculated with HYPERCHEM. The obtained QSAR equations pointed out the fact that the CA inhibitory activity decreased for unsubstituted (at the organic scaffold) aromatic/heteroaromatic sulfonamides, but was favored by the presence of alkyl groups substituting the scaffold, which led to a higher internal topological diversity, as well as by the presence of condensed aromatic rings in the structure of these enzyme inhibitors.

Keywords: *QSAR, Carbonic anhydrase XIV, sulfonamides, PRECLAV, HyperChem, NCSS, inhibition*

Introduction

Quantitative Structure-Activity Relationships (QSARs) are mathematical models relating measured biological activity of series of structurally related chemical compounds/pharmacological agents to the variation in their chemical structure. In cases in which some physico-chemical properties or toxicities of such compounds are correlated to their structures, the methodology is called quantitative structure-property relationships (QSPRs) or quantitative structure-toxicity relationships (QSTRs), respectively. Such methodologies are widely used in environmental toxicology to understand the adverse effects of chemical compounds. These predictions are important to be done due to the large number of untested chemicals present in nature ultimately, and because of the high costs of biological testing.

QSAR models are nowadays regarded as a scientifically credible tool for predicting and classifying biological activities of untested chemicals. QSAR has become inexorably embedded as an essential tool in the pharmaceutical industry, from lead discovery, optimization to lead development and computer –aided drug designing.A growing trend is to use QSAR early in the drug discovery process as a screening and enrichment tool to estimate from further development those chemicals lacking drug like properties or those chemicals predicted to elicit a toxic response. The fundamental assumption of QSAR is that variations in the biological activity of a series of chemicals that target a common mechanism of action are correlated

Correspondence: P. V. Khadikar, Research Division, Laxmi Fumigation and Pest Control, PVT. Ltd., 3, Khatipura, Indore 452 007, India. E-mail: pvkhadikar@rediffmail.com; C. T. Supuran, Laboratorio di Chimica Bioinorganica, Dipartimento di Chimica, University of Florence, via della Lastruccia, 3, Rm.188, Polo Scientifico, 50019 Sesto Fiorentino, Firenze, Italy. E-mail: claudiu.supuran@unifi.it

ISSN 1475-6366 print/ISSN 1475-6374 online © 2009 Informa UK Ltd.
DOI: 10.1080/14756360802187489

with variations in their structural, physical and chemical properties.

Since presumably these structurally related properties of a chemical can be determined by experimental or computational mean much more efficiently that its biological activity using *in-vivo* or *in-vitro* approaches, a statistically validated QSAR model is capable of predicting the biological activity of a new chemical within the same series in lieu of the time-consuming and lab our-intensive processes of chemical synthesis and biological evaluation. Applied judiciously, QSAR can save substantial amount of time, money, and human resources.

The ubiquitous metalloenzymes carbonic anhydrase (CAs, EC 4.2.1.1) catalyze the interconversion between carbon dioxide and the bicarbonate ion at the physiological pH. [1,2].These enzymes are inhibited by aromatic and heterocyclic sulphonamides, among which acetazolamide represents the prototype of a class of pharmacological agents with apparently limited therapeutic usefulness nowadays, but which played a major role in the development of fundamental renal physiology and pharmacology, as well as for the design of many of the presently widely used diuretic agents, such as among others the thiazide and high ceiling diuretics [1–7]. In mammals, 16 different α-CA isozymes or CA-related proteins (CARP) were described so far, with very different catalytic activity, sub-cellular localization, tissue distribution and susceptibility to be inhibited by sulfonamides [1–5,7,8]. Basically, there are several cytosolic forms (CA I-III, CA VII), five membrane-bound isozymes (CA IV, CA IX, CA XII, CA XIV and CA XV), two mitochondrial forms (CA VA and VB), as well as one secreted CA isozyme, CA VI. The acatalytic isozymes (CARPs) are CA VIII, CA X and CA XI. [1–7]. These enzymes are involved in crucial physiological processes connected with respiration and transport of CO_2/bicarbonate between metabolizing tissues and lungs, pH and CO_2 homeostasis, electrolyte secretion in a variety of tissues/organs, biosynthetic reactions (such as gluconeogenesis, lipogenesis and ureagenesis), bone resorption, calcification, tumorigenicity, and many other physiologic/pathologic processes [1–7].

Out of the known mammalian isoforms, the isozyme hCA XIV was one of the last to be discovered, by Nishimori's group in 1999 [8]. It has been subsequently reported that hCA XIV is highly abundant in the brain, kidney, urinary bladder,liver etc. [9–12].

Many CA isozymes are druggable targets, and their inhibitors may have applications in the design of antiglaucoma, antitumor, anticonvulant or antiobesity agents among others [13–22]. At this stage it is worh mentioning that the sulfonamides substituted at the sulfamoyl moiety ($ArSO_2NHR$) are known for about 70 years as antibacterial agents. In the last years, there have been many studies regarding the inhibitory effect of unsubstituted sulfonamides ($ArSO_2NH_2$) on various CA isozymes. In the last years a part of our research was directed towards detecting new classes of sulfonamides, with effects in the treatment of glaucoma or other pathologies [23–25].

Recently one of our groups [26] reported for the first inhibition study of the transmembrane isozyme CA XIV with a library of sulfonamides. We report here a QSARs study for the inhibition of human CA XIV with a series of aromatic and heterocyclic sulkfonamides investigated earlier [26]. The objective of the present study is to use the PRECLAV [27,28] program for the identification of: (i) significant molecular descriptors, (ii) significant molecular virtual fragments, (iii) outliers if any and finally to obtain (iv) statistically significant models. In addition, we have also used parameters generated from the HyperChem [29] software for obtaining statistically significant models and compared the results so obtained with those obtained using PRECLAV software.

Materials and methods

Calculations

The virtual construction of the molecules and minimization was performed to remove close atom contacts by 1000 cycles of minimization using standard Tripos force field [30] (with 0.005 kcal/mole energy gradient convergence criterion) by using Sybyl 7 [31] molecular modeling software package from Tripos,Inc. The next step was the rigorous optimization of the geometry using the Mopac software [32,33] for quantum mechanics calculation using the keywords string "AM1 Pulay gnorm = 0.01 shift = 50 geo-ok camp-king mmok bonds vectors".The output file from Mopac was used as input file for PRECLAV program [27,28] using these data, PRECLAV program calculated for each molecule the values for almost 400 whole molecule descriptors, specific for the program in question. The statistic calculations, including the QSAR equations, have been performed using the PRECLAVE program and best set of descriptor chosen/calculated. Ridge regression analyses [34] and the standard deviations of the correlating parameters in the proposed models were carried out using NCSS software [35].

The virtual constructions of the molecules and the geometry optimization have been done using the molecular mechanics software [28]. The next step was the rigorous optimization of the geometry using the Mopac software for quantum mechanics calculations [29] using the key-words string "am1 pulay gnorm = 0.01 shift = 50 geo-ok camp-king mmok bonds vectors". The output file from Mopac was used as input file for PRECLAV program [27,28]. Using these data, PRECLAV calculated for each molecule the values for almost 400 whole molecule descriptors, specific for the program in question. Separately, for each molecule,

the values for another set of descriptors, specific to this program have been calculated, using the Hyper Chem. program [30]. The statistic calculations, including the QSAR equations, have been performed using the PRECLAV program [27,28].

Recently, Tarko [14] has advocated that the PRECLAV program has used only "significant" descriptors in calculating the QSAR equations, descriptors that fulfill criteria (1).

$$r^2 > 4/N \quad (1)$$

Where:

r^2 is the square of the Pearson linear correlation between the values of the analyzed descriptor and the values of the dependent property and N is the number of molecules in the calibration set (here $N = 30$).

Furthermore, Tarko [14] also argued as below:

The program combines successively sets with 2, 3... k significant descriptors ($1 < k < 11$). A set of descriptors contains only descriptors that are sufficiently low intercorrelated and fulfill criteria (2).

$$r_{ij}^2 < N^{-1/2} \quad (2)$$

Where:

r_{ij}^2 is the square of Pearson linear correlation between the values of two descriptors present in the same set

Each set of descriptors has been used to calculate a multilinear QSAR equation of type (3).

$$IC = c_0 + \sum_{i=2}^{k} c_i \cdot D_i \quad (3)$$

Where:

IC is the dependent property (inhibition constant), c_0 is the free term (intercept), c_i are the coefficients (weighting factors) of the descriptors, D_i are some significant descriptors, k is the number of descriptors in the set.

The c_i coefficients have been calculated with OLSM (Ordinary Least Square Method). Tens of thousands of equations of type (3) have been calculated. With each of the type (3) equation, the values of IC have also been calculated. These computed values have been compared with the observed values. The concordance between the calculated/observed values has been calculated using the quality function Q.

$$Q = K_{CV} \cdot (N - k)/N \quad (4)$$

where:

K_{CV} [14] is Kendall cross-validated rank correlation between computed/observed values

As the value of k increases, the quality Q of the equations increases, reaches a maximum, then decreases. For the prediction, the equation with the highest value of Q is used and the descriptors used in this equation are called "predictors".

For every predictor, the "utility" U has been calculated.

$$U = (R^2 - r^2)/(1 - r^2) \quad (5)$$

Where:

R^2 is the square of the Pearson linear correlation between the observed values of IC and, the values calculated with the equation containing k predictors, r^2 is the square of the Pearson linear correlation between the observed values of IC and the values calculated with the equation containing $k - 1$ predictors, that is without the analyzed descriptor.

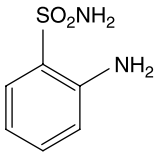
The values of U are then normalized (the highest becoming 1000). The "useful" predictors ($U > 400$) are sufficiently well correlated with IC and not too well correlated with the other predictors. Every "useful" predictor describes (reasonably) well the variation in the values of IC, and, at the same time, describes a different aspect than the other predictors.

Results and discussion

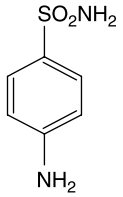
Some of the sulfonamides/sulfametes investigated in the present study for the inhibition of hCAXIV are given in Table I. The transmembrane isozyme hCAXIV, of the type **1**–**30** used have **22**–**30** as clinically used drugs, compounds **1**–**4,9,10,16, 21**–**27,29** and **3** are commercially available from Sigma-Aldrich, Merck, Alcon, Johnson & Johnson or DaiNippon, whereas compounds **5**–**8,11**–**15,20** and **28** were reported by one of the authors(Supuran) [26]. As stated above for the QSAR studies reported here in we have used two different sets of PRECLAV [Table II] and HyperChem [Table III] descriptors and obtained two different types of models [Table IV and V] depending upon these descriptors respectively. The dependent property was the inhibition constant i.e. $\log K_i$(hCAXIV), where K_i is the concentration of inhibitor where the inhibitory effect on hCAXIV isozyme starts. The values for inhibition constant (K_i, nM), used in the computation, have been taken from literature [26]. Greater the value of the concentration of inhibitor (IC), smaller the value of inhibitory activity. Furthermore, the external validation of all the proposed models was performed by cross-validation tests [34]. Also, due care is made to investigate the chance correlation [32–34].

First, we shall present the results of applying the PRECLAV algorithm in the two QSAR studies containing **30** compounds. The quality of the equations is given by the value of some usual statistical functions. Table I gives structural details of 30 sulfonamides/sulfamates used in the present study. The activities of the CA inhibitors ($\log K_i$(hCAXIV)) for this set of **30** compounds are recorded in Table II. This Table II also summaries PRECLAV descriptors used in the present

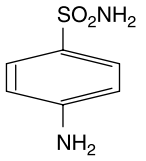
Table I. Structural details of sulfonamides/sulfamates used in the present study.



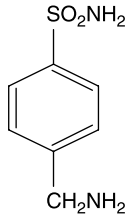
1



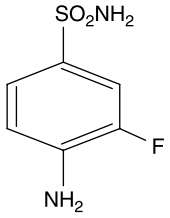
2



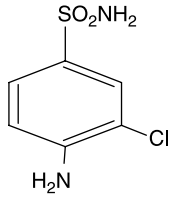
3



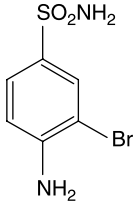
4



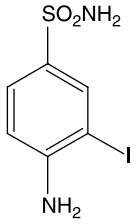
5



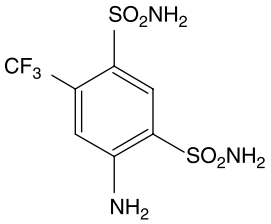
6



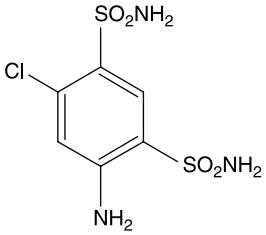
7



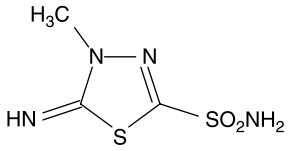
8



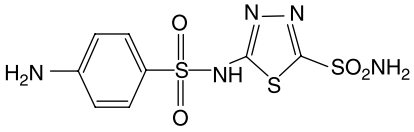
9



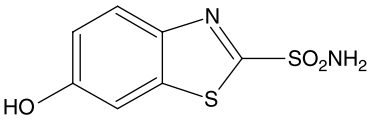
10



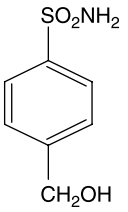
11



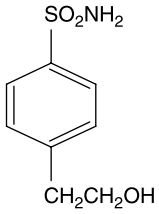
12



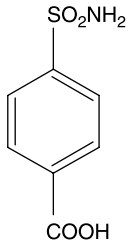
13



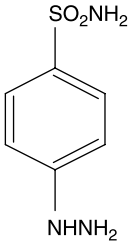
14



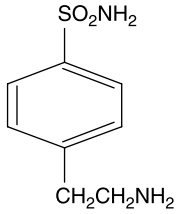
15



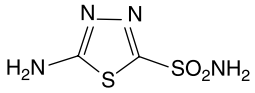
16



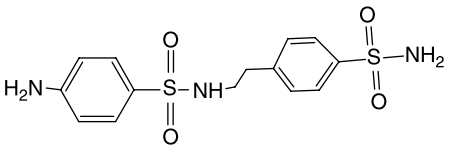
17



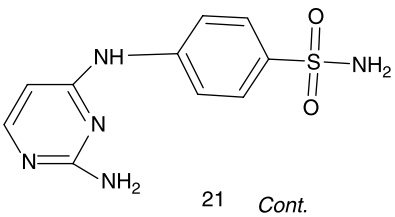
18



19

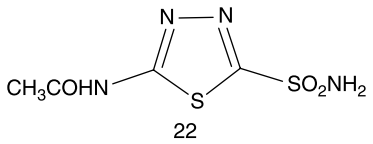


20



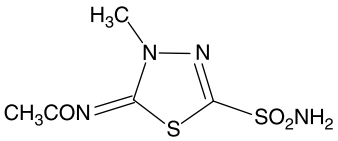
21 *Cont.*

Table I – *Continued*



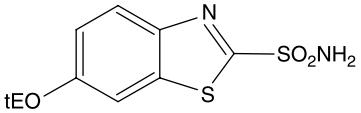
22

AAZ



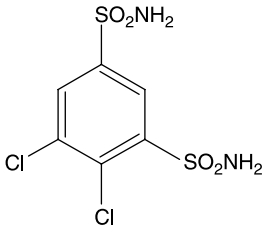
23

MZA



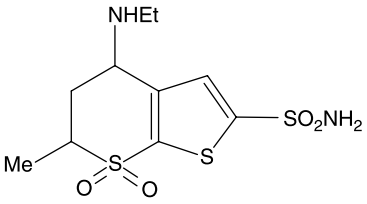
24

EZA



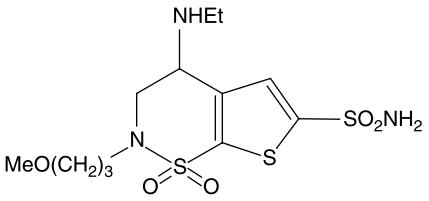
25

DCP



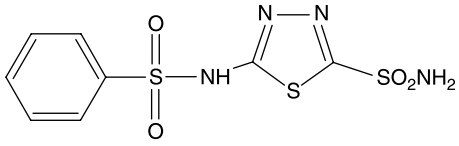
26

DZA



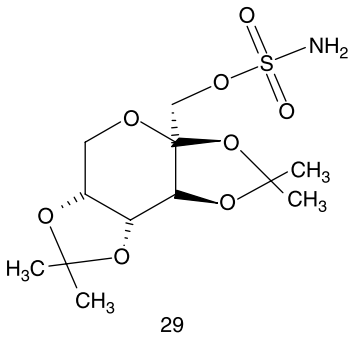
27

BRA



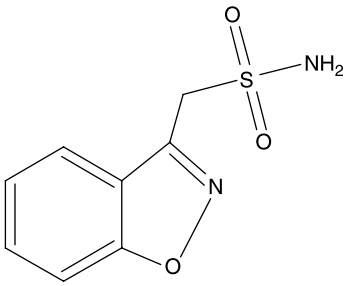
28

BZA



29

TPM



30

ZNC

study. The best 4- descriptor model obtained for the set of **30** compounds is discussed below:

Best set (by r2) of 4 descriptors: pca / ray / qgf / lhb / for the set of **30** compounds

Dependent property(IC) = Inhibition Constant against hCA XIV (Table II)

C0 = −5.4844
C1 = .0475
D1 = pca, is Percent of carbon,
C2 = −0.8167
D2 = ray, is Gyration radius Y of molecular area
C3 = 0.2305
D3 = qgf, is QSPR of mass fragment percents
C4 = 0.7443
D4 = lhb, is E(lumo) − E(homo-1) gap
Number of molecules in calibration set = 30
Number of significant descriptors: 264
Standard error (Se) = 0.4012
Pearson square correlation (r2) = 0.8083
Fisher function (*F*) = 27.4115
Pearson cross-validated square correlation leave one out method (r2CV) = 0.7600
Kendall cross-validated rank correlation leave one out method (rKCV) = 0.6184
Quality (*Q*) = 0.5359
Max. Quality = 0.6062
Outliers = 2(Compound no. 6 and 7)

The above information can be written in Equation type (3) form as given below:

$$\log K_i(\text{hCAXIV}) = -5.4844 + 0.4175\,(\text{pca}) - 0.8167(\text{ray}) + 0.2305(\text{qgf}) + 0.7443(\text{lhb}) \quad (6)$$

$N = 30$, $r^2 = 0.8083$, Se = 0.4012,
F = 27.4115, r2cv = 0.7600, rKCV = 0.6184

Table II. PRECLAV descriptors used in the significant models.

Compd. no.	ray	pca	qgf	lhb	pon	lhf	raz	nss
1	1.2824	41.841	2.6591	9.937	0.5399	0.1006	1.423	2.8632
2	1.0703	41.841	2.6591	9.931	0.5401	0.1007	1.4086	2.8635
3	1.4718	41.841	2.6591	9.841	0.6228	0.1016	1.405	2.8763
4	1.5874	45.138	5.1481	9.569	0.8518	0.1045	1.545	2.8673
5	1.5224	37.884	2.4272	9.58	0.5978	0.1044	1.6681	2.8745
6	1.4973	34.868	2.2506	9.617	0.582	0.104	1.7877	2.8761
7	1.8039	28.696	1.8869	9.635	0.578	0.1038	2.0582	2.8764
8	2.0144	24.172	1.6233	9.656	0.579	0.1036	2.2465	2.8767
9	1.8595	26.329	2.8516	10.066	0.6534	0.0993	2.1076	2.8731
10	1.6011	25.218	3.1626	9.724	0.5358	0.1028	2.0372	2.8623
11	1.4879	18.548	2.3798	10.468	0.8406	0.0955	1.6309	0.3799
12	2.1487	28.646	1.57	8.608	1.4195	0.1162	2.3609	0.6642
13	1.8718	36.509	2.0398	8.408	0.912	0.1189	1.9781	0.6408
14	1.6063	44.902	2.8662	9.654	0.7215	0.1036	1.6641	2.8673
15	1.6305	47.739	2.6798	9.611	0.9454	0.104	1.6945	2.8673
16	1.7993	41.782	2.6825	9.623	0.8601	0.1039	1.8709	2.8623
17	1.6197	38.485	4.7176	9.664	0.8289	0.1035	1.6736	2.8731
18	1.5679	47.972	4.8034	9.553	0.8767	0.1047	1.6198	2.8679
19	1.5065	13.328	2.5485	10.052	0.0641	0.0995	1.6195	0.6418
20	2.7002	45.728	1.6634	9.552	2.1254	0.1047	2.7193	2.8675
21	2.2726	45.266	1.7973	8.939	1.511	0.1119	2.2406	2.873
22	1.9678	21.614	2.2638	9.901	1.3758	0.101	1.921	0.6508
23	1.9392	21.614	2.2638	9.753	1.2638	0.1025	2.0261	0.5729
24	2.1419	41.841	1.8421	8.44	1.8195	0.1185	2.2176	0.6329
25	1.4156	23.613	2.9754	9.703	0.7054	0.1031	2.2148	2.8619
26	1.4374	37.015	1.5074	8.442	1.336	0.1185	1.9521	1.0666
27	2.2049	37.379	1.3018	8.939	1.7383	0.1119	2.3458	0.4052
28	1.9996	35.99	0.3224	8.988	1.4455	0.1113	2.1721	0.5377
29	1.0322	38.081	1.5443	10.229	0.9975	0.0978	1.6977	2.8208
30	1.0221	45.269	2.1953	9.268	0.9268	0.1079	1.7484	2.8255

ray = Gyration radius Y of molecular area; pca = Percent of carbon; qgf = QSPR of mass fragment percents; lhb = E(lumo) - E(homo-1) gap; pon = RMS of distance to geometric center(H and Halogen atom); lhf = 1/[E(lumo) - E(homo-1)] ratio; raz = Gyration radius Z of molecular area; nss = minimum net charge of S atom.

The aforementioned results indicted that the model (Equation 6) contains compounds **6** and **7** as outliers. The deletion of these compounds as outliers from the regression process yielded the following statistically significant models:

(1) *Two variable model containing raz and nss as the Correlating parameters*

Dependent property = Inhibitory activity against hCA XIV (Table I)
Number of molecules in calibration set = 28
Number of significant descriptors: 264
c_0 = 4.6814
c_1 = −1.6277
D_1 = raz, is Gyration radius Z of molecular area (U = 1000)
c_2 = 0.4543
D_2 = nss, is minim net charge of S atom (U = 965)
Standard error s = 0.3162
Pearson square correlation r2 = 0.8732
Fisher function F = 89.4955
Pearson cross-validated square correlation r2CV = 0.846
Kendall cross-validated rank correlation rKCV = 0.7513
Quality Q = 0.6977
max. quality = 0.6977

This two-variable model is expressed by the following expression:

$$\log K_i(\text{hCA XIV}) = 4.6814 - 1.677\,\text{raz} + 0.4543\,\text{nss} \quad (7)$$

(2) Best set (by r2) of 3 descriptors containing pca, pon, lhb,as the correlating parameters

Dependent property = Inhibitory activity against hCA XIV (Table I)
Number of molecules in calibration set = 28
Number of significant descriptors: 264
c_0 = −6.2596
c_1 = 0.0642
D_1 = pca, is Percent of carbon (U = 1000)
c_2 = 0.7902
D_2 = pon, is RMS of distance to geometric

Table III. Quantum and Chemical descriptors computed by HyperChem used in the present study.

Compd. no.	MPC	LNC	MNC	SPC	SNC	SAC	E_{HOMO}	E_{LUMO}	χ	η	S	ω
1	2.879	0.03	0.957	4.572	4.571	9.143	−210.1	−9.76	100.21	−109.9	−0.009	−45.66
2	2.863	0.066	0.952	4.332	4.332	8.664	−207.1	−8.1	99.5	−107.6	−0.009	−46
3	2.877	0.237	0.953	4.576	4.576	9.154	−211.1	−7.52	101.82	−109.3	−0.009	−47.4
4	2.867	0.02	0.953	4.415	4.418	8.833	−229	−4.05	112.52	−116.5	−0.009	−54.3
5	2.875	0.003	0.953	4.42	4.419	8.839	−211.9	−3.3	104.33	−107.6	−0.009	−50.56
6	2.864	0.004	0.954	4.324	4.331	8.655	−228.8	−17.78	105.54	−123.5	−0.008	−45.16
7	2.877	0.29	0.953	4.649	4.648	9.297	−212.1	−11.01	100.57	−111.5	−0.009	−45.32
8	2.877	0.247	0.953	4.737	4.754	9.491	−213	−10.75	101.14	−111.8	−0.009	−45.74
9	2.865	0.151	0.959	8.476	8.477	16.953	−227.3	−29.6	98.69	−128.4	−0.008	−37.9
10	2.881	0.045	0.32	8.126	8.124	16.25	−224.6	−22.21	101.24	−123.4	−0.008	−41.51
11	0.306	0.045	0.299	1.44	1.439	2.879	−208.1	−26.54	90.78	−117.3	−0.009	−35.12
12	2.904	0.109	1.17	8.651	8.651	17.302	−224.1	−23.99	100.07	−124	−0.008	−40.36
13	2.884	0.017	0.996	4.812	4.813	9.625	−214.8	−29.81	92.53	−122.3	−0.008	−34.99
14	2.867	0.017	0.953	4.332	4.331	8.663	−232.9	−13.35	109.77	−123.1	−0.008	−48.93
15	2.868	0.007	0.954	4.497	4.497	8.994	−235.2	−14.84	110.18	−125	−0.008	−48.55
16	2.863	0.038	0.954	4.599	4.599	9.198	−241.6	−29.87	105.86	−135.7	−0.007	−41.28
17	2.873	0.206	0.954	4.576	4.576	9.152	−216.8	−4.08	106.37	−110.4	−0.009	−51.22
18	2.868	0.015	0.953	4.736	4.736	9.472	−230.1	−13.82	108.17	−121.9	−0.008	−47.95
19	2.903	0.0124	0.949	4.567	4.567	9.134	−223.3	−29.35	97	−126.3	−0.008	−37.23
20	2.867	0.013	0.954	8.664	8.665	17.329	−213.9	−15.93	98.9	−114.9	−0.009	−42.62
21	2.873	0.21	0.954	5.552	5.552	11.104	−209.8	−17.04	96.38	−113.4	−0.009	−40.94
22	2.899	0.162	1.129	5.062	5.062	10.124	−230.1	−37.9	96.09	−134	−0.007	−34.45
23	2.889	0.012	1.026	4.875	4.873	9.748	−220.6	−27.86	96.37	−124.2	−0.008	−37.37
24	2.885	0.016	1	5.028	5.027	10.055	−212.3	−19.96	96.18	−116.1	−0.009	−39.82
25	2.867	0.115	0.954	7.369	7.369	14.738	−240.7	−33.85	103.4	−137.3	−0.007	−38.97
26	2.901	0.048	1.174	8.862	8.862	17.724	−224.2	−37.74	93.21	−130.9	−0.008	−33.18
27	2.902	0.003	1.232	10.197	10.197	20.394	−223.8	−37.27	93.28	−130.5	−0.008	−33.32
28	2.967	0.013	1.833	9.25	9.25	18.501	−234.5	−35.16	99.65	−134.8	−0.007	−36.83
29	2.312	0.004	0.906	4.973	4.973	9.946	−202.7	−4.19	99.24	−103.4	−0.008	−47.61
30	2.824	0.007	0.953	4.302	4.299	8.601	−228.9	−20.11	104.42	−124.5	−0.008	−43.78

Compd. no.	DM_t	DM_x	DM_Y	DM_z	Hf	SA	VOL	Ref	Pol.	log p	HE
1	4.879	3.676	0.187	3.203	−53.08	257.04	496.95	44.1	14.27	0.15	−11.33
2	4.605	3.384	1.597	2.684	−48.72	271.97	490.52	47.76	14.27	0.15	−13.02
3	6.239	4.664	3.132	3.03	−51.64	267.69	489.21	44.1	14.27	0.15	−11.39
4	4.723	2.672	2.871	2.632	−49.75	304.64	541.71	51.92	16.11	0.05	−11.66
5	5.666	3.024	3.861	2.838	−95.64	280.64	498.19	48.39	14.18	0.29	−12.48
6	4.619	2.914	19.65	2.998	−55.1	334.69	579.76	52.68	18.04	0.57	−10.93
7	6.311	3.622	3.99	3.285	−46.35	311.5	545.77	55.8	16.9	0.01	−12.13
8	6.358	3.713	3.883	−3.401	−34.64	308.26	554.13	52.65	20.36	1.41	−10.73
9	7.822	2.11	4.361	6.141	−270	359.22	696.31	63.42	18.33	−0.08	−16.04
10	5.243	0.97	5.108	−0.677	−128	367.33	673	63.25	18.69	3.42	−17.47
11	1.258	−1.167	−0.349	−0.313	−95.91	272	485.22	39.14	15.27	0.2	−12.12
12	7.995	5.159	−6.065	−0.721	−49.83	462.72	840.15	79.01	25.07	3.24	−24.39

Table III – *Continued*

Compd. no.	DM_t	DM_x	DM_Y	DM_z	Hf	SA	VOL	Ref	Pol.	log p	HE
13	4.486	2.346	−2.423	−2.958	−62.67	312.8	568.39	55.5	18.56	2.78	−17.95
14	5.812	4.223	1.229	3.799	−98.04	337.04	551.07	46.22	15.4	0.4	−13.62
15	5.739	1.727	2.344	4.945	−106	362.18	612.26	53.21	18.14	0.98	−23.84
16	4.129	1.206	0.167	3.945	−135	340.66	558.1	43.16	15.48	3.43	−14.67
17	5.554	1.975	2.848	4.34	−23.9	316.06	547.8	49.18	15.63	0.25	−19.43
18	4.501	3.25	0.33	−3.094	−57.13	340.43	595.97	53.63	17.94	4.17	−11.95
19	5.497	−4.285	2.844	1.941	−16	263.03	440.13	38.34	12.38	0.62	−21.39
20	4.911	3.321	−2.597	1.298	−97.9	479.42	923.3	90.52	30.09	1.28	−19.24
21	5.843	3.471	3.726	2.865	13.4	357.47	715	70.48	23.87	4.98	−18.75
22	3.325	−6.432	−1.456	−2.957	−50.8	345.52	549	47.78	16.78	0.44	−16.33
23	4.104	−1.489	2.125	−3.18	−31.5	319.86	542.75	47.78	16.13	0.78	−13.14
24	5.237	3.298	−2.754	−2.994	−62.32	412.2	702.05	65.02	22.22	3.16	−13.62
25	7.238	0.794	1.406	−7.056	−123.26	388.02	646.97	62.35	19.27	0.85	−12.19
26	4.863	−1.387	−3.065	−3.512	−119.9	461.52	804.1	77	25.17	−0.04	−9.62
27	5.788	−5.033	−2.223	1.799	−164.5	569.72	977.01	91.82	30.83	−0.75	−10.98
28	1.849	1.321	−0.79	−1.026	−47.33	408.42	739.69	80.13	23.18	6.56	−18.33
29	6.321	4.136	−1.048	4.664	−329.4	421.91	731.02	64.03	22.83	−0.36	−8.9
30	3.217	3.102	−0.281	0.806	−11.5	274.74	561.29	56.01	17.39	2.517	−10.99

MPC-Most positive charges; LNC-Least negative charges; MNC-Most Negative charges; SPC-Sum of positive charges; SNC-Sum of negative charges; SAC-Sum of absolute charges; E_{HOMO}(Kcal\mol)-Highest occupied molecular orbital energy; E_{LUMO}, Kcal\mol-Lowest unoccupied molecular orbital energy; χ,Kcal\mol Electronegativity ($\chi = -0.5(E_{HOMO} - E_{LUMO})$); η,Kcal\mol Hardness ($\eta = 0.5\ (E_{HOMO} + E_{LUMO})$); S-Mol\Kcal, Softness ($S = 1\backslash\eta$); Pol(A^3)Kcal\mol-Electrophilicity ($\omega = \chi^2/2\eta$); DM_tDebyes-Total molecular Dipole moment; DM_xDebyes-Molecular dipole moment at x-direction; DM_yDebyes-Molecular dipole moment at y-direction; DM_zDebyes-Molecular dipole moment at z-direction; H_fKcal\mol-Heat of formation; SA(A^2)-Surface area; V(A^3)-Volume; Ref(A^3)-Ref: refractivity; log P-The log of octanol-water partition coefficient; HE Kcal\mol-Hydration energy.

Table IV. Statically significant models for the set of 28 compounds using PRECLAV parameters.

Model	Parameters used	R^2	R^2A	Se	F
1	raz,nss	0.8732	0.8460	0.3162	89.4955
2	pca,pon,lhb	0.9096	0.8891	0.2670	83.8137
3	pca,ray,pon,lhf	0.9259	0.8985	0.2416	75.0196

Table V. Statically significant models for the set of 28 compounds using HyperChem descriptors.

Model	Parameters Number used	R^2	R^2A	Se	F
1	ω	0.4906	0.4710	0.2669	25.036
2	POL, ω	0.6383	0.6094	0.2294	22.059
3	SA, HF, ω	0.7095	0.6732	0.2098	19.540
4	SA, H_f, χ, DM_y	0.7598	0.7181	0.1949	18.192
5	SA, H_f, ω, SPC, SNC	0.8453	0.8101	0.1599	24.042
6	SA, H_f, ω, SPC, SNC, DM_y	0.8698	0.8326	0.1502	23.382
7	SA, H_f, ω, SPC SNC, DM_y, DMx	0.8840	0.8434	0.1452	21.780

center(H and Halogenalions)
$c_3 = -1.0173$
D_3 = lhb, is E(lumo) - E(homo-1) gap
Standard error s = 0.267
Pearson square correlation r2 = 0.9096
Fisher function F = 83.8137
Pearson cross-validated square correlation r2CV = 0.8891
Kendall cross-validated rank correlation rKCV = 0.836
Quality Q = 0.7464
max. quality = 0.7559

This three-variable model is found as below:

$$\log K_i(\text{hCA XIV}) = -6.2596 + 0.0642\,\text{pca} + 0.7902\,\text{pon} - 1.0173\,\text{lhb} \quad (8)$$

(3) Best set (by r2) of 4 descriptors pca, ray, pon,and lhf containing, as the correlating Parameters

Dependent property = Inhibitory activity against hCA XIV (Table I)
Number of molecules in calibration set = 28
Number of significant descriptors: 264
c_0 = 8.7924
c_1 = 0.0583
D_1 = pca, is Percent of carbon (U = 1000)
c_2 = −64.7685
D_2 = ray, is Gyration radius Y of molecular area (U = 1000)
c_3 = −0.4487
D_3 = pon, is RMS of distance to geometric center (H and Halogenalions) (U = 965)
c_4 = −0.7438
D_4 = lhf, is 1/[E(lumo)-E(homo-1)] ratio
Standard error s = 0.2416
Pearson square correlation r2 = 0.9259
Fisher function F = 75.0196
Pearson cross-validated square correlation r2CV = 0.8985
Kendall cross-validated rank correlation rKCV = 0.8095
Quality Q = 0.6939
max. quality = 0.7559

This model is found as below:

$$\log K_i(\text{hCA XIV}) = 8.7924 + 0.0583(\pm 0.0057)\text{pca} - 64.7685((\pm 9.6861)\text{ray} - 0.4487((\pm 0.1900)\,\text{pon}_{-0.7438(0.1762)\text{lhf}} \quad (9)$$

n = 28, s = 0.2416, r2 = 0.9259, F = 75.0196

It is worth to mention that the PRECLAV program does not compute the error of coefficients of the parameters involved in the proposed models. We have, therefore, used the NCSS program [35] for performing these calculations. Also tthe PRECLAV program indicated that none of the models for the set of **28** compounds contain any outlier. Here it is worth mentioning that if the residue i.e. the difference between the observed and calculated activity for a particular compound is two – times or more than its standard deviation then it is considered as an outlier and is, therefore, deleted from the regression analysis.

The positive coefficient of pca in the aforementioned model [Equation 9] indicates that increase in percent of carbon is favorable for the exhibition of log hCA xiv, while negative coefficients of ray, pon, and lhf indicate that decrease in Gyration radius Y of molecular area, RMS of distance to geometric center(H and Halogenations) and 1/[E (lumo) - E (homo-1)] ratio respectively

Table VI. Observed and estimated values of log hCAXIV their residue(for 28 compound's) using 4-varible models obtained using PRECLAV and HyperChem parameters.

		log hCAXIV			
		PRECLAVE		HyperChem	
Compd. no.	obs.	Est.	Res.	Est.	Res.
1	3.813	3.742	0.071	3.224	0.589
2	3.681	3.831	−0.149	2.906	0.775
3	3.732	3.53	0.203	3.071	0.662
4	3.505	3.312	0.193	3.753	−0.248
5	2.255	3.113	−0.858	–	–
6	1.322	–	–	1.706	−0.384
7	1.176	–	–	–	–
8	1.898	2.157	−0.259	2.509	−0.612
9	2.833	2.576	0.256	3.002	−0.17
10	2.756	2.488	0.268	2.39	0.365
11	2.389	2.397	−0.007	2.466	−0.077
12	1.114	0.916	0.197	1.806	−0.692
13	1.681	1.702	−0.021	2.242	−0.561
14	3.584	3.445	0.139	3.57	0.015
15	3.513	3.407	0.105	3.35	0.163
16	2.892	3.053	−0.161	3.434	−0.542
17	3.161	2.991	0.17	2.946	0.215
18	3.462	3.455	0.008	3.242	0.22
19	2.447	2.401	0.046	2.498	−0.051
20	1.881	1.886	−0.005	1.557	0.324
21	1.949	2.041	−0.092	1.394	0.556
22	1.613	1.605	0.008	2.158	−0.545
23	1.633	1.604	0.029	2.066	−0.432
24	1.398	1.243	0.155	1.725	−0.327
25	2.389	2.333	0.056	2.64	−0.251
26	1.431	1.638	−0.207	1.328	0.104
27	1.38	1.443	−0.062	0.548	0.832
28	1.519	1.71	−0.192	1.871	−0.353
29	3.164	3.476	−0.312	3.176	−0.012
30	3.72	3.298	0.422	3.281	0.439

are favorable for modeling the activity $\log K_i$ (hCAXIV). Also, the information recorded in Table IX indicates that large mass percent of H2N and C6H4 fragments increase the activity value, while large mass percent of C3HN3S and NH fragments decrease the activity value.

In order to confirm our results we have estimated (calculated) $\log K_i$(hCAXIV)using the best model expressed by Equation 9 and compared them with the observed values $\log K_i$(hCAXIV). The results are shown in Table VI. The data gave the following correlation between estimated and observed $\log K_i$ (hCAXIV):

$$\log K_i(\text{hCA XIV})\text{stimated} = \text{constant} + c(\log K_i(\text{hCA XIV}))\text{observed} \quad (10)$$

$$s = 0.2799,\ r2 = 0.908,\ F = 82.2279,$$
$$r\,\text{Kendall} = 0.8677,\ r2\text{CV} = 0.887,\ \text{quality} = 0.7559$$

In order to at the final conclusion we have to examine the aforementioned 2-, 3- and 4-descriptors models for the set of 28 compounds with regard to the co-linearity problem. The models will be acceptable only when they are free from co-linearity defect. This can be made by investigating VIF (Variance Inflection Factor), Tolerance, Eigenvalues and Condition Number [34] for each of the descriptors present in a given model. These parameters are calculated using Ridge regression and are presented in Table VII.

Application of Ridge statistics [34] provides important statistical parameters namely variance inflation factors (*VIF*s) for each of the parameters involved in the model. The *VIF* is defined for each variable in the equation, and not for the equation as a whole, so there should be as many *VIF*s, as there are correlating parameters. The *VIF* is defined as:

$$VIF = 1(1 - R_i^2) \quad (11)$$

Where R_i is the multiple correlation coefficient of the i^{th} independent variable on all of the other independent variables. In the proposed models, all these *VIF*s should be less than 10 indicating that no co-linearity problem exists in the model.

A perusal of Table VII shows that in all the three proposed models the descriptors evolved in them have *VIF* considerably smaller than 10 indicating that the models are free from the defect due to co-linearity.

The Ridge regression analysis also provides λ-statistics helping us to resolve the problem of

Table VII. VIF values, Tolerance, Eigenvalue and Condition Number for the descriptors contained in the proposed models using PRECLAV descriptors for the set of 28 compounds.

Model	Variables Used	Variance Inflation factor	R-Squared	Tolerance	Eigenvalue	Condition Number
1	nss	1.0826	0.0763	0.9237	1.276186	1.00
	raz	1.0826	0.0763	0.9237	0.723814	1.76
2	pca	1.1242	0.1105	0.8895	1.782082	1.00
	Pon	1.4900	0.3289	0.6711	0.784004	2.27
	lhb	1.5270	0.3451	0.6549	0.433915	4.11
3	pca	1.2600	0.2063	0.7937	2.196745	1.00
	pon	2.6287	0.6196	0.3804	1.050654	2.09
	lhf	1.5196	0.3420	0.6580	0.518117	4.24
	Ray	2.1934	0.5441	0.4559	0.234484	9.37

Since all VIF's are less than 10, multicollinearity is not a problem. All Condition Numbers less than 100. Multicollinearity is NOT a problem.

Table VIII. VIF values, Tolerance, Eigenvalue and Condition Number for the best four - variable model for the set of 28 compounds using HyperChem parameters.

Independent Variable	Variance Inflation factor	R-Squared	Tolerance	Eigenvalue	Condition Number
SA	1.4232	0.2974	0.7026	0.666215	2.71
H_f	1.2281	0.1857	0.8143	0.483897	3.74
χ	1.1691	0.1394	0.8606	1.042437	1.73
DM_y	1.2669	0.2107	0.7893	1.807450	1.00

Since all VIF's are less than 10, multicollinearity is not a problem. All Condition Numbers less than 100. Multicollinearity is NOT a problem.

co-linearity. The λ-statistics is defined as below:

$$\lambda = \frac{1}{n}\sum_{i=1}\frac{1}{\lambda_i} \qquad (12)$$

Where n is the number of variables in the model (regression expression) and λ_i is the eigen-values of the correlation matrix of the independent variables.

If $\lambda < 5.0$, the sub-set is considered free from co linearity problem, and the Equation (model) is accepted. If λ is not < 5.0, then eigen-vector matrix is examined. The eigen-values presented in Table VIII confirm the above findings that the models are free from co-linearity problem.

We now discuss the modeling of $\log K_i$ (hCAXIV) using quantum and chemical (HyperChem) descriptors (Table III). Table V summarizes statistically significant models indicating that the maximum quantum and chemical descriptors that can be used for modeling logKi(hCAXIV)is seven. We observed that the quality of this 7-descriptor model is much inferior to the 4-descriptor model using PRECLAV descriptors. However, even the 7-descriptor model is free from defect due to co-linearity (Table VIII). Further examination of these models indicted 5-,6-, and 7-descriptor model contain one or two correlating descriptors in that their standard deviations were much larger than the respective coefficients. Such models are not allowed statistically and need not to be considered. The deletion of these compounds from the regression analysis gave a 4-descriptor model whose quality is very much poor compared to the 4-descriptor model using PRECLAV descriptors.

$$\begin{aligned}\log K_i(\text{hCA XIV}) = &-4.4337\\ &+ 0.0935(\pm 0.0208)\chi\\ &- 0.0753((\pm 0.0273)\text{DMy}\\ &- 0.0082((\pm 0.0017)\text{SA}\\ &-_{0.0050(0.0016)}\, H_f \end{aligned} \qquad (13)$$

$$n = 30,\ s = 0.2323,\ r2 = 0.6694,\ F = 12.657$$

A careful examination of this model indicated that it contains two outliers (compounds **6** and **7**).Deletion of these compounds from the process of regression yielded a model with improved statistics:

$$\begin{aligned}\log K_i(\text{hCA XIV}) = &-4.1311\\ &+ 0.0935(\pm 0.0179)\chi\\ &- 0.0688((\pm 0.0235)\text{DMy}\\ &- 0.0088((\pm 0.0015)\text{SA}\\ &-_{0.0051(0.0014)}\, H_f \end{aligned} \qquad (14)$$

It is clear, therefore, that in the present case the PRECLAV descriptors are far superior descriptors compare to the quantum and chemical descriptors for modeling logKi(hCAXIV). This is conformed by the correlations of observed activity with the estimated activity using PRECLAV (Figure 1) and quantum and chemical parameters (Figure 2).

Table IX. Significant molecular virtual fragments for the Set of 28 compounds. Molecules of analyzed database include 32 virtual fragments. The mass percents of 4 fragments are significant. THE MOST SIGNIFICANT VIRTUAL FRAGMENTS by square correlation of 'the mass percent of fragment' and 'Property values' are

Fragment atoms	Specimen (Com.no.)	Correlation sign	Square correlation
H_2N	1	1	.3701
C_2HN_3S	2	−1	.223
C_6H_4	4	1	.1751
HN	20	−1	.1445

Correlation sign 1 means 'large mass percent of this fragment increase the Property value'. Correlation sign − 1 means 'large mass percent of this fragment decrease the Property value'.

Conclusions

The QSAR studies of the inhibitory activity over carbonic anhydrase (hCAXIV isozyme) using 30 sulfonamides lead to the following conclusions:

– the PRECLAV descriptors are more useful than HyperChem descriptors
– the inhibitory activity is disfavored by C3HN3S and NH fragments

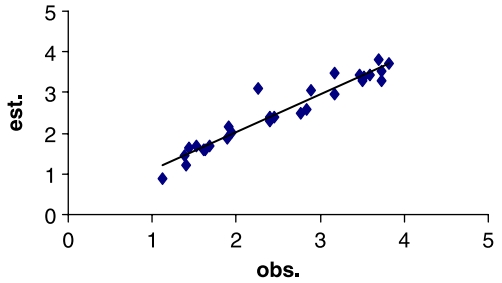


Figure 1. Correlation of observed and calculated activity [log K_i (hCA XIV)] using PRECLAV descriptors (pca, ray, pon, and lhf).

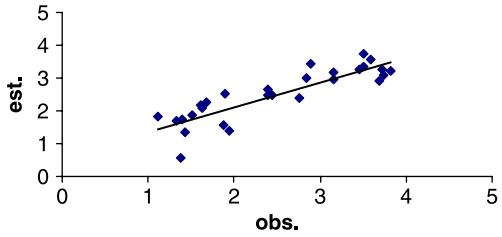


Figure 2. Correlation of observed and calculated activity [log K_i (hCA XIV)] using Quantum chemical (HyperChem) descriptors (Hf, SA, X, Debys Y).

– the inhibitory activity is favored by H2N and C6H4 fragments.
– the obtained results by using descriptors of PRECLAV have been compared with those in which descriptors have been calculated with HYPERCHEM. The obtained QSAR equations pointed to the fact that the CA inhibitory activity decreased for unsubstituted (at the organic scaffold) aromatic/heteroaromatic sulfonamides, but was favored by the presence of alkyl groups substituting the scaffold, which led to a higher internal topological diversity, as well as by the presence of condensed aromatic rings in the structure of these enzyme inhibitors.

Acknowledgements

One of the authors (S. Singh) expresses her thanks to the Department of Science & Technology, Government of India, New Delhi, for awarding DST project SR / WOS-A / CS-61 / 2004 under Woman Scientists scheme, and to Dr. R.P.Singh, Principal, Bareilley College, Bareilley, India for his personal interest,encouragement and for providing facilities for carrying out the present study.The project was also in part financed by an EU grant to CTS and AS of the 6th framework programme (DeZnIT project).

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

[1] Supuran CT. Carbonic anhydrases: Novel therapeutic applications for inhibitors and activators. Nature Rev Drug Disc 2008;7:168–181.
[2] Mincione F, Menabuoni L, Supuran CT. Clinical applications of the Carbonic anhydrase inhibitors in ophthalmology. In: Supuran C, Scozzafava A, Conway J, editors. Carbonic anhydrase, its inhibitors and activators. Boca Raton (FL): CRC Press; 2004. p 243–254.
[3] Supuran CT, Scozzafava A, Conway J, editors. Carbonic anhydrase–its inhibitors and activators. Boca Raton (USA): CRC Press; 2004. p 1–363.
[4] Clare BW, Supuran CT. QSAR studies of sulfonamide carbonic anhydrase inhibitors. In: Supuran CT, Scozzatava A, Conway J, editors. Carbonic andydrase, its inhibitors and activators. Boca Raton, FI, USA: CRC Press; 2004.
[5] Supuran CT. Carbonic anhydrases: Catalytic and inhibition mechanisms, distribution and physiological roles. In: Pacscas I, editor. Carbonic anhydrase and modulation of physiologic and pathologic process in the organism. Timisoara (Romania): Helicon; 1994. p 29–111.
[6] Supuran CT, Scozzafava A, Casini A. Development of sulfonamide carbonic anhydrase inhibitors (CAIs). In: Supuran CT, Scozzafava A, Conway J, editors. Carbonic anhydrase, its inhibitors and activators. Boca Raton (FL): CRC Press; 2004. p 67–148.
[7] Supuran CT, Scozzafava A, Casini A. Carbonic anhydrase inhibitors. Med Res Rev 2003;23:146–189.
[8] Fujikawa-adachi K, Nishimori I, Taguchi T, Onishi S. J Biol Chem 1999;274:21228.
[9] Asida S, Nishimori I, Tanimura M, Onishi S, Shuri T. J Cancer Res Clin Oncol 2002;128:561.
[10] Clare BW, Supuran CT. A perspective on quantitative structure-activity relationships and carbonic anhydrase inhibitors. Expert Opin Drug Metab Toxicol 2006;2:113.
[11] Singh J, Singh S, Khadikar PV, Supuran CT, Clare BW, Lakhwani M. Estimation of human carbonic anhydrase II inhibition using topological indices and their combination with quantum- theoretical descriptors. Med Chem 2008;4:30–66.
[12] Supuran CT, Scozzafava A, Conway J, editors. Carbonic anhydrase–its inhibitors and activators. Boca Raton (USA): CRC Press; 2004. p 1–363.
[13] Agrawal VK, Singh J, Pandey A, Khadikar PV. Modeling of inhibitory activities of sulfonamide schiff bases using physicochemical parameters. Oxid Commun 2006;29: 803–816.
[14] Tarko L. QSAR studies regarding the inihibition of the carbonic anhydrase by the sulfonamides containing a picolinyl group. Rev Chim (Bucuresti) 2007;58:191–194.
[15] Khadikar PV, Clare BW, Balaban AT, Supuran CT, Agrawal VK, Singh J, Joshi AK, Lakhwani M. QSAR predection of CAI, CAII, CAIV inhibitory activities: Relative potential of balaban and balaban type indices. Rev Roum Chem 2006;51: 703–717.
[16] Jaiswal M, Khadikar PV, Scozzafava A, Supuran CT. Carbonic anhydrase inhibitors: The first QSAR study on inhibition of tumour associated isoenzyme IX with aromatic and heterocyclic sulfonamides. Bioorg Med Chem Lett 2004;14: 3283–3290.
[17] Khadikar PV, Diudea MV, Singh J, Shrivastva A, Singh S, Karmarkar S, Lakhwani M, Thakur P. Use of PI index in computer-aided designing of bioactive compounds. Curr Bioact Comp 2006;2:19–56.
[18] Agrawal VK, Bano S, Supuran CT, Khadikar PV. QSAR study on carbonic anhydrase inhibitors: Aromatic/heterocyclic sulfonamides containing 8-quinoline moieties with topical activity as anti-glauoma agents. Eur J Med Chem 2004;39:593–600.

[19] Khadikar PV, Sharma V, Karmarkar S, Supuran CT. QSAR studies on benzene sulfonamide carbonic anhydrase inhibitors: Need of hydrophobic parameter for topological modeling of binding constants of sulfonamides to human CA-II. Bioorg Med Chem Lett 2005;15:923–930.

[20] Khadikar PV, Clare BW, Balaban AT, Supuran CT, Agrawal VK, Singh J, Joshi AK, Lakhwani M. QSAR predection of CAI,CAII, CAIV inhibitory activities: Relative potential of balaban and balaban type indices. Rev Roum Chem 2006;51: 703–717.

[21] Khadikar PV, Jaliwala YA, Lakhwani M, Thakur P, Joshi S, Joshi AK. QSAR studies on inhibitoty properties of benzinesulphonamides towards CA-I: Domaminating role of van der Waals repulsion energy. Bioinformatics Trends 2006;1: 125–136.

[22] Khadikar PV, Clare BW, Balaban AT, Supuran CT, Agrawal VK, Singh J, Joshi AK, Lakhwani M. Estimation of human carbonic anhydrase II inhibition using topological indices and their combination with quantum- theoretical descriptors. J Med Chem Res 2006;15:156–160.

[23] Agrawal VK, Singh J, Pandey A, Khadikar PV. Modeling of inhibitory activities of sulfonamide schiff bases using physico-chemical parameters. Oxid Commun 2006;29:803–816.

[24] Khadikar PV, Deeb O, Jaber A, Singh J, Lakhwani M. Development of quantitative structure- activity relationship for a set of carbonic anhydrase inihibitors: Use of quantum and chemical descriptors. Lett Drug Design Discov 2006;3: 622–635.

[25] Balaban AT, Khadikar PV, Supuran CT, Thakur A, Thakur M. Study on supramolecular complexing ability vis-à-vis estimation of pKa of substituted sulfonamides: Dominating role of balaban index (J). Bioorg Med Chem Lett 2005;15: 3966–3973.

[26] Nishimori I, Vullo D, Innocenti A, Scozzafava A, Mastrolorenzo A, Supuran CT. Carbonic anhydrase inihibitors: Inihibition of the transmembrane isozyme XIV with sulfomamides. Bioorg Med Chem Lett 2005;15:3828–3833.

[27] Tarko L, Lupescu I, Groposila-Constantinescu D. Sweetness power QSARs by PRECLAV software. Arkivoc 2005;10: 254–271.

[28] PRECLAV v.0601 program is available from Center of Organic Chemistry (CCO) Bucharest, Romanian Academy; Director CCO E-mail address: pfilip@cco.ro

[29] HyperChem7.52 Software, Hypercube Inc. Florida Science and Technology Park,1115 N.W. 4th street Gainesvilla Florida, 326001, USA; 2006. http://www.hyper.com/

[30] ADDREIT T., STANCU S., Statistica, Ed. ALL, Bucuresti, 1995, p 341.

[31] Sybyl 7.2 Tripos Inc., 1699 Hanley Road, St.Louis MO 63144see www.tripos.com

[32] Vinter JG, Davis A, Saunders MR. J Comp-Aided Design Mol 1987;1:31.

[33] Stewart JJP. Mopac93, Fujitsu Ltd., Tokyo, Japa.

[34] Charterjee S, Hadi AS, Price B. Regression analysis by examples. 3rd ed. New York: Wiley; 2000.

[35] NCSS(statistical Software Delux package), 329 North 1000 East; Kaysville, UT, USA, 2004.