Revue Roumaine de Chimie, **2006**, *51*(7-8), 691–701

Dedicated to the memory of
Professor Mircea D. Banciu (1941–2005)

QSAR STUDY ON THE INHIBITION OF THE HUMAN CARBONIC ANHYDRASE CYTOSOLIC ISOZYME VII

Jyoti SINGH,[a] Meenakshi LAKHWANI,[b] Padmakar V. KHADIKAR,[c*] Alexandru T. BALABAN, [d] Brian W. CLARE[e] and Claudiu T. SUPURAN[f]

[a]Department of Chemistry, A.P.S. University, Rewa, 486 003, India, e-mail: jyoti_singh07@rediffmail.com
[b]Department of Chemistry, Holkar Model & Autonomous College, Indore, India, e-mail: meenakashi@yahoo.com
[c]Research Division, Laxmi Fumigation and Pest Control, Pvt. Ltd., 3, Khatipura, Indore 452 007, India, pvkhadikar@rediffmail.com
[d]Texas A&M University at Galveston,5007 Avenue U, Galveston, TX 77551,USA, e-mail: balabana@tamug.edu
[e]School of Biomedical and Chemical Sciences, University of Western Australia, 35, Stirling Highway Crawltywa 6009, Australia, e-mail: bwc@crystal.uwa.edu.au
[f]Laboratorio di Chimica Bioinorganica, Departmento di Chimica, University of Florence, via della Lastruccia, 3, RM-188, Polo Scientifico, 50019 Sesto Fioventino, Firenze, Italy, e-mail: claudiu.supuran@unifi.it

Received April 13, 2006

The inhibition of a newly cloned human carbonic anhydrase (CA, EC 4-2-1-1) isozyme VII (hCAII) by a series of 32 aromatic and heterocyclic sulfonamides, including some of the clinically used derivatives (acetazolamide, methazolamide, ethoxzolamide, dichlorophenamide, dorzolamide, brinzolamide, and benzolamide, as well as the sulfamate antiepileptic drug topiramate), has been modeled using a set of distance-based topological indices and structural descriptors. The results have shown that the Balaban-type indices play a dominant role. Excellent results were obtained in multi-parametric regressions. The results are discussed with a variety of statistical indicators.

INTRODUCTION

Recently it has been reported that the human carbonic anhydrase (CA, EC 4.2.1.1) isozyme VII is inhibited by aromatic and heterocyclic sulfonamides,[1] the classical CA inhibitors with clinical applications.[2,3] This is one of the recently known 15 isozymes present in humans. It is worthy to mention that out of these 15 isozymes, twelve are catalytically active: CAs I-Va, CA-Vb, CA-VI, CA-VII, CA-IX and CAs XII-XIV; whereas the CA-related proteins: CARPs VIII, X and XI are devoid of catalytic activity due to the fact that one or more histidine zinc ligand are replaced by other amino acid derivatives.[2,3] Among the active isozymes, five show a cytosolic subcellular localization: CAs I-III, CA-VII and CA-XIII [4-14]. The isozyme CA-VII appears to be the less studied and understood among the cytosolic CAs; CA-VII has been shown to be highly expressed in the brain, by *in situ* hybridization.[15,16] In view of the above, we have undertaken the present investigation in that we have modeled inhibition of hCA-VII with aromatic and heterocyclic sulfonamides (Table 1) using distance-based topological indices and structural descriptors (Tables 2 – 4). The inhibition data as reported by Vullo et al.[1] have been used after converting them into logarithmic units *i.e.* as $\log K_i$(hCA-VII), and the experimental data are presented in Table 2. Among the 25 descriptors used by us, 14 are distance-based topological indices, of which 6 are Balaban (J)[17] and Balaban-type (J_{hetZ}, J_{hetM}, J_{hetV}, J_{hetE}, J_{hetP},) indices.[18-21] Statistical treatment of data for QSAR followed well-known procedures.[22-25]

* Corresponding author: Tel: +91-731-2531906, E-mail: pvkhadikar@rediffmail.com

Table 1

Structural details of carbonic anhydrase-VII inhibitors used in the present investigation

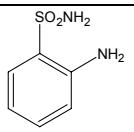
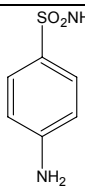
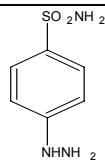
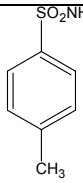
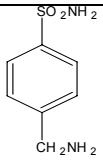
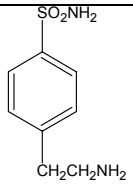
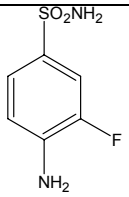
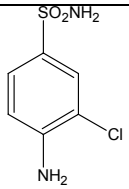
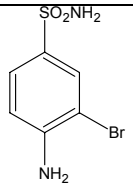
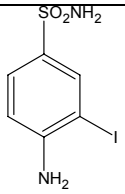
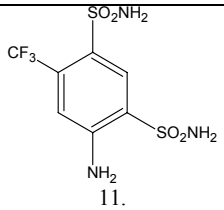
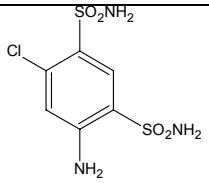
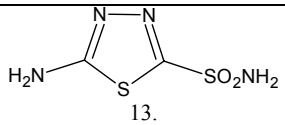
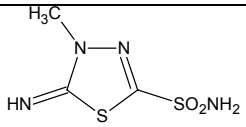
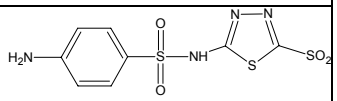
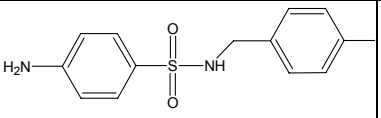
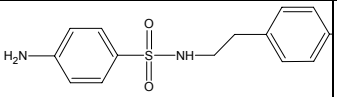
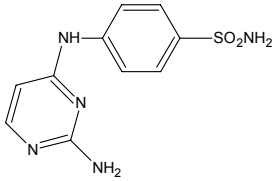
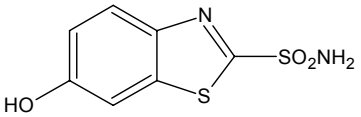
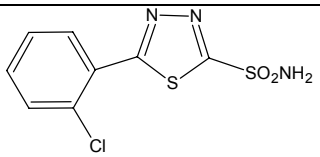
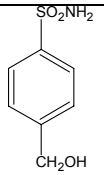
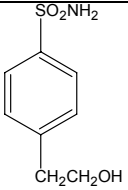
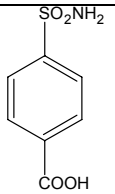
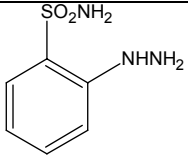
 1.	 2.	 3.
 4.	 5.	 6.
 7.	 8.	 9.
 10.	 11.	 12.
 13.	 14.	 15
 16.	 17.	 18.
 19.	 20.	 21.
 22.	 23.	 24.

Table 1 (continues)

Table 1 (continued)

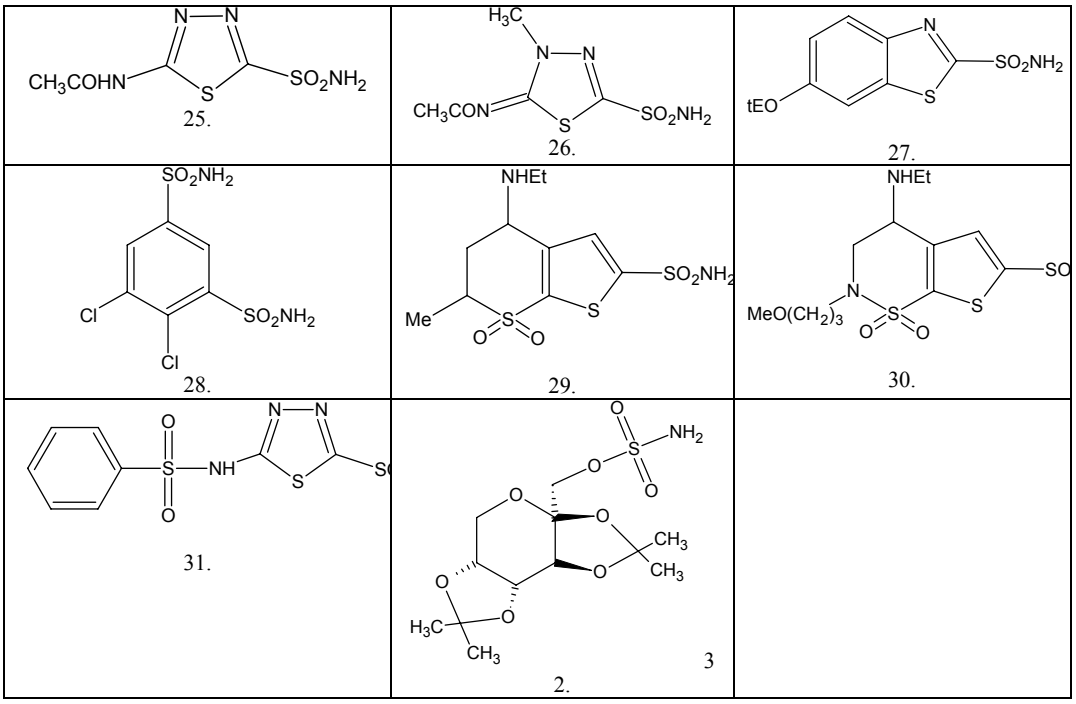


Table 2

Various topological descriptors used in the present study and their values

Comp.	logK$_i$(hCAVII)	W	$^0\chi$	$^1\chi$	$^2\chi$	$^0\chi^v$	$^1\chi^v$	$^2\chi^v$
1.	1.6532	144	8.4831	5.015	5.2343	5.688	2.2889	2.0828
2.	1.8450	152	8.4831	4.999	5.3226	5.688	2.8832	2.1140
3.	1.9493	201	9.1902	5.537	5.4919	6.1889	3.1332	2.2137
4.	1.9994	152	8.4831	4.9999	5.323	6.9280	4.6410	3.9190
5.	1.8750	201	9.1902	5.537	5.4919	6.3960	3.3563	2.3913
6.	1.9030	262	9.8973	6.037	5.8723	7.1031	3.8563	2.7276
7.	1.8750	189	9.3534	5.4097	5.8306	5.9895	2.9889	2.2283
8.	2.0791	189	9.3534	5.4097	5.8306	5.9895	2.9889	2.2283
9.	1.7853	189	9.3534	5.4097	5.8306	5.9895	2.9889	2.2283
10.	2.1760	189	9.3534	5.4097	5.8306	5.9895	2.9889	2.2283
11.	2.0000	624	15.2236	3.2430	10.3662	8.9701	4.3066	3.4740
12.	2.3222	399	12.7236	7.0317	8.3710	7.7142	3.6787	2.9105
13.	0.7160	113	7.7760	4.4990	4.9812	4.6821	2.1173	1.4512
14.	0.6334	146	8.6463	4.9097	5.4889	5.6048	2.5288	1.8417
15.	0.8450	853	14.9663	9.1825	9.8915	9.7663	4.9302	3.5927
16.	0.7481	1195	16.3805	10.1825	10.5877	11.4301	6.1533	4.5825
17.	0.8129	1408	17.0876	10.6825	10.9123	12.1372	6.6533	4.9508
18.	0.8325	669	13.1734	8.4885	8.4572	9.2380	4.9342	3.4945
19.	0.6020	287	10.3449	6.4653	6.9727	6.8368	3.5903	2.6768
20.	0.7323	434	11.759	7.4820	7.5620	10.181	6.8360	6.0180
21.	1.7781	201	9.1902	5.5370	5.4919	5.4919	6.6258	3.2643
22.	1.8915	262	9.9872	6.0370	5.8723	6.9729	3.7643	2.6626
23.	1.7160	252	10.0605	5.9097	6.2217	6.4670	3.2723	2.3745
24.	1.8325	185	9.1900	5.5554	5.4250	7.0050	4.6860	3.7150
25.	0.3979	257	10.605	5.8929	6.3716	6.5130	3.0317	2.0993
26.	0.3222	304	10.9307	6.3035	6.0013	7.4062	3.4272	2.4500
27.	-0.1079	442	11.7591	7.5033	7.5223	8.5050	4.5666	3.0873
28.	1.4232	398	12.7238	7.0370	8.3433	7.5146	3.5790	2.8072
29.	0.5440	634	14.4223	8.6749	9.3126	10.5814	5.8318	4.5548
30.	0.4471	1142	17.2507	10.7129	10.6302	12.9809	7.1044	5.1729
31.	-0.0604	731	14.0960	8.7890	9.2700	11.6660	8.5770	7.7320
32.	-0.3407	799	15.6820	9.4170	11.7860	13.0007	7.9950	7.7320

W = Wiener index [26]; $^0\chi$, $^1\chi$, $^2\chi$ = zero-, first-, and second-order Randić connectivity indices [27]; $^0\chi^v$, $^1\chi^v$ and $^2\chi^v$ = zero-, first- and second-order Kier and Hall's valence connectivity indices [28,29].

Table 3

Structural details of the sulfonamides, their Balaban and Balaban-type indices

Comp.	J	J_{hetZ}	J_{hetM}	J_{hetV}	J_{hetE}	J_{hetP}	BAC
1.	2.545	4.788	4.789	3.092	3.614	3.531	18
2.	2.394	4.425	4.426	2.953	3.415	3.438	18
3.	2.359	4.121	4.122	2.695	3.331	2.884	21
4.	2.349	4.372	4.373	3.032	3.383	3.487	18
5.	2.359	4.000	4.009	2.895	3.257	3.216	21
6.	2.305	3.632	3.633	2.791	3.672	3.046	22
7.	2.512	4.526	4.540	2.878	3.592	3.083	27
8.	2.512	4.645	4.652	3.101	3.568	3.509	27
9.	2.512	4.720	4.730	3.105	3.553	3.562	27
10.	2.512	4.745	4.754	3.178	3.523	3.623	27
11.	3.069	5.305	5.326	3.299	4.188	3.500	110
12.	2.853	5.633	5.640	3.456	4.014	4.073	69
13.	2.449	6.035	6.039	2.416	3.218	2.913	18
14.	2.538	5.769	5.772	2.360	3.437	2.654	27
15.	1.861	3.779	3.781	1.848	2.401	2.159	38
16.	1.736	2.828	2.828	1.882	2.650	2.068	38
17.	1.686	2.549	2.549	1.788	2.108	1.943	38
18.	1.713	2.650	2.650	1.837	2.394	1.827	11
19.	1.987	3.926	3.955	2.610	2.749	2.583	18
20.	1.923	3.801	3.805	2.301	2.628	2.680	18
21.	2.347	2.359	4.036	4.030	3.276	3.120	21
22.	2.305	3.651	3.652	2.737	3.086	2.973	22
23.	2.392	4.000	4.000	2.858	3.340	3.096	30
24.	2.587	4.612	4.631	2.869	3.623	3.088	21
25.	2.364	4.553	4.555	2.161	3.007	2.372	31
26.	2.515	5.015	5.015	2.347	3.376	2.526	42
27.	1.958	3.737	3.736	2.109	2.696	2.331	21
28.	2.859	5.814	5.825	3.511	4.035	4.217	69
29.	2.222	3.885	3.886	2.349	2.690	2.742	55
30.	2.163	3.109	3.104	1.997	2.405	2.169	27
31.	1.867	3.902	3.909	1.637	2.549	2.113	65
32.	1.915	2.528	5.525	1.245	2.357	1.192	52

J = Balaban distance-based connectivity index [17]; J_{hetZ} = Balaban-type index from Z-weighted distance matrix (Barysz matrix) [21]; J_{hetM} = Balaban-type index from mass weighted distance matrix; J_{hetV} = Balaban-type index from van der Waals weighted distance matrix; J_{hetE} = Balaban-type index from electro negativity weighted distance matrix; J_{hetP} = Balaban-type index from polarizability weighted distance matrix 18-20]; BAC = Balaban centric index.

Table 4

Values of physicochemical parameters calculated for compounds used in the present study

Comp.	MW	MR	MV	Pol	IR	ST	d	PC	MIM	NM	AM
1.	172.204	42.80	120.6	340.6	1.627	63.7	1.427	16.97	172.03	172	172.02
2.	172.204	42.80	120.6	340.6	1.627	63.7	1.427	16.97	172.03	172	172.02
3.	187.219	46.46	125.0	369.2	1.665	76.0	1.497	18.41	187.04	187	187.22
4.	171.21	43.81	134.6	351.2	1.564	46.3	1.271	17.36	171.03	171	171.21
5.	186.23	47.43	138.3	381.0	1.601	57.4	1.345	18.80	186.04	186	186.23
6.	200.25	52.07	154.6	421.1	1.587	54.6	1.293	20.64	200.06	200	200.26
7.	190.19	42.92	124.8	348.3	1.603	60.5	1.523	17.01	190.02	190	190.19
8.	206.64	47.63	132.6	378.0	1.637	66.0	1.558	18.88	205.99	206	206.65
9.	251.10	50.52	136.8	392.0	1.660	67.3	1.834	20.03	249.94	250	251.10
10.	298.10	55.73	147.0	414.9	1.709	71.3	2.088	22.09	297.92	298	298.10
11.	319.28	60.23	183.1	510.5	1.570	60.4	1.743	23.87	318.99	319	319.28
12.	285.72	60.07	161.5	485.7	1.666	81.6	1.768	23.81	284.96	285	285.72
13.	180.20	36.54	96.5.0	316.9	1.681	115.9	1.865	14.48	179.97	180	180.02
14.	194.24	42.04	95.2.0	306.5	1.835	107.4	2.030	16.66	193.99	194	194.23
15.	335.28	73.72	185.7	596.6	1.724	106.7	1.805	29.22	337.98	335	335.30
16.	340.41	85.62	245.9	671.3	1.613	56.5	1.384	33.94	340.05	340	340.42
17.	354.44	90.25	262.4	711.4	1.603	53.9	1.350	35.77	354.07	354	354.44

Table 4 (continues)

Table 4 (continues)

18.	250.24	63.14	170.8	504.4	1.66	75.9	1.464	25.03	250.05	250	20.28
19.	230.28	54.21	133.6	413.5	1.745	91.7	1.723	21.49	229.80	230	230.26
20.	275.50	62.46	171.5	498.1	1.648	71.1	1.607	24.76	274.95	275	25.73
21.	187.80	45.35	132.1	368.2	1.602	60.2	1.416	17.97	187.03	187	187.21
22.	201.24	49.98	148.7	408.3	1.587	56.8	1.353	19.01	201.45	201	201.29
23.	201.19	45.46	130.9	373.4	1.611	66.2	1.536	18.02	201.00	201	201.2
24.	187.21	46.46	125.0	369.2	1.666	76.0	1.497	18.41	187.04	187	187.22
25.	222.24	45.95	127.3	400.7	1.640	97.9	1.744	18.21	221.98	222	222.24
26.	236.27	53.66	131.9	394.6	1.737	80.0	1.790	21.03	236.00	236	236.27
27.	244.29	59.05	159.2	450.5	1.864	67.8	1.534	23.41	243.99	244	244.20
28.	305.15	61.28	171.2	494.4	1.634	69.7	1.782	24.29	303.91	304	305.15
29.	322.44	75.48	211.0	609.0	1.634	69.9	1.530	29.29	324.01	324	324.40
30.	320.36	70.10	183.4	568.9	1.689	92.5	1.746	27.79	319.97	320	320.36
31.	365.46	87.80	255.5	717.0	1.603	62.3	1.430	34.80	385.10	365	365.47
32.	325.33	69.69	217.1	596.2	1.553	56.2	1.490	27.63	325.08	325	325.89

where: MW = Molecular weight; MR = Molar refractivity; MV = Molecular volume;
Pol = polarizability; IR = index of refraction; ST = surface tension; d = density; PC = Parachor; MIM = Monoisotopic Mass; NM = Nominal Mass; AM = Average Mass.[35]

RESULTS AND DISCUSSION

The set of 32 CA-VII inhibitors is presented in Table 1. For these compounds, values of distance-based topological indices: Wiener's index (W);[26] Randić-Kier-Hall indices ($^{0}\chi$, $^{1}\chi$, $^{2}\chi$, $^{0}\chi^{v}$, $^{1}\chi^{v}$, $^{2}\chi^{v}$);[27-29] Balaban index *J;* [17] Balaban-type indices (J_{hetZ}, J_{hetM}, J_{hetV}, J_{hetE}, J_{hetP}, BAC);[18-21] and structural descriptors (MW, MR, PC, IR, ST, density, Pol, MIM, NM, AM) are presented in Tables 2, 3, and 4, respectively. Further details on these molecular descriptors are provided in footnotes of these Tables.

The inhibition activity versus $\log K_i$(hCA-VII) adopted from our earlier work, shown in Table 2, is first tested in a monoparametric regression analysis (Table 5). The results show that out of 25 descriptors used by us, J_{hetV} (boldface characters) yields the best statistically significant mono-parametric model:

$$\log K_i(\text{hCA-VII}) = -0.901 + 0.827(\pm 0.112)\, J_{hetV} \quad (1)$$
$$n = 32,\ \text{SE} = 0.445,\ \text{R} = 0.824,\ \text{F} = 63.481,\ \text{Q} = 1.851$$

where n is the number of sulfonamidic carbonic anhydrase VII inhibitors, SE is the standard error, R is the correlation factor, F is the Fisher criterion, and Q is the quality factor of the correlation.

Table 5

Monoparametric correlations between the 21 parameters and $\log K_i$(hCA-VII) *

Model	Parameter	SE	R	F	Q
1.	W	0.683	-0.495	9.748	-0.725
2.	$^{1}\chi$	0.624	-0.608	17.620	-0.975
3.	$^{2}\chi$	0.675	-0.512	10.633	-0.759
4.	$^{1}\chi^{v}$	0.673	-0.517	10.926	-0.769
5.	$^{2}\chi^{v}$	0.661	-0.541	12.392	-0.818
6.	*J*	0.616	0.621	18.809	1.008
7.	J_{hetZ}	0.749	0.303	3.040	0.404
8.	$\boldsymbol{J_{hetV}}$	**0.445**	**0.824**	**63.481**	**1.851**
9.	J_{hetE}	0.532	0.737	35.587	1.386
10.	J_{hetP}	0.459	0.812	57.938	1.770
11.	MW	0.698	-0.460	8.037	-0.660
12.	MR	0.680	-0.502	10.128	-0.739
13.	MV	0.710	-0.429	6.773	-0.604
14.	PC	0.679	-0.503	10.154	-0.740
15.	RI	0.735	-0.353	4.281	-0.480
16.	ST	0.743	-0.328	3.610	-0.441
17.	d	0.781	-0.112	0.382	-0.144
18.	Pol	0.679	-0.503	10.140	-0.740
19.	MIM	0.695	-0.468	8.414	-0.674
20.	NM	0.698	-0.461	8.091	-0.660
21.	AM	0.763	-0.242	1.859	-0.317

SE is the standard error, R is the correlation coefficient, F is the Fisher criterion, and Q is the quality factor of the correlation.

In order to go further, we first examined intercorrelations between the 21 molecular descriptors. Table 6 indicates that J_{hetV} (boldface characters in the corresponding row and column) is strongly correlated with J_{hetP} (an index that will be excluded from further multiparametric models involving J_{hetV}) fairly-correlated with J_{hetE}, and weakly correlated with the other two Balaban-type indices, J_{hetM} and J_{hetZ}. Also, it is weakly correlated with all the other descriptors.

Table 6

Intercorrelation matrix for molecular descriptors applied to $\log K_i$(hCA-VII)

	logCA	W	$^0\chi$	$^1\chi$	$^2\chi$	$^0\chi^v$	$^1\chi^v$	$^2\chi^v$	J	J_{hetZ}	J_{hetM}	J_{hetV}	J_{hetE}
logCA	1												
W	-0.496	1											
$^0\chi$	-0.499	0.951	1										
$^1\chi$	-0.609	0.878	0.827	1									
$^2\chi$	-0.512	0.915	0.981	0.789	1								
$^0\chi^v$	-0.6	0.91	0.933	0.876	0.923	1							
$^1\chi^v$	-0.517	0.712	0.721	0.741	0.721	0.859	1						
$^2\chi^v$	-0.54	0.667	0.697	0.694	0.731	0.875	0.946	1					
J	0.62	-0.602	-0.45	-0.745	-0.441	-0.604	-0.613	-0.56	1				
J_{hetZ}	0.304	-0.57	-0.464	-0.643	-0.438	-0.574	-0.695	-0.55	0.784	1			
J_{hetM}	0.158	-0.545	-0.391	-0.584	-0.293	-0.455	-0.456	-0.304	0.743	0.801	1		
J_{hetvV}	**0.824**	**-0.637**	**-0.57**	**-0.707**	**-0.575**	**-0.704**	**-0.511**	**-0.585**	**0.773**	**0.441**	**0.387**	**1**	
J_{hetE}	0.737	-0.674	-0.56	-0.8	-0.545	-0.694	-0.66	-0.608	0.938	0.732	0.656	**0.835**	1
J_{hetP}	0.811	-0.629	-0.551	-0.686	-0.55	-0.67	-0.609	-0.589	0.824	0.663	0.484	**0.915**	0.88
BAC	-0.055	0.346	0.551	0.089	0.587	0.383	0.248	0.32	0.35	0.257	0.341	**0.044**	0.244
MW	-0.46	0.83	0.901	0.739	0.903	0.838	0.664	0.683	-0.37	-0.279	-0.228	**-0.484**	-0.46
MR	-0.503	0.909	0.906	0.863	0.884	0.9	0.766	0.753	-0.597	-0.496	-0.49	**-0.595**	-0.627
MV	-0.43	0.885	0.884	0.807	0.876	0.897	0.799	0.8	-0.565	-0.544	-0.478	**-0.547**	-0.576
PC	-0.503	0.913	0.925	0.853	0.913	0.919	0.791	0.787	-0.577	-0.5	-0.458	**-0.601**	-0.623
IR	-0.354	-0.115	-0.123	0.013	-0.146	-0.159	-0.268	-0.307	-0.061	0.271	0.105	**-0.157**	-0.104
ST	-0.328	-0.085	-0.068	-0.019	-0.073	-0.159	-0.283	-0.29	0.051	0.401	0.297	**-0.205**	-0.103
d	-0.112	-0.141	-0.039	-0.172	-0.019	-0.187	-0.339	-0.296	0.369	0.589	0.544	**0.077**	0.217
Pol	-0.503	0.911	0.906	0.863	0.885	0.9	0.767	0.755	-0.599	-0.495	-0.487	**-0.595**	-0.625
MIM	-0.468	0.824	0.894	0.736	0.895	0.837	0.675	0.697	-0.378	-0.277	-0.23	**-0.493**	-0.466
NM	-0.46	0.83	0.901	0.739	0.903	0.839	0.664	0.683	-0.37	-0.279	-0.229	**-0.486**	-0.463
AM	-0.241	0.562	0.616	0.433	0.617	0.491	0.337	0.344	-0.02	-0.037	0.047	**-0.21**	-0.13

	J_{hetP}	BAC	MW	MR	MV	PC	IR	ST	d	Pol	MIM	NM	AM
J_{hetP}	1												
BAC	0.12	1											
MW	-0.4	0.628	1										
MR	-0.538	0.449	0.928	1									
MV	-0.505	0.486	0.885	0.97	1								
PC	-0.548	0.493	0.929	0.989	0.98	1							
IR	-0.116	-0.237	-0.047	-0.099	-0.299	-0.19	1						
ST	-0.145	-0.109	-0.028	-0.168	-0.348	-0.169	0.641	1					
d	0.169	0.216	0.187	-0.123	-0.278	-0.154	0.59	0.72	1				
Pol	-0.539	0.448	0.929	0.999	0.97	0.989	-0.097	-0.164	-0.118	1			
MIM	-0.406	0.629	0.999	0.931	0.89	0.934	-0.051	-0.03	0.175	0.933	1		
NM	-0.402	0.628	0.999	0.928	0.886	0.929	-0.047	-0.027	0.186	0.929	0.999	1	
AM	-0.141	0.64	0.709	0.62	0.609	0.626	-0.042	-0.037	0.185	0.62	0.712	0.71	1

In order to raise its correlation factor with $\log K_i$(hCA-VII) above 0.9, we tried multiparametric correlations, using J_{hetV} and other descriptors. We tried the descriptors that are next in line in the

monoparametric correlation, namely the index of refraction IR, the Balaban-type indices J_{hetE}, J_{hetM}, J_{hetZ}, the surface tension ST, the index of refraction IR, and the density d.

During the stepwise regression analysis, we observed in biparametric corretaions that the addition of the IR term to eq (1) resulted into a model (Table 7, model 59, boldface characters) with significantly improved quality:

$$\log K_i(\text{hCAVII}) = 0.137 + 1.034(\pm 0.117)\, J_{hetV} - 0.975\ (\pm 0.302)\ \text{IR} \quad (2)$$
$$n = 32,\ SE = 0.423,\ R = 0.855,\ R^2A = 0.713,\ F = 40.754$$

We now comment on adjustable-R^2 (denoted as R^2A). This value takes into account the adjustment of R^2. Therefore, if a variable is added that does not contribute its fair share, R^2A will actually decline. It also takes into account the relationship between sample size and the number of variables: R^2 may appear artificially high if the number of variables is high compared with the sample size. That is, R^2 will always increase when an independent variable is added, and R^2A will decrease if the added variables do not reduce the unexplained variation enough to offset the loss of degrees of freedom.

Table 7

Bi-parametric models

Model	Parameters	SE	R	R^2A	F
22	J, $^1\chi$	0.602	0.657	0.393	11.036
23	J_{hetZ}, $^1\chi$	0.628	0.619	0.340	9.000
24	J_{hetM}, $^1\chi$	0.604	0.655	0.390	10.894
25	J_{hetV}, $^1\chi$	0.452	0.825	0.658	30.870
26	J_{hetE}, $^1\chi$	0.540	0.737	0.512	17.266
27	J_{hetP}, $^1\chi$	0.463	0.815	0.641	28.648
28	$^1\chi$, $^2\chi$	0.633	0.610	0.329	8.613
29	$^1\chi^v$, $^2\chi$	0.665	0.554	0.259	6.426
30	J, $^2\chi$	0.590	0.675	0.418	12.130
31	J_{hetZ}, $^2\chi$	0.683	0.519	0.219	5.350
32	J_{hetM}, $^2\chi$	0.687	0.512	0.211	5.141
33	J_{hetV}, $^2\chi$	0.451	0.825	0.659	30.987
34	J_{hetE}, $^2\chi$	0.530	0.748	0.530	18.449
35	J_{hetP}, $^2\chi$	0.463	0.815	0.642	28.771
36	W, $^1\chi^v$	0.669	0.546	0.250	6.173
37	J, $^1\chi^v$	0.611	0.644	0.375	10.291
38	J_{hetZ}, $^1\chi^v$	0.682	0.522	0.223	5.444
39	J_{hetM}, $^1\chi^v$	0.681	0.524	0.225	5.489
40	J_{hetV}, $^1\chi^v$	0.444	0.831	0.670	32.477
41	J_{hetE}, $^1\chi^v$	0.540	0.938	0.513	17.317
42	J_{hetP}, $^1\chi^v$	0.466	0.812	0.636	28.097
43	W, $^2\chi^v$	0.657	0.570	0.278	6.982
44	J, $^2\chi^v$	0.598	0.663	0.401	11.376
45	J_{hetV}, W	0.452	0.825	0.658	30.886
46	J_{hetV}, $^0\chi$	0.452	0.825	0.658	30.854
47	J_{hetV}, $^0\chi^v$	0.452	0.825	0.658	30.794
48	J_{hetV}, $^1\chi^v$	0.444	0.831	0.670	32.477
49	J_{hetV}, $^2\chi^v$	0.449	0.827	0.662	31.402
50	J_{hetV},, J	0.452	0.824	0.658	30.770
51	J_{hetV}, J_{hetM}	0.450	0.827	0.662	31.329
52	J_{hetV}, J_{hetM}	0.431	0.842	0.689	35.397
53	Jhetv, J_{hetE}	0.447	0.829	0.665	31.821
54	J_{hetV}, J_{hetP}	0.438	0.836	0.679	33.767
55	J_{hetV}, MW	0.449	0.827	0.662	31.375
56	J_{hetV}, MR	0.453	0.824	0.657	30.717
57	J_{hetV}, MV	0.452	0.824	0.658	30.771
58	J_{hetV}, PC	0.453	0.824	0.657	30.694
59	**J_{hetV}, IR**	**0.415**	**0.855**	**0.712**	**39.338**
60	J_{hetV}, ST	0.434	0.840	0.685	34.748
61	J_{hetV}, d	0.431	0.843	0.690	35.498
62	J_{hetV}, MIM	0.449	0.827	0.662	31.411
63	J_{hetV}, NM	0.449	0.827	0.662	31.372
64	J_{hetV}, AVM	0.449	0.827	0.662	31.380

In both eqs (1) and (2), the coefficient of J_{hetV} is positive. This means that the increase in van der Waals weighted distance favors the correlation with $logK_i$(hCA-VII). The coefficient of IR on the other hand is negative; showing that a decrease in the index of refraction is favorable for the activity.

In the following we present an abbreviated description of multiparametric correlations that were tested sequentially, numbering separately equations and models. The results for triparametric correlations are displayed in Table 8. Surprisingly, although eq. (2) is the best biparametric model, adding to J_{hetV} and IR a third parameter (models 65 – 67 in Table 8) afforded lower R values than the model 68 of Table 8 (boldface characters) with three Balaban-type indices.

Table 8

Tri- and tetra-parametric models

Model	Parameters	SE	R	R^2A	F	Q
65	J_{hetV}, IR , J_{hetM}	0.408	0.865	0.722	27.842	2.120
66	J_{hetV}, IR, ST	0.423	0.855	0.703	25.404	2.021
67	J_{hetV}, IR, d	0.420	0.856	0.704	25.503	2.038
68	**J_{hetV}, J_{hetM}, J_{hetE}**	**0.380**	**0.884**	**0.758**	**33.398**	**2.304**
69	**J_{hetV}, J_{hetM}, J_{hetE}, IR**	**0.360**	**0.905**	**0.783**	**29.038**	**2.513**

Finally, as expected, adding back the index of refraction (IR) in a tetraparametric model (model 69 in Table 8, boldface characters), the correlation coefficient R could be raised above 0.9, as shown in the final eq. (3).

$$logK_i(hCAVII) = 1.753 + 0.550(\pm0.204)\, J_{hetV} - 0.335\,(\pm0.105)\, J_{hetM} + 0.855(\pm0.286)\, J_{hetE} - 1.947(\pm0.942)\ IR \quad (3)$$

$$n = 32,\ SE = 0.360,\ R = 0.905,\ R^2A = 0.783,\ F = 29.038,\ Q = 2.513$$

We observe that R^2A increases with the increase in number of variables, so that in each case the added variable has a high enough contribution to the developed model.

Again, as in eq. (2), the coefficient of J_{hetV} is positive, like the coefficient of J_{hetE}. The coefficient of IR on the other hand is negative, like the coefficient of J_{hetM}.

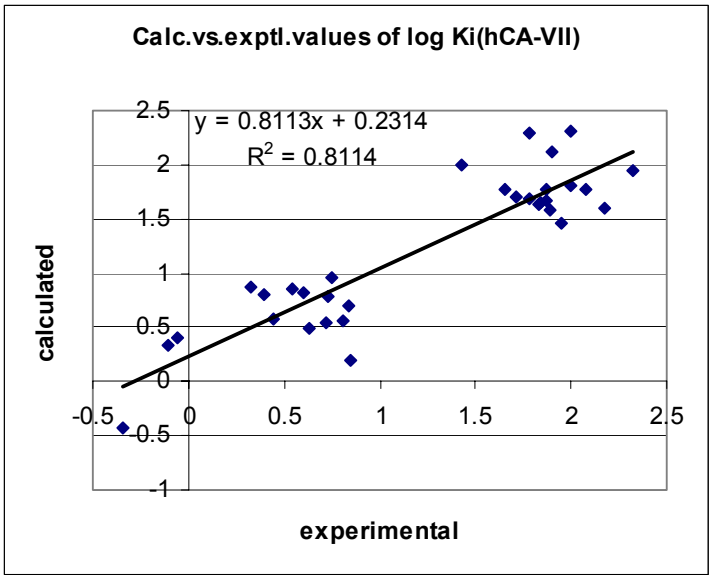


Fig. 1 – A plot between experimental and calculated $logK_i$ (hCAVII) using eq.(3) and model 69.

In Table 9 we present the experimental and calculated logKi(hcA-VII) values together with the calculated ones obtained via eq. (3), *i.e.*, model 69, as well as the residuals. Fig.1 displays a plot of these values and the linear correlation obtained therefrom.

Table 9

Experimental and calculated values of $logK_i$ (hCAVII) and residuals using eq (3), *i. e.* model 69

Comp.	Exp.	Calc.	Residual
1.	1.653	1.772	-0.119
2.	1.845	1.647	0.198
3.	1.949	1.461	0.488
4.	1.999	1.803	0.196
5.	1.875	1.670	0.205
6.	1.903	2.121	-0.218
7.	1.875	1.765	0.110
8.	2.079	1.764	0.315
9.	1.785	1.682	0.103
10.	2.176	1.593	0.583
11.	2.000	2.307	-0.307
12.	2.322	1.953	0.369
13.	0.716	0.537	0.179
14.	0.633	0.484	0.149
15.	0.845	0.200	0.645
16.	0.748	0.966	-0.218
17.	0.813	0.564	0.249
18.	0.833	0.691	0.142
19.	0.602	0.817	-0.215
20.	0.732	0.783	-0.051
21.	1.778	2.299	-0.521
22.	1.892	1.584	0.308
23.	1.716	1.704	0.012
24.	1.833	1.634	0.199
25.	0.398	0.794	-0.396
26.	0.322	0.869	-0.547
27.	-0.108	0.338	-0.446
28.	1.423	2.001	-0.578
29.	0.544	0.862	-0.318
30.	0.447	0.580	-0.133
31.	-0.060	0.403	-0.463
32.	-0.341	-0.421	0.080

COLINEARITY INVESTIGATION

The correlation matrix for the entire set of descriptors allows investigating if there are any significant correlations among the descriptors. The correlation matrix is also used to discover the correlation of the descriptors used with the activity. We have used Randić's recommendations[30, 31] in support of the proposed model. Randić stated that if a descriptor strongly correlates with another descriptor already used in a regression, such a descriptor in most studies should be discarded. However, although two highly correlated descriptors overall depict the same features of molecular structure, it is important to recognize that even highly interrelated descriptors differ in some other structural traits. The difference between them may be relatively small but nevertheless very important for structure-property regression.

The criteria for inclusion or exclusion of descriptors should not be based on colinearity between descriptors, but should be based on whether the part in which two descriptors disagree is or is not relevant for the characterization of the property under consideration. If the part in which the second descriptor differs from the first, regardless of how small it is, is relevant for the property under consideration, then the descriptor should be included. Randić further stated that the selection of descriptors to be used in structure-property-activity studies should not be delegated solely to computers, although statistical criteria will continue to be useful for preliminary screening of descriptors taken from a large pool. Often in an automated selection of descriptors, a descriptor will be discarded because it is highly correlated with another descriptor already selected. But what is important is not whether two descriptors parallel one another, i.e. duplicate much of the same structural information, but whether they are complementary in those parts that are important for structure-property-activity correlations. In our case, most of the variance of the biological activity is explained by J_{hetV}, and a non-negligible minor amount of the variance is explained by the other three descriptors, one of which (J_{hetE}) has a sizable but not unduly large (0.835) with J_{hetV}.

CONCLUSION

Among the molecular descriptors, Balaban-type indices play an important role in modeling $\log K_i$(hCAVII) activity. The best model is the tetraparametric model containing three Balaban-type indices and the index of refraction (eq. 3).

EXPERIMENTAL

$\log K_i$(hCA-VII). The inhibition data for the cytosolic isozyme VII as K_1 (nm) hCA-VII were taken from the earlier report[1] and used after converting them into logarithmic units, *i.e*. we used as $\log K_i$(hCAVII).

Topological indices. All the topological indices were calculated from the hydrogen depleted graphs of aromatic and heterocyclic sulfonamides used. In obtaining hydrogen depleted graphs all the carbon-hydrogen as well as hetero-atoms-hydrogen bonds are deleted from the molecular of the sulfonamides used. The used topological indices[32-34] are calculated using Todeschini's DRAGON Software.[35]

Structural descriptors. All the structural descriptors were calculated using ACD Labs Software.[36]

Regression analysis. All the regression analysis was performed using maximum R^2 method. This is made using the software provided by Dr. Istvan Lukovits.

ACKNOWLEDGEMENT. The authors are thankful to Dr. Istvan Lukovits for providing the Regress-1 program for performing regression analysis.

REFERENCES

1. D. Vullo, J. Voipio, A. Innocenti, C. Rivera, H. Ranki, A. Scozzafava, K. Kaila and C.T. Supuran, *Bioorg. Med. Chem. Lett.,* **2005**, *15,* 971-976.
2. I. Nishimori, "Carbonic anhydrase – its inhibitors and activators*",* C.T. Supuran, A. Scozzafava, J. Conway(eds.), CRC Press, Boca Raton, 2004, p. 25-43.
3. C.T. Supuran, "Carbonic anhydrase – its inhibitors and activators*",* C.T. Supuran, A. Scozzafava, J. Conway(eds.), CRC Press, Boca Raton, 2004, p. 1-4.
4. C.T. Supuran, A. Scozzafava, J. Conway (eds.), "Carbonic anhydrase – its inhibitors and activators*",* CRC Press, Boca Raton, 2004, p. 1-337 and references cited therein.
5. S. Pastorekova, S. Parkkila, J. Pastorek and C.T. Supuran, *J. Enzym. Inhib. Med. Chem.,* **2004,** *19,* 199-299.
6. C.T. Supuran, D. Vullo, G. Manole, A. Casini and A. Scozzafava, *Curr. Med. Chem.-Cardiovasc. Hematol Agents,* **2004,** *2,* 49-68.
7. C.T. Supuran and A. Scozzafava, *Expert Opin. Ther. Pat.,* **2000,** *10,* 575-600.
8. C.T. Supuran and A. Scozzafava, *Expert Opin. Ther. Pat,* **2002,** *12,* 217-242.
9. C.T. Supuran and A. Scozzafava, *Curr. Med. Chem.-Imm. Endocrinol. Metab. Agents,* **2001,** *1*, 61-97.
10. C.T. Supuran, A. Scozzafava and A. Casini, *Med. Res. Rev. ,* **2003,** *23,* 146-189.
11. A. Scozzafava, A. Mastrolorenzo and C.T. Supuran, *Expert. Opin. Ther. Pat.*, **2004,** *14,* 667-702.
12. C.T. Supuran, *Expert. Opin. Ther. Pat,* **2003,** *13,* 1545-1552.
13. J. Lehtonen, B. Shen, M. Vihinen, A. Casini, A. Scozzafava, C.T. Supuran, A.K. Parkkila, J. Saarnio, A. Kivela, A. Waheed, W.S. Sly and S. Parkkila, *J. Biol.-Chem.,* **2004,** *279,* 2719-2727.
14. J. Lehtonen, S. Parkkia, D. Vullo, A. Casini, A. Scozzafava and C. T. Supuran, *Bioorg. Med. Chem. Lett.,* **2004,** *14,* 3757-3762.
15. M.M. Lakkis, K.S. Osea and R.E. Tashian, *J. Histochem. Cytochem.,* **1997,** *45,* 657-662.
16. E. Ruusuvuor, L.H. Hutta, J.M. Palva, S. Smirnov, C. Rivera, K. Kaila and J. Voipio, *J. Neurosci.*, **2004,** *24*, 2699-2707.
17. A.T. Balaban, *Chem. Phys. Lett.,* **1982,** *89*, 399-404.
18. A.T. Balaban and O. Ivanciuc, "MATH/CHEM/COMP 1988. Studies in physical and theoretical chemistry", A. Graovac(ed.), Elsevier, Amsterdam, 1989, No. 63, p. 193-211.
19. A.T. Balaban, *Math. Chem. (MATCH)*, **1986**, *21*, 115-122.
20. O. Ivanciuc, T. Ivanciuc and A.T. Balaban, *J. Chem. Inf. Comput. Sci.,* **1998**, *38*, 395-401.
21. M. Barysz, G. Jashari, R. S. Lall, V.K. Srivastava, and N. Trinajstić, "Chemical applications of topology and graph theory", R.B. King (ed.), Elsevier, Amatsrdam, p. 222-227.
22. M.S. Tute, "History and objectives of quantitative drug design in advances in drug research"*,* N. J. Harter, A. B. Simmord(eds.), Vol. 6, Academic Press, London, 1971, p. 1.
23. C. Hansch, C. (Ed.) "Comprehensive drug design"*,* Pergamon Press, New York, 1990, p.19.
24. C. Hansch and A. Leo, "Exploring QSAR: fundamental and applications in chemistry and Biology", American Chemical Society, Washington, DC, 1995.
25. S. Chaterjee, A. S. Hadi and B. Price, "Regression Analysis by Example" , 3rd Ed., Wiley: New York, 2000.
26. H. Wiener, *J. Am. Chem. Soc.,* **1947**, *69*, 17.

27. M. Randić, *J. Am. Chem. Soc.,* **1975**, *97,* 6609.
28. L.B. Kier and L.H. Hall, "Molecular connectivity in structure-activity analysis", Research Studies Press – Wiley, Chichester, 1986.
29. L.B. Kier and L.H. Hall, "Molecular connectivity in chemistry and drug research", Academic Press, New York, 1976.
30. M. Randić, *J. Chem. Inf. Comput. Sci.,* **1997,** *37,* 672.
31. M. Randić, *Acta Chem. Slov.,* **1998**, *45,* 239.
32. J. Devillers, A.T., Balaban (editors), "Topological indices and related descriptors in QSAR and QSPR", Gordon and Breach, The Netherlands, 1999.
33. M. Karelson, "Molecular descriptors in QSAR/QSPR", Wiley-Interscience, New York, 2000.
34. R. Todeschini and V. Consonni, "Handbook of molecular descriptors", Wiley-VCH, Weinheim, 2000.
35. DRAGON software for calculation of topological indices: www.disat.unimib.it
36. ACD-Lab software for calculating the referred physicochemical parameters; Chem Sketch 3.0, www.acdlabs.com