Revue Roumaine de Chimie, **2006**, *51*(7-8), 703–717

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

QSAR MODELING OF CARBONIC ANHYDRASE-I, -II AND -IV INHIBITORY ACTIVITIES: RELATIVE CORRELATION POTENTIAL OF SIX TOPOLOGICAL INDICES

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

[a] Research Division, Laxmi Fumigation and Pest Control, Pvt. Ltd., 3, Khatipura, Indore 452007, India, e-mail: pvkhadikar@rediffmail.com
[b] School of Biomedical and Chemical Sciences, University of Western Australia,35, Stirling Highway Crawltywa 6009, Australia, e-mail: bwc@crystal.uwa.edu.au
[c] Texas A&M University at Galveston, 5007 Avenue U, Galveston, TX, 77551, USA, e-mail: balabana@tamug.edu
[d] Laboratorio di Chimica Bioinorganica, Departmento di Chimica, University of Florence, Via della Lastruccia, 3, Polo Scientifico, 50019 Sesto Fioventino, Firenze, Italy.
[e] Department of Chemistry, A.P.S. University, Rewa, 486 003, India, e-mail: vijay- agrawal@Iycos.com
[f] Upzon Drugs Pvt. Ld, MPLUN Plot No,32, Sector 'F' Sanwer Road, Indore 452015, India.
[g] Department of Chemistry, Holkar Model & Autonomous College, Indore, India.

Received December 9, 2005

QSAR studies on modeling the biological activities of 26 benzenesulfonamide derivatives as inhibitors of carbonic anhydrases CA-I, CA-II and CA-IV were performed using six distance-based topological indices, including Balaban-type indices J_{hetV} and J_{hetE}. Satisfactory multiparametric correlations were obtained.

INTRODUCTION

Carbonic anhydrases (CAs) are important enzymes found in red blood cells, gastric mucosa, pancreatic cells, and renal tubules. The physiological and physio-pathological processes in which carbonic anhydrases are involved were thoroughly investigated due to pharmacological applications of their inhibitors, chiefly sulfonamides.[1-3] It has been shown that such inhibitors are important clinical agents.[4] As a result a large number of aromatic and heterocyclic sulfonamides were synthesized and tested for their CA inhibitory potential.[5-20] Among these carbonic anhydrases, benzenesulfonamides have attracted much attention.[5-20] One of the authors [CTS] has published extensive work on such inhibitors,[2-4,14-25] and has reported CA inhibition data K_1 (nmol) for hCAI, hCAII and bCAIV. However, till now, few QSAR (quantitative structure activity relationship) studies employing distance- based topological indices have been reported. This is, one of the objectives of present study. It is worth mentioning that QSAR methodology is very useful in screening a large library of possible drug candidates for selectivity and potency, arriving at models that correlate molecular structure to bioactivity activity.[4-20]

A current interest in predictive QSAR is the estimation of biological activities of organic compounds acting as drugs from their calculated structural parameters. Molecular structure is encoded through numerical descriptors, which correspond to topological, geometrical, chemical, or electronic structural features. During the last decades, QSAR modeling based on topological (graph-theoretical) indices has undergone an explosive growth due to rapid progress in chemical graph theory and to advances of computer technology.

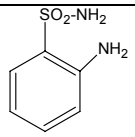
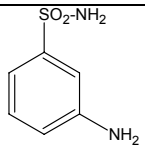
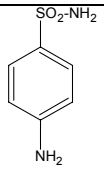
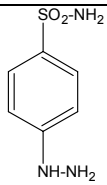
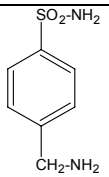
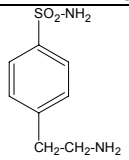
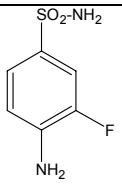
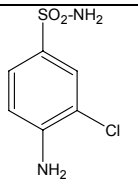
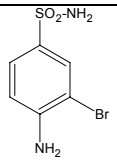
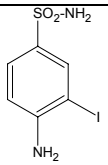
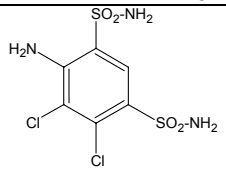
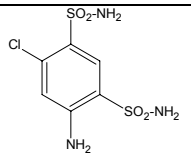
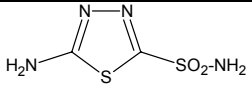
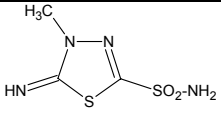
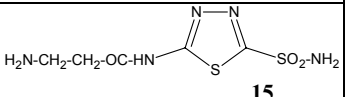
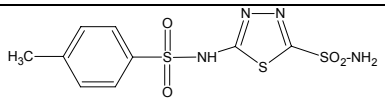
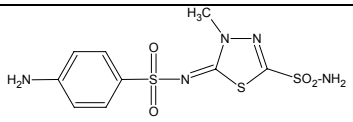
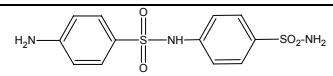
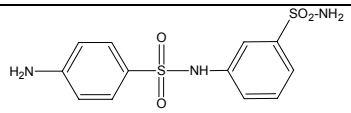
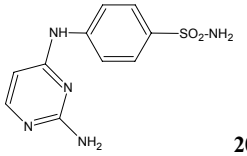
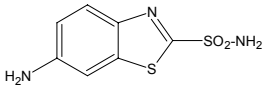
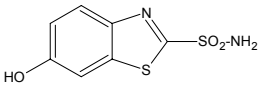
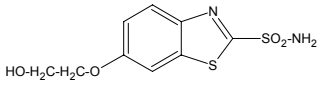
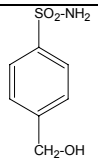
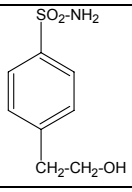
Looking to the potential of QSAR methodology, we have recently investigated the relative activity of carbonic anhydrase inhibitors.[21-36] The parameters used in these studies were chiefly distance-based topological indices. In a few cases, information-theoretic indices were also used.[21-41] Having noted that in the

*Authors for correspondence: Phone + 91-731-531906 (PVK): +1-409-741-4313 (ATB); +39-055-4573005 (CTS)

literature there are only a few QSAR studies based on the Balaban index (J),[42] and its extensions for taking into account the presence of heteroatoms and/or multiple bonds,[43-45] we decided to undertake the present study for investigating the relative potential of Balaban-type indices for modeling the CA inhibitory activities of the set of compounds indicated in Table 1.

Table 1

Structural details of carbonic anhydrase inhibitors used in 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**
 **25**		

STRUCTURES AND MOLECULAR DESCRIPTORS

Structural details of 25 carbonic anhydrase sulfonamide inhibitors used in the present study are presented in Table 1. The CA inhibitory data (K_i, nmol) reported by one of us (CTS) against isozymes I, II and IV were used by converting them into their log units.[13] Distance-based molecular descriptors selected for the present study are **topostructural indices**, namely Balaban (*J*),[42] Wiener (*W*),[46] Szeged (*Sz*),[47-49] first - order Randić connectivity ($^1\chi$),[59,51] and **topochemical indices**. Several such distance-based Balaban-type topochemical indices have been described by accounting for heteroatoms via their atomic number (J_{hetZ}),[52] atomic weight (J_{hetM}), atomic radius (J_{hetV}), electronegativity (J_{hetE}), and polarizability (J_{hetP}). Table 2 contains the values of molecular descriptors for the 25 sulfonamides from Table 1.

Table 2

Various topological descriptors used in the present study and their values*

Comp.	*W*	*J*	J_{hetZ}	J_{hetM}	J_{hetV}	J_{hetE}	J_{hetP}	$^1\chi$	*Sz*
1	144	2.545	4.788	4.789	3.092	3.614	3.531	5.016	220
2	148	2.461	4.585	4.586	3.015	3.504	3.430	4.999	228
3	152	2.394	4.425	4.426	2.953	3.415	3.348	4.999	236
4	201	2.359	4.121	4.122	2.695	3.331	2.884	5.537	306
5	201	2.359	4.008	4.009	2.895	3.257	3.216	5.537	306
6	262	2.305	3.632	3.633	2.791	3.072	3.046	6.037	388
7	189	2.512	4.526	4.540	2.878	3.592	3.083	5.410	291
9	189	2.512	4.645	4.652	3.101	3.568	3.509	5.410	291
10	189	2.512	4.720	4.730	3.150	3.553	3.562	5.410	291
11	189	2.512	4.745	4.754	3.178	3.523	3.623	5.410	291
12	458	2.991	5.883	5.894	3.629	4.203	4.260	7.459	668
13	399	2.853	5.633	5.640	3.456	4.014	4.073	7.032	582
14	113	2.449	6.035	6.039	2.416	3.271	2.913	4.499	132
15	146	2.538	5.769	5.772	2.360	3.437	2.654	4.910	171
16	403	2.304	3.826	3.827	2.064	2.805	2.189	6.931	452
17	853	1.861	3.779	3.781	1.848	2.401	2.159	9.182	1124
18	948	1.937	3.741	3.743	1.836	2.452	2.093	9.593	1237
19	1004	1.816	3.225	3.226	1.984	2.473	2.205	9.682	1502
20	960	1.900	3.409	3.410	2.049	2.575	2.287	9.682	1414
21	669	1.731	2.651	2.652	1.818	2.419	1.788	8.449	1057
22	287	1.987	3.927	3.929	2.304	2.736	2.648	6.465	430
23	287	1.987	3.953	3.955	2.261	2.749	2.583	6.465	430
24	543	1.856	3.228	3.228	1.816	2.500	1.909	8.003	776
25	201	2.359	4.036	4.036	2.827	3.276	3.120	5.537	306
26	262	2.305	3.651	3.652	2.737	3.086	2.973	6.037	388

* *W* – Wiener index; *J*- Balaban distance connectivity index; J_{hetZ} – Balaban-type index from Z-weighted distance matrix (Barysz et al. matrix); 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; $^1\chi$ – First order Randić connectivity index; and *Sz* – Szeged index.

We have carried out regression analysis [53-55] for modeling the inhibitory activity of sulfonamides against CA-I, CA-II, and CA-IV using the maximum R^2 method,[53] and the results are discussed below.

RELATIVE POWER OF BALABAN-TYPE INDICES FOR MODELING LOG K_I(HCA-I), LOG K_I(HCA-II, AND LOG K_I(BCA-IV) ACTIVITIES

The first step is obviously to investigate uniparametric modeling separately for each descriptor, although the Balaban indices were designed to account for "topological shape" (degree of branching, centricity of branches, and cyclicity), and not for molecular size. Consequently, it was emphasized that for physical-chemical properties or biological activities that depend on molecular size, *J* and *J*-type indices should always be used in multiparametric correlations.[42-45]

As expected, Table 3 shows that in uniparametric correlations the size-independent *J* and *J*-type indices lead to poorer correlations (lower values of the correlation coefficient R) than indices *W*, $^1\chi$ and *Sz*, which increase with graph size. Among the J_{het} indices, the best results (comparable to *J*) are observed for J_{hetV} and J_{hetE}.

Table 3

Regression parameters and quality of correlations for uniparametric modeling

Para-meter	CA-I			CA-II			CA-IV		
	R	St.err.	F	R	St.err.	F	R	St.err.	F
W	-0.812	0.763	44.36	-0.721	0.482	24.83	-0.653	0.641	17.14
J	0.757	0.853	30.91	0.615	0.548	13.99	0.526	0.720	9.25
J_{hetZ}	0.493	1.136	7.38	0.359	0.648	3.40	0.280	0.818	1.96
J_{hetM}	0.493	1.137	7.37	0.359	0.649	3.39	0.279	0.819	1.94
J_{hetV}	0.775	0.826	34.53	0.748	0.461	29.18	0.567	0.702	10.91
J_{hetE}	0.788	0.805	37.60	0.689	0.505	20.63	0.561	0.702	7.55
J_{hetP}	0.712	0.917	23.63	0.690	0.499	21.60	0.502	0.737	7.78
$^1\chi$	-0.830	0.729	50.81	-0.750	0.460	29.61	-0.694	0.613	21.25
Sz	-0.780	0.818	35.68	-0.678	0.511	19.58	-0.620	0.669	14.39

The next step was to proceed to biparametric correlations associating one of the *W*, $^1\chi$ and *Sz* indices to each of the *J* and *J*-type indices. In particular, we will concentrate of six topological indices, namely the foremost three Balaban indices (*J*, J_{hetV} and J_{hetE}) and three distance-based indices (*W*, $^1\chi$ and *Sz*). In Tables 4 and 5 we present the results for biparametric correlations involving these pairwise associations. It is evident that models 5, 14, and 23 involving J_{hetV} and $^1\chi$ (highlighted by boldface characters) have the highest R values and the lowest standard errors.

Table 4

Regression parameters for biparametric modeling of log K_i(hCA-I), log K_i(hCA-II), and log K_i(bCA-IV)

Model No.	Parameters used	Regression Coefficients	Constant	St.error	R^2A	R	F-ratio
(i) For modeling of log K_i(hCAI)							
1.	W J	-0.0025(±0.0006) 1.5451(±0.5497)	0.6902	0.6692	0.7259	0.8653	32.785
2.	$^1\chi$ J	-0.4518(±0.1033) 1.4770(±0.5207)	2.8764	0.6378	0.7511	0.8785	37.207
3.	Sz J	-0.0016(±0.0004) 1.6762(±0.5924)	0.3072	0.7158	0.6865	0.8442	27.277
4.	W Jhetv	-0.0024(±0.0005) 1.0114(±0.3141)	1.5676	0.6432	0.7468	0.8763	36.396
5.	**$^1\chi$ Jhetv**	**-0.4383(±0.0954) 0.9852(±0.2892)**	**3.6084**	**0.6031**	**0.7774**	**0.8922**	**42.914**
6.	Sz Jhetv	-0.0015(±0.0004) 1.1144(±0.3167)	1.2345	0.6687	0.7264	0.8656	32.858
7.	W Jhete	-0.0023(±0.0006) 1.0967(±0.3476)	0.7117	0.6473	0.7436	0.8746	35.796
8.	$^1\chi$ Jhete	-0.4237(±0.1023) 1.0474(±0.3290)	2.7775	0.6167	0.7673	0.8869	40.566
9.	Sz Jhete	-0.0015(±0.0004) 1.2032(±0.3585)	0.2981	0.6798	0.7172	0.8607	31.430

Table 4 (continues)

Table 4 (continues)

(ii) For modeling of log K_i(hCAII)							
10.	W J	-0.0013(±0.0004) 0.5229(±0.3891)	0.8625	0.4737	0.5152	0.7454	13.753
11.	$^1\chi$ J	-0.2494(±0.0735) 0.4619(±0.3708)	2.1383	0.4541	0.5545	0.7691	15.934
12.	Sz J	-0.0008(±-0.0003) 0.6159(±0.4118)	0.5898	0.4975	0.4652	0.7140	11.437
13.	W Jhetv	-0.0009(±0.0003) 0.6054(±0.2031)	0.3527	0.4159	0.6263	0.8108	21.111
14.	**$^1\chi$ Jhetv**	**-0.1909(±0.0625) 0.5711(±0.1896)**	**1.3267**	**0.3954**	**0.6623**	**0.8309**	**24.532**
15.	Sz Jhetv	-0.0005(±0.0002) 0.6638(±0.2029)	0.1545	0.4283	0.6037	0.7979	19.277
16.	W Jhete	-0.0011(±0.0004) 0.4898(±0.2431)	0.4417	0.4527	0.5571	0.7707	16.096
17.	$^1\chi$ Jhete	-0.2161(±0.0722) 0.4452(±0.2320)	1.5762	0.4349	0.5914	0.7909	18.371
18.	Sz Jhete	-0.0006(±0.0003) 0.5629(±0.2479)	0.1511	0.4700	0.5227	0.7500	14.139
(iii) For modeling of log K_i(bCAIV)							
19.	W J	-0.0015(±0.0006) 0.4832(±0.5320)	1.7555	0.6476	0.3975	0.6691	8.916
20.	$^1\chi$ J	-0.3016(±0.1004) 0.3783(±0.5060)	3.3901	0.6197	0.4483	0.7030	10.750
21.	Sz J	-0.0009(±0.0004) 0.5694(±0.5525)	1.5034	0.6676	0.3598	0.6428	7.744
22.	W Jhetv	-0.0014(±0.0005) 0.3842(±0.3116)	1.8217	0.6380	0.4153	0.6812	9.523
23.	**$^1\chi$ Jhetv**	**-0.2810(±0.0964) 0.3352(±0.2924)**	**3.2498**	**0.6096**	**0.4662**	**0.7146**	**11.479**
24.	Sz Jhetv	-0.0008(±0.0004) 0.4584(±0.3086)	1.5748	0.6516	0.3900	0.6640	8.674
25.	W Jhete	-0.0015(±0.0006) 0.3582(±0.3459)	1.7074	0.6442	0.4039	0.6735	9.132
26.	$^1\chi$ Jhete	-0.2910(±0.1024) 0.2846(±0.3292)	3.2911	0.6172	0.4529	0.7060	10.932
27.	Sz Jhete	-0.0009(±0.0004) 0.4360(±0.3483)	1.4019	0.6604	0.3735	0.6525	8.155

The next steps consist in looking at triparametric correlations involving J_{hetV} and two from the distance-based indices (W, $^1\chi$ and Sz). Results are displayed in Table 6 as models 28–33. A slight increase in R and a slight decrease in the standard error may be seen by comparing Tables 4 and 6. A final possible refinement may be obtained by introducing an indicator variable I_1 for taking into account the large residuals observed in Table 3 for the calculated versus observed inhibitory activity of CA-I for compounds 21, 22, and 23: the indicator parameter I_1 signifies the presence (=1) or absence (=0) of an electron-donating group (NH_2 or OR) attached to the benzene ring of a thiazole group. The resulting tetraparametric correlations are presented in the lower part of Table 6 as models 34 –39. Then Tables 7 and 8 show the observed and calculated data, with the corresponding residuals, for tri- and tetraparametric regressions, respectively.

PREDICTIVE POWER OF THE PROPOSED MODELS

We now discuss the predictive power of the best models for logK_i(hCA-I), logK_i(hCA-II), and logK_i(bCA-IV). The correlation coefficients R^2_{pred} are 0.8942 (model 35), 0.7709 (model 36), and 0.6663 (model 38), for logK_i(hCAI), logK_i(hCAII), and logK_i(bCAIV), respectively.

In order to investigate the intercorrelations between descriptors in the proposed models, in Table 9 we present the correlation matrix from the data in Table 2. This can be useful in determining if certain variables are redundant and therefore not needed in the model. Because J_{hetZ} and J_{hetM} are so highly intercorrelated with one another, and because J_{hetP} correlates with J_{hetV} we decided to keep only the topostructural index J, and the topochemical indices J_{hetV} and J_{hetE}, along with indices W, $^1\chi$, and Sz.

Table 5

Observed and calculated values and their residuals

A. Estimated log K_i (hCA-I) using biparametric models

Comp.	Obs.	Model	1	Model	2	Model	3	Model	4	**Model**	**5**	Model	6	Model	7	Model	8	Model	9
		Estim.	Residual	Estim.	Residual	Estim.	Residual	Estim.	Resid.	**Estim.**	**Resid.**	Estim.	Resid.	Estim.	Resid.	Estim.	Resid.	Estim.	Resid.
1.	4.657	4.267	0.39	4.369	0.288	4.229	0.428	4.351	0.306	**4.456**	**0.201**	4.341	0.316	4.344	0.313	4.438	0.219	4.327	0.330
2.	4.398	4.127	0.271	4.253	0.145	4.076	0.322	4.264	0.134	**4.388**	**0.010**	4.243	0.155	4.214	0.184	4.330	0.068	4.183	0.215
3.	4.447	4.014	0.433	4.154	0.293	3.951	0.496	4.191	0.256	**4.326**	**0.121**	4.162	0.285	4.108	0.339	4.236	0.211	4.064	0.383
4.	4.895	3.839	1.056	3.859	1.036	3.782	1.113	3.813	1.082	**3.836**	**1.059**	3.766	1.129	3.903	0.992	3.920	0.975	3.861	1.034
5.	4.398	3.839	0.559	3.859	0.539	3.782	0.616	4.016	0.382	**4.033**	**0.365**	3.989	0.409	3.822	0.576	3.843	0.555	3.772	0.626
6.	4.322	3.604	0.718	3.553	0.769	3.564	0.758	3.765	0.557	**3.712**	**0.610**	3.747	0.575	3.479	0.843	3.437	0.885	3.430	0.892
7.	3.919	4.105	-0.186	4.142	-0.223	4.062	-0.143	4.027	-0.108	**4.072**	**-0.153**	3.994	-0.075	4.217	-0.298	4.248	-0.329	4.197	-0.278
8.	3.991	4.105	-0.114	4.142	-0.151	4.062	-0.071	4.253	-0.262	**4.292**	**-0.301**	4.242	-0.251	4.190	-0.199	4.223	-0.232	4.168	-0.177
9.	3.813	4.105	-0.292	4.142	-0.329	4.062	-0.249	4.302	-0.489	**4.340**	**-0.527**	4.297	-0.484	4.174	-0.361	4.207	-0.394	4.150	-0.337
10.	3.778	4.105	-0.327	4.142	-0.364	4.062	-0.284	4.331	-0.553	**4.368**	**-0.590**	4.328	-0.550	4.141	-0.363	4.175	-0.397	4.114	-0.336
11.	3.785	4.180	-0.395	3.924	-0.139	4.275	-0.490	4.144	-0.359	**3.914**	**-0.129**	4.250	-0.465	4.269	-0.484	4.019	-0.234	4.384	-0.599
12.	3.924	4.113	-0.189	3.913	0.011	4.179	-0.255	4.110	-0.186	**3.931**	**-0.007**	4.189	-0.265	4.197	-0.273	4.002	-0.078	4.282	-0.358
13.	3.934	4.195	-0.261	4.461	-0.527	4.206	-0.272	3.741	0.193	**4.017**	**-0.083**	3.724	0.210	4.039	-0.105	4.297	-0.363	4.042	-0.108
14.	3.968	4.251	-0.283	4.406	-0.438	4.294	-0.326	3.606	0.362	**3.781**	**0.187**	3.601	0.367	4.146	-0.178	4.297	-0.329	4.185	-0.217
15.	2.658	3.255	-0.597	3.148	-0.490	3.462	-0.804	2.693	-0.035	**2.604**	**0.054**	2.838	-0.180	2.862	-0.204	2.779	-0.121	3.016	-0.358
16.	0.778	1.458	-0.680	1.476	-0.698	1.667	-0.889	1.400	-0.622	**1.404**	**-0.626**	1.562	-0.784	1.385	-0.607	1.402	-0.624	1.553	-0.775
17.	0.954	1.341	-0.387	1.403	-0.449	1.618	-0.664	1.161	-0.207	**1.212**	**-0.258**	1.375	-0.421	1.222	-0.268	1.281	-0.327	1.450	-0.496
18.	1.623	1.016	0.607	1.184	0.439	1.000	0.623	1.177	0.446	**1.319**	**0.304**	1.132	0.491	1.116	0.507	1.265	0.358	1.090	0.533
19.	1.643	1.254	0.389	1.308	0.335	1.279	0.364	1.347	0.296	**1.383**	**0.260**	1.340	0.303	1.329	0.314	1.372	0.271	1.341	0.302
20.	2.839	1.712	1.127	1.615	1.224	1.554	1.285	1.809	1.030	**1.696**	***1.143***	1.632	1.207	1.827	1.012	1.731	1.108	1.672	1.167
21.	1.845	3.051	-1.206	2.890	-1.045	2.965	-1.120	3.213	-1.368	**3.044**	***-1.199***	3.140	-1.295	3.053	-1.208	2.904	-1.059	2.965	-1.120
22.	1.740	3.051	-1.311	2.890	-1.150	2.965	-1.225	3.169	-1.429	**3.002**	***-1.262***	3.092	-1.352	3.067	-1.327	2.918	-1.178	2.981	-1.241
23.	1.699	2.217	-0.518	2.002	-0.303	2.204	-0.505	2.108	-0.409	**1.889**	**-0.190**	2.063	-0.364	2.206	-0.507	2.005	-0.306	2.178	-0.479
24.	4.380	3.839	0.541	3.859	0.521	3.782	0.598	3.947	0.433	**3.966**	**0.414**	3.914	0.466	3.843	0.537	3.863	0.517	3.795	0.585
25.	4.255	3.604	0.651	3.553	0.702	3.564	0.691	3.710	0.545	**3.659**	**0.596**	3.687	0.568	3.494	0.761	3.452	0.803	3.447	0.808

B. Estimated log K_i (hCA-II) using biparametric models

Comp.	Obs.	Model	10	Model	11	Model	12	Model	13	**Model**	**14**	Model	15	Model	16	Model	17	Model	18
		Estim.	Residual	Estim.	Residual	Estim.	Residual	Estim.	Resid.	**Estim.**	**Resid.**	Estim.	Resid.	Estim.	Resid.	Estim.	Resid.	Estim.	Resid.
1.	2.470	2.003	0.467	2.063	0.407	1.979	0.491	2.084	0.386	**2.135**	**0.335**	2.077	0.393	2.050	0.42	2.101	0.369	2.038	0.432
2.	2.380	1.954	0.426	2.028	0.352	1.921	0.459	2.034	0.346	**2.094**	**0.286**	2.021	0.359	1.992	0.388	2.056	0.324	1.971	0.409
3.	2.477	1.913	0.564	1.997	0.480	1.873	0.604	1.993	0.484	**2.059**	**0.418**	1.975	0.502	1.944	0.533	2.016	0.461	1.916	0.561
4.	2.505	1.830	0.675	1.847	0.658	1.795	0.710	1.789	0.716	**1.809**	**0.696**	1.763	0.742	1.848	0.657	1.862	0.643	1.822	0.683
5.	2.230	1.830	0.400	1.847	0.383	1.795	0.435	1.910	0.320	**1.923**	**0.307**	1.895	0.335	1.812	0.418	1.830	0.400	1.780	0.450
6.	2.204	1.721	0.483	1.697	0.507	1.695	0.509	1.787	0.417	**1.768**	**0.436**	1.778	0.426	1.653	0.551	1.639	0.565	1.621	0.583
7.	1.778	1.926	-0.148	1.949	-0.171	1.901	-0.123	1.911	-0.133	**1.937**	**-0.159**	1.893	-0.115	1.989	-0.211	2.006	-0.228	1.979	-0.201
8.	2.041	1.926	0.115	1.949	0.092	1.901	0.140	2.046	-0.005	**2.065**	**-0.024**	2.041	0.000	1.977	0.064	1.995	0.046	1.965	0.076
9.	1.602	1.926	-0.324	1.949	-0.347	1.901	-0.299	2.076	-0.474	**2.093**	**-0.491**	2.073	-0.471	1.970	-0.368	1.989	-0.387	1.957	-0.355
10.	1.845	1.926	-0.081	1.949	-0.104	1.901	-0.056	2.093	-0.248	**2.109**	**-0.264**	2.092	-0.247	1.955	-0.110	1.975	-0.130	1.940	-0.095
11.	1.447	1.821	-0.374	1.660	-0.213	1.891	-0.444	2.104	-0.657	**1.975**	**-0.528**	2.169	-0.722	1.987	-0.540	1.835	-0.388	2.071	-0.624
12.	1.875	1.827	0.048	1.702	0.173	1.876	-0.001	2.057	-0.182	**1.958**	**-0.083**	2.105	-0.230	1.961	-0.086	1.843	0.032	2.022	-0.147
13.	1.778	1.994	-0.216	2.147	-0.369	1.991	-0.213	1.705	0.073	**1.847**	**-0.069**	1.680	0.098	1.917	-0.139	2.060	-0.282	1.904	-0.126
14.	1.279	1.996	-0.717	2.086	-0.807	2.015	-0.736	1.639	-0.360	**1.737**	**-0.458**	1.620	-0.341	1.961	-0.682	2.045	-0.766	1.972	-0.693
15.	0.477	1.534	-1.057	1.474	-0.997	1.643	-1.166	1.210	-0.733	**1.182**	**-0.705**	1.257	-0.780	1.364	-0.887	1.327	-0.850	1.428	-0.951
16.	0.301	0.708	-0.407	0.708	-0.407	0.826	-0.525	0.642	-0.341	**0.629**	**-0.328**	0.717	-0.416	0.662	-0.361	0.661	-0.360	0.752	-0.451
17.	0.778	0.622	0.156	0.641	0.137	0.781	-0.003	0.542	0.236	**0.544**	**0.234**	0.642	0.136	0.580	0.198	0.595	0.183	0.705	0.073
18.	0.778	0.484	0.294	0.562	0.216	0.492	0.286	0.577	0.201	**0.611**	**0.167**	0.584	0.194	0.528	0.250	0.585	0.193	0.540	0.238
19.	0.954	0.586	0.368	0.601	0.353	0.615	0.339	0.660	0.294	**0.648**	**0.306**	0.679	0.275	0.627	0.327	0.630	0.324	0.656	0.298
20.	1.079	0.883	0.196	0.831	0.248	0.800	0.279	0.803	0.276	**0.752**	**0.327**	0.736	0.343	0.877	0.202	0.827	0.252	0.807	0.272
21.	0.954	1.522	-0.568	1.444	-0.490	1.465	-0.511	1.468	-0.514	**1.408**	**-0.454**	1.430	-0.476	1.460	-0.506	1.397	-0.443	1.404	-0.450
22.	0.903	1.522	-0.619	1.444	-0.541	1.465	-0.562	1.442	-0.539	**1.384**	**-0.481**	1.401	-0.498	1.467	-0.564	1.403	-0.500	1.411	-0.508
23.	0.845	1.115	-0.27	1.000	-0.155	1.105	-0.260	0.924	-0.079	**0.836**	**0.009**	0.901	-0.056	1.058	-0.213	0.960	-0.115	1.040	-0.195
24.	2.097	1.830	0.267	1.847	0.250	1.795	0.302	1.869	0.228	**1.884**	**0.213**	1.850	0.247	1.821	0.276	1.838	0.259	1.791	0.306
25.	2.041	1.721	0.320	1.697	0.344	1.695	0.346	1.755	0.286	**1.737**	**0.304**	1.742	0.299	1.660	0.381	1.645	0.396	1.629	0.412

C. Estimated log K_i (bCA-IV) using biparametric models

Comp.	Obs.	Model	19	Model	20	Model	21	Model	22	**Model**	**23**	Model	24	Model	25	Model	26	Model	27
		Estim.	Residual	Estim.	Residual	Estim.	Residual	Estim.	Resid.	**Estim.**	**Resid.**	Estim.	Resid.	Estim.	Resid.	Estim.	Resid.	Estim.	Resid.
1.	3.117	2.763	0.354	2.840	0.277	2.738	0.379	2.803	0.314	**2.877**	**0.240**	2.794	0.323	2.790	0.327	2.860	0.257	2.777	0.340
2.	3.342	2.716	0.626	2.814	0.528	2.683	0.659	2.768	0.574	**2.856**	**0.486**	2.752	0.590	2.744	0.598	2.834	0.508	2.722	0.620
3.	3.477	2.677	0.800	2.788	0.689	2.637	0.840	2.738	0.739	**2.835**	**0.642**	2.716	0.761	2.707	0.770	2.808	0.669	2.676	0.801
4.	3.507	2.585	0.922	2.613	0.894	2.549	0.958	2.569	0.938	**2.597**	**0.910**	2.535	0.972	2.605	0.902	2.628	0.879	2.575	0.932
5.	3.447	2.585	0.862	2.613	0.834	2.549	0.898	2.645	0.802	**2.664**	**0.783**	2.627	0.82	2.578	0.869	2.607	0.840	2.543	0.904
6.	3.389	2.465	0.924	2.442	0.947	2.438	0.951	2.518	0.871	**2.489**	**0.900**	2.505	0.884	2.422	0.967	2.409	0.980	2.387	1.002
7.	2.255	2.677	-0.422	2.709	-0.454	2.651	-0.396	2.656	-0.401	**2.694**	**-0.439**	2.632	-0.377	2.716	-0.461	2.739	-0.484	2.703	-0.448
8.	2.505	2.677	-0.172	2.709	-0.204	2.651	-0.146	2.742	-0.237	**2.769**	**-0.264**	2.735	-0.230	2.707	-0.202	2.732	-0.227	2.692	-0.187
9.	1.820	2.677	-0.857	2.709	-0.889	2.651	-0.831	2.761	-0.941	**2.785**	**-0.965**	2.757	-0.937	2.702	-0.882	2.728	-0.908	2.686	-0.866
10.	2.097	2.677	-0.580	2.709	-0.612	2.651	-0.554	2.771	-0.674	**2.795**	**-0.698**	2.770	-0.673	2.691	-0.594	2.720	-0.623	2.672	-0.575
11.	2.243	2.493	-0.250	2.272	-0.029	2.556	-0.313	2.558	-0.315	**2.370**	**-0.127**	2.638	-0.395	2.538	-0.295	2.317	-0.074	2.625	-0.382
12.	2.204	2.518	-0.314	2.349	-0.145	2.562	-0.358	2.577	-0.373	**2.432**	**-0.228**	2.636	-0.432	2.558	-0.354	2.387	-0.183	2.621	-0.417
13.	2.732	2.764	-0.032	2.960	-0.228	2.769	-0.037	2.588	0.144	**2.795**	**-0.063**	2.564	0.168	2.713	0.019	2.913	-0.181	2.708	0.024
14.	2.550	2.756	-0.206	2.870	-0.320	2.782	-0.232	2.519	0.031	**2.661**	**-0.111**	2.503	0.047	2.723	-0.173	2.841	-0.291	2.744	-0.194
15.	2.097	2.246	-0.149	2.172	-0.075	2.375	-0.278	2.036	0.061	**1.994**	**0.103**	2.115	-0.018	2.119	-0.022	2.073	0.024	2.212	-0.115
16.	0.699	1.337	-0.638	1.325	-0.626	1.469	-0.770	1.307	-0.608	**1.289**	**-0.590**	1.411	-0.712	1.312	-0.613	1.303	-0.604	1.423	-0.724
17.	0.903	1.227	-0.324	1.230	-0.327	1.402	-0.499	1.166	-0.263	**1.169**	**-0.266**	1.304	-0.401	1.190	-0.287	1.198	-0.295	1.342	-0.439
18.	1.699	1.082	0.617	1.157	0.542	1.076	0.623	1.143	0.556	**1.194**	**0.505**	1.134	0.565	1.115	0.584	1.178	0.521	1.110	0.589
19.	1.724	1.191	0.533	1.189	0.535	1.209	0.515	1.231	0.493	**1.216**	**0.508**	1.243	0.481	1.216	0.508	1.207	0.517	1.234	0.490
20.	2.188	1.559	0.629	1.497	0.691	1.460	0.728	1.560	0.628	**1.485**	**0.703**	1.458	0.730	1.589	0.599	1.521	0.667	1.492	0.696
21.	1.279	2.272	-0.993	2.192	-0.913	2.216	-0.937	2.295	-1.016	**2.205**	**-0.926**	2.244	-0.965	2.265	-0.986	2.189	-0.910	2.202	-0.923
22.	1.230	2.272	-1.042	2.192	-0.962	2.216	-0.986	2.278	-1.048	**2.191**	**-0.961**	2.225	-0.995	2.269	-1.039	2.192	-0.962	2.208	-0.978
23.	1.176	1.814	-0.638	1.679	-0.503	1.805	-0.629	1.740	-0.564	**1.609**	**-0.433**	1.710	-0.534	1.803	-0.627	1.674	-0.498	1.784	-0.608
24.	2.748	2.585	0.163	2.613	0.135	2.549	0.199	2.619	0.129	**2.641**	**0.107**	2.596	0.152	2.585	0.163	2.612	0.136	2.551	0.197
25.	2.653	2.465	0.188	2.442	0.211	2.438	0.215	2.497	0.156	**2.471**	**0.182**	2.481	0.172	2.427	0.226	2.413	0.240	2.393	0.260

Table 6

Regression parameters and quality of correlations for modeling logK_i(hCA-I), logK_i(hCA-II), and logK_i(bCA-IV) using tri- and tetra-parametric regressions

(A) Tri-parametric regressions

Model No.	Parameters used	Regression Coefficients	Constant	Se	R^2A	R	F-ratio
(i) For modeling of log K_i(hCA-I)							
28	Jhetv $^1\chi$ Sz	0.9873(±0.2766) -1.0371(±0.3540) 0.0025(±0.0014)	6.1738	0.5766	0.7965	0.9066	32.315
29	Jhetv $^1\chi$ W	1.0445(±0.2947) 0.0026(±0.0025) -0.8661(±0.4288)	5.2838	0.6025	0.7779	0.8976	29.019
(ii) For modeling of log K_i(hCA-II)							
30	Jhetv $^1\chi$ Sz	0.5730(±0.1681) -0.7409(±0.2152) 0.0023(±0.0008)	3.6827	0.3505	0.7346	0.8762	23.144
31	Jhetv $^1\chi$ W	0.6285(±0.1874) -0.6049(±0.2727) 0.0025(±0.0016)	2.9484	0.3832	0.6828	0.8500	18.221
(iii) For modeling of log K_i(bCA-IV)							
32	Jhetv $^1\chi$ Sz	0.3376(±0.2723) -0.9850(±0.3485) 0.0029(±0.0014)	6.2659	0.5677	0.5371	0.7713	10.281
33	Jhetv $^1\chi$ W	0.4176(±0.2912) -0.8756(±0.4237) 0.0036(±0.0025)	5.5788	0.5953	0.4909	0.7447	8.715

(B) Tetra-parametric regressions

Model No.	Parameters used	Regression Coefficients	Constant	Se	R^2A	R	F-ratio
(i) For modeling of log K_i(hCAI)							
34	Jhetv $^1\chi$ Sz I_1	0.6932(±0.2582) -0.6037(±0.3387) 0.0005(±0.0014) -1.0854(±0.3734)	5.2972	0.4954	0.8498	0.9353	34.949
35	Jhetv $^1\chi$ W I_1	0.3921(±0.2739) 0.3795(±0.4446) -0.0053(±0.0027) -1.7484(±0.4272)	2.0072	0.4554	0.8731	0.9456	42.281
(ii) For modeling of log K_i(hCAII)							
36	Jhetv $^1\chi$ Sz I_1	0.5353(±0.1860) -0.6853(±0.2439) 0.0020(±0.0010) -0.1391(±0.2689)	3.5704	0.3568	0.7250	0.8780	16.820
37	Jhetv $^1\chi$ W I_1	0.5359(±0.2335) -0.4282(±0.3790) 0.0014(±0.0023) -0.2481(±0.3641)	2.4835	0.3881	0.6745	0.8537	13.433
(iii) For modeling of log K_i(bCAIV)							
38	Jhetv $^1\chi$ Sz I_1	0.1146(±0.2752) -0.6562(±0.3610) 0.0014(±0.0015) -0.8234(±0.3979)	5.6009	0.5279	0.5997	0.8163	9.987
39	Jhetv $^1\chi$ W I_1	-0.0409(±0.3218) -0.0001(±0.5224) -0.0020(±0.0032) -1.2289(±0.5019)	3.2758	0.5350	0.5888	0.8107	9.590

Table 7

Observed and calculated log K_i(hCA-I), log K_i(hCA-II), and log K_i(bCA-IV) *values* using tri-parametric models 28-33.

Comp.	CA-I logK_i					CA-II logK_i					CA-IV logK_i				
	Obs.	Model 28		Model 29		Obs.	Model 30		Model 31		Obs.	Model 32		Model 33	
		Calc.	Residual	Calc.	Residual		Calc.	Residual	Calc.	Residual		Calc.	Residual	Calc.	Residual
1.	4.657	4.573	0.084	4.541	0.116	2.470	2.243	0.227	2.217	0.253	3.117	3.015	0.102	2.995	0.122
2.	4.398	4.535	-0.137	4.486	-0.088	2.380	2.229	0.151	2.189	0.191	3.342	3.029	0.313	2.992	0.350
3.	4.447	4.494	-0.047	4.431	0.016	2.477	2.212	0.265	2.160	0.317	3.477	3.032	0.445	2.981	0.496
4.	4.895	3.856	1.039	3.822	1.073	2.505	1.826	0.679	1.795	0.710	3.507	2.620	0.887	2.578	0.929
5.	4.398	4.053	0.345	4.031	0.367	2.230	1.941	0.289	1.921	0.309	3.447	2.687	0.760	2.661	0.786
6.	4.322	3.637	0.685	3.647	0.675	2.204	1.699	0.505	1.705	0.499	3.389	2.401	0.988	2.399	0.990
7.	3.919	4.131	-0.212	4.093	-0.174	1.778	1.991	-0.213	1.957	-0.179	2.255	2.763	-0.508	2.722	-0.467
8.	3.991	4.351	-0.360	4.325	-0.334	2.041	2.119	-0.078	2.097	-0.056	2.505	2.838	-0.333	2.815	-0.310
9.	3.813	4.399	-0.586	4.377	-0.564	1.602	2.147	-0.545	2.128	-0.526	1.820	2.855	-1.035	2.836	-1.016
10.	3.778	4.427	-0.649	4.406	-0.628	1.845	2.163	-0.318	2.145	-0.300	2.097	2.864	-0.767	2.847	-0.750
11.	3.785	3.688	0.097	3.797	-0.012	1.447	1.768	-0.321	1.862	-0.415	2.243	2.105	0.138	2.207	0.036
12.	3.924	3.746	0.178	3.834	0.090	1.875	1.788	0.087	1.864	0.011	2.204	2.214	-0.010	2.297	-0.093
13.	3.934	4.223	-0.289	4.203	-0.269	1.778	2.037	-0.259	2.028	-0.250	2.732	3.037	-0.305	3.054	-0.322
14.	3.968	3.838	0.130	3.873	0.095	1.279	1.789	-0.510	1.826	-0.547	2.550	2.728	-0.178	2.789	-0.239
15.	2.658	2.152	0.506	2.478	0.180	0.477	0.767	-0.290	1.060	-0.583	2.097	1.462	0.635	1.819	0.278
16.	0.778	1.282	-0.504	1.465	-0.687	0.301	0.516	-0.215	0.687	-0.386	0.699	1.145	-0.446	1.373	-0.674
17.	0.954	1.126	-0.172	1.341	-0.387	0.778	0.464	0.314	0.669	0.109	0.903	1.067	-0.164	1.349	-0.446
18.	1.623	1.841	-0.218	1.563	0.060	0.778	1.090	-0.312	0.848	-0.070	1.699	1.808	-0.109	1.534	0.165
19.	1.643	1.685	-0.042	1.518	0.125	0.954	0.926	0.028	0.779	0.175	1.724	1.571	0.153	1.403	0.321
20.	2.839	1.845	0.994	1.593	1.246	1.079	0.888	0.191	0.652	0.427	2.188	1.660	0.528	1.342	0.846
21.	1.845	2.817	-0.972	2.832	-0.987	0.954	1.199	-0.245	1.203	-0.249	1.279	1.938	-0.659	1.910	-0.631
22.	1.740	2.774	-1.034	2.787	-1.047	0.903	1.175	-0.272	1.176	-0.273	1.230	1.923	-0.693	1.892	-0.662
23.	1.699	1.604	0.095	1.652	0.047	0.845	0.753	0.092	0.606	0.239	1.176	1.274	-0.098	1.279	-0.103
24.	4.380	3.986	0.394	3.960	0.420	2.097	1.902	0.195	1.878	0.219	2.748	2.665	0.083	2.633	0.115
25.	4.255	3.583	0.672	3.591	0.664	2.041	1.668	0.373	1.671	0.370	2.653	2.382	0.271	2.376	0.277

Table 8

Observed and calculated log K_i(hCA-I), log K_i(hCA-II),and log K_i(bCA-IV) values using tetra-parametric models 34-39.

Comp.	CA-I logK_i					CA-II logK_i					CA-IV logK_i				
	Obs.	Model Calc.	34 Residual	Model Calc.	35 Residual	Obs.	Model Calc.	36 Residual	Model Calc.	37 Residual	Obs.	Model Calc.	38 Residual	Model Calc.	39 Residual
1.	4.657	4.529	0.128	4.358	0.299	2.470	2.237	0.233	2.191	0.279	3.117	2.981	0.136	2.866	0.251
2.	4.398	4.490	-0.092	4.300	0.098	2.380	2.224	0.156	2.163	0.217	3.342	2.995	0.347	2.861	0.481
3.	4.447	4.451	-0.004	4.254	0.193	2.477	2.207	0.270	2.135	0.342	3.477	2.999	0.478	2.856	0.621
4.	4.895	3.984	0.911	4.097	0.798	2.505	1.843	0.662	1.834	0.671	3.507	2.717	0.790	2.770	0.737
5.	4.398	4.123	0.275	4.175	0.223	2.230	1.950	0.280	1.941	0.289	3.447	2.740	0.707	2.762	0.685
6.	4.322	3.792	0.530	4.000	0.322	2.204	1.719	0.485	1.755	0.449	3.389	2.518	0.871	2.647	0.742
7.	3.919	4.180	-0.261	4.184	-0.265	1.778	1.997	-0.219	1.970	-0.192	2.255	2.800	-0.545	2.786	-0.531
8.	3.991	4.334	-0.343	4.271	-0.280	2.041	2.117	-0.076	2.089	-0.048	2.505	2.826	-0.321	2.777	-0.272
9.	3.813	4.368	-0.555	4.291	-0.478	1.602	2.143	-0.541	2.116	-0.514	1.820	2.831	-1.011	2.775	-0.955
10.	3.778	4.388	-0.610	4.302	-0.524	1.845	2.158	-0.313	2.131	-0.286	2.097	2.834	-0.737	2.774	-0.677
11.	3.785	3.662	0.123	3.826	-0.041	1.447	1.764	-0.317	1.866	-0.419	2.243	2.085	0.158	2.228	0.015
12.	3.924	3.755	0.169	3.910	0.014	1.875	1.789	0.086	1.874	0.001	2.204	2.221	-0.017	2.351	-0.147
13.	3.934	4.326	-0.392	4.061	-0.127	1.778	2.050	-0.272	2.007	-0.229	2.732	3.116	-0.384	2.955	-0.223
14.	3.968	4.059	-0.091	4.020	-0.052	1.279	1.818	-0.539	1.847	-0.568	2.550	2.896	-0.346	2.892	-0.342
15.	2.658	2.782	-0.124	3.304	-0.646	0.477	0.848	-0.371	1.177	-0.700	2.097	1.941	0.156	2.400	-0.303
16.	0.778	1.628	-0.850	1.682	-0.904	0.301	0.561	-0.260	0.718	-0.417	0.699	1.407	-0.708	1.525	-0.826
17.	0.954	1.431	-0.477	1.328	-0.374	0.778	0.503	0.275	0.667	0.111	0.903	1.299	-0.396	1.339	-0.436
18.	1.623	1.619	0.004	1.122	0.501	0.778	1.062	-0.284	0.785	-0.007	1.699	1.639	0.060	1.223	0.476
19.	1.643	1.618	0.025	1.381	0.262	0.954	0.917	0.037	0.759	0.195	1.724	1.520	0.204	1.307	0.417
20.	2.839	2.014	0.825	2.370	0.469	1.079	0.910	0.169	0.762	0.317	2.188	1.788	0.400	1.888	0.300
21.	1.845	2.133	-0.288	2.090	-0.245	0.954	1.111	-0.157	1.097	-0.143	1.279	1.419	-0.140	1.389	-0.110
22.	1.740	2.103	-0.363	2.073	-0.333	0.903	1.088	-0.185	1.074	-0.171	1.230	1.414	-0.184	1.390	-0.160
23.	1.699	1.048	0.651	1.121	0.578	0.845	0.502	0.343	0.530	0.315	1.176	0.852	0.324	0.906	0.270
24.	4.380	4.076	0.304	4.148	0.232	2.097	1.913	0.184	1.905	0.192	2.748	2.732	0.016	2.765	-0.017
25.	4.255	3.755	0.500	3.979	0.276	2.041	1.690	0.351	1.726	0.315	2.653	2.512	0.141	2.649	0.004

Table 9

Intercorrelation matrix for topological indices according to data in Table 2.

	W	J	J_{hetZ}	J_{hetM}	J_{hetV}	J_{hetE}	J_{hetP}	$^1\chi$	Sz
W	1								
J	-0.6527	1							
J_{hetZ}	-0.5352	0.8467	1						
J_{hetM}	-0.5348	0.8472	0.9999	1					
J_{hetV}	-0.6423	0.9037	0.6511	0.652	1				
J_{hetE}	-0.6736	0.9771	0.8143	0.815	0.9475	1			
J_{hetP}	-0.6057	0.9079	0.7198	0.7205	0.9883	0.9445	1		
$^1\chi$	0.4284	0.3537	0.3419	0.343	0.2477	0.2955	0.3055	1	
Sz	0.9848	-0.6473	-0.5642	-0.5638	-0.6252	-0.6683	-0.5943	0.4642	1

Eq. (35) and (36) are the best models used for modeling logK_i(hCA-I) This eq. (35) contains J_{hetV}, 1χ, I_1 and W as parameters whose correlation matrix is shown in Table 10. These results show that topological indices $^1\chi$ and W are highly correlated. We have proposed eq. (36) for modeling log K_i(hCA-II) and eq. (38) for modeling log K_i(bCA-IV). The correlation matrices (Table 10) show that indices $^1\chi$ and Sz appearing in equations (36) and (38) also suffer from the same collinearity defect.

Table 10

Correlation matrices for the parameters of eq. (35), (36), and (38)

Equation (35)	logK_i(hCA-I).	W	$^1\chi$	J_{hetV}	I_1
log K_i(hCA-I).	1.000				
W	-0.812	1.000			
$^1\chi$	-0.830	0.985	1.000		
J_{hetV}	0.775	-0.642	-0.625	1.000	
I_1	-0.455	-0.005	0.098	-0.331	1.000

Equation (36)	logK_i(hCA-II)	Sz	$^1\chi$	J_{hetV}	I_1
logK_i(hCA-II)	1.000				
Sz	-0.678	1.000			
$^1\chi$	-0.750	0.980	1.000		
J_{hetV}	0.748	-0.613	-0.625	1.000	
I_1	-0.368	0.004	0.098	-0.331	1.000

Equation (38)	logK_i(bCA-IV)	Sz	$^1\chi$	J_{hetV}	I_1
logK_i (bCA-IV)	1.000				
Sz	-0.620	1.000			
$^1\chi$	-0.694	0.979	1.000		
J_{hetV}	0.567	-0.613	-0.625	1.000	
I_1	-0.477	0.004	0.098	-0.331	1.000

A thorough investigation of collinearity involves examining the values of R^2 that result from regressions for each of the predictor variables against all others. The relationship between predictor variables can be judged by examining the variance inflation factor (VIF) which is defined as VIF = $1/(1- R_i^2)$, where R_i is the multiple correlation coefficient of the i'th independent variable versus all other independent variables. A VIF is defined for each variable in the equation and not for the equation (model) as a whole. Therefore, there should be as many VIFs as there are independent variables in the equation (model), and all should be less than 10. Any independent variable having VIF > 10 is indicative of the occurrence of collinearity. We observed that: $^1\chi$ and W, as well as $^1\chi$ and Sz involved in eqs. (35), (36), and (38) have VIF values higher than 10 and therefore, there is a collinearity problem in these equations (models).

The above results show that collinearity exists in all the three proposed models, and thus statistically they are disputable. However, Randić has investigated [56,57] such a problem and he recommended that under certain situations even highly correlated descriptors could be retained in the model. We will, therefore, use Randić's recommendations in the present case. 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. For

example $^{1}\chi$ and $^{2}\chi$ often strongly correlate and in many structure-property-activity studies 2χ has been discarded. This is not theoretically justified and despite the widespread practice should be stopped. 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 parallelism between descriptors even if overwhelming, but should be based on whether the part in which two descriptors disagree is or is not relevant for the characterization of the property considered. If the part in which the second descriptor differ 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,[56] 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. Hence, the residual of the correlation between two descriptors should be examined and kept or discarded depending on how well it can improve the correlation based on already selected descriptors. In view of Randić's recommendations and the fact that $^{1}\chi$, *W*, and *Sz* indices have different information contents, these highly correlated descriptors can be retained in the proposed models.

At this stage, it is worth mentioning that problems caused by colinearity; and how to deal with them, continue to be of prime concern to theoretical statistician. From a decision maker's viewpoint, one should be aware of that collinearity can (and usually does) exist and recognize the basic problems it can cause. Some of the most obvious problems and indications of severe multi-collinearity are:

Incorrect signs on the coefficient,

A change in the values of the previous coefficient when a new variable is added to the model.

Change to insignificant of a previously significant variable when a new variable is added to the model

An increase in the standard error of the estimate when a variable is added to the model.

In the present case most the correlating variables have their coefficients smaller than the respective standard error.

We now comment on adjustable-R^2 (R^2_A). These values take into account of adjustment of R^2. Therefore, if a variable is added that does not contribute its fair share, R^2_A will actually decline. It also takes into account the relationship between sample size and number of variables. The correlation coefficient R^2 may appear artificially high if the number of variables is high compared with sample size. That is, R^2 will always increase when a new independent variable is added, but R^2_A will decrease if the added variables do not reduce the unexplained variation enough to offset the loss of degrees of freedom. From Tables 4 and 6 we observe that in each case R^2_A increases with the increase in the number of variables (an exception is provided by data on CA-II from tri- to tetraparametric regressions). Thus, the added variable has a significant contribution to the developed model. All these points indicate that multi-collinearity is not that serious in the proposed models.

EXPERIMENTAL

1. Inhibitory activities. All three values of inhibitory activities logK_i(hCA-I), logK_i(hCA-II), and logK_i(bCA-IV) were taken from our earlier publication after converting into their log unit.[13]

2. Topological indices. All topological indices used in this paper were calculated from the hydrogen suppressed molecular graph of the benzenesulfonamides presented in Table 1. Their calculations are well documented in the literature.[58-63] We have used the Luko-1 program of Lukovits, Hungarian Academy of Sciences, Budapest for the calculation of Szeged index (Sz), while other indices are calculated using Todeschini's **Dragon software.**[64]

3. Regression Analysis. The maximum-R^2 method [53-55] was adopted for implementing regression analysis. The Regress-1 program of Lukovits as well as Origin-6 and NCSS programs were used.

ACKNOWLEDGEMENTS. Authors are thankful to Professor Istvan Lukovits, Hungarian Academy of Sciences, Budapest, Hungary for providing software to carry out regression analysis. Authors are also thankful to CSIR New Delhi, India for providing financial support through project No 01(1785)/02/EMR-II.

REFERENCES

1. A. T. Balaban, S. C. Basak, A. Beteringhe, D. Mills and C. T. Supuran, *Molecular Diversity*, **2004**, *8,* 401.
2. C. T. Supuran, A. Scozzafava and J. Conway (Eds.), "Carbonic Anhydrase Inhibitors and Activators", CRC Press, Boca Raton, 2004.
3. C. T. Supuran and A. T. Balaban, *Rev. Roum. Chem*., **1994**, *39*, 107.
4. C. T. Supuran, A. Casini, A. Mastrolorenzo and A. Scozzafava, *Mini-Rev. Med. Chem.,* **2004**, *4,* 625.
5. P. M. Bell and R. O. Roblin, *J. Am. Chem. Soc.,* **1942,** *64*, 2905.
6. C. Silipo and A. Vittoria, *Farmaco, Ediz. Sci.*, **1979**, *34*, 858.
7. G. H. Miller, P. M. Doukas and J. K. Seydel, *J. Med. Chem.*, **1972**, *15*, 700.
8. J. K. Seydel, *J. Med. Chem*., **1971**, *14*, 714.
9. W. Walter and R. F. Becker, *Liebigs Ann. Chem.* **1969**, *727*, 71.
10. N. Kakey, M. Aoki, A. Kamada and N. Yata, *Chem. Pharm. Bull.,* **1969**, *17*, 1010.
11. G. Dauphin and A. Kergomard, *Bull. Soc. Chim. Fr.,* **1961**, 486.
12. M. Yoshika, K. Hamamoto and T. Kubota, *Bull. Chem. Soc. Jpn.,* **1962**, *35*, 1723.
13. V. K. Agrawal, J. Singh, M. Gupta, P. V. Khadikar and C. T. Supuran, *Eur. J. Med. Chem.,* **2005**, *40*, 1002.
14. C. T. Supuran, A. Scozzafava and A. Casini, "Carbonic Anhydrase, Its Inhibitors and Activators", C.T. Supuran, A. Scozzafava, J. Conway (Eds.) CRC Press, Boca Raton, 2004, p. 67.
15. F. Mincione, L. Menabuoni and C. T. Supuran, *Ibid.,* p. 243-254.
16. C. T. Supuran, F. Briganti, L. Menabuoni, G. Mincione, F. Mincione and A. Scozzafava, *Eur. J. Med. Chem.,* **2000,** *35*, 309.
17. C. T. Supuran and B. W. Clare, *Eur. J. Med. Chem.,* **1995,** *30*, 687.
18. C. T. Supuran and B. W. Clare, *Eur. J. Med. Chem.,* **1999,** *34*, 41.
19. B. W. Clare and C. T. Supuran, *Eur. J. Med. Chem.,* **1997,** *32*, 311.
20. C. T. Supuran and B. W. Clare, *Eur. J. Med. Chem.,* **1998,** ***33***, 489.
21. C. T. Supuran and A. Scozzafava, *J. Enz. Inhib.,* **2000**, *15*, 597.
22. A. Casini, J. Antel, F. Abbate, A. Scozzafava, S. David, H. Waldeck, S. Schafer and C. T. Supuran, *Bioorg. Med. Chem. Lett*., **2003**, *13*, 841.
23. B. W. Clare and C. T. Supuran, *J. Pharm. Sci*., **1994**, *83*, 768.
24. C. T. Supuran and A. Scozzafava, *SAR QSAR Environ. Res*., **2001**, *12***,** 17.
25. C. T. Supuran and A. Scozzafava, *Eur. J. Med. Chem.,* **2000**, *35*, 867.
26. A. Thakur, M. Thakur, P. V. Khadikar and C. T. Supuran, *Bioorg. Med. Chem. Lett.,* **2005,** *15,* 203.
27. D. Mandoli, S. Joshi, P. V. Khadikar and N. Khosla, *Bioorg. Med. Chem. Lett.,* **2005,** *15***,** 405.
28. P. V. Khadikar, V. Sharma, S. Karmarkar and C. T. Supuran, *Bioorg. Med. Chem. Lett.,* **2005**, *15*, 931.
29. P. V. Khadikar, V. Sharma, S. Karmarkar and C. T. Supuran, *Bioorg. Med. Chem. Lett.,* **2005**, *15*, 923.
30. A. T. Balaban, P. V. Khadikar, C. T. Supuran, A. Thakur and M. Thakur, *Bioorg. Med. Chem. Lett.,* **2005,** *15*, 3966.
31. V. K. Agrawal, M. Banerji, M.Gupta, J. Singh, P. V. Khadikar and C. T. Supuran, *Eur. J. Med. Chem.,* **2005**, *40*, 1002.
32. M. Jaiswal, P. V. Khadikar and C. T. Supuran, *Bioorg. Med. Chem. Lett.,* **2004**, *14*, 5661.
33. M. Jaiswal, P. V. Khadikar, A. Scozzafava and C. T. Supuran, *Bioorg. Med. Chem. Lett.,* **2004,** *14*, 3283.
34. V. K. Agrawal, S. Bano, C. T. Supuran and P. V. Khadikar, *Eur. J. Med. Chem.*, **2004**, *39,* 593.
35. M. Jaiswal, P. V. Khadikar and C. T. Supuran, *Bioorg. Med. Chem.,* **2004,** *12,* 2477.
36. A. Thakur, M. Thakur, P. V. Khadikar, C. T. Supuran and P. Sudele, *Bioorg. Med. Chem.*, **2004**, *12,* 789.
37. A. Saxena, V. K. Agrawal and P. V. Khadikar, *Oxid. Commun.,* **2003**, *26*, 9.
38. V. K. Agrawal, S. Shrivastava, P. V. Khadikar and C. T. Supuran, *Bioorg. Med. Chem.*, **2003**, *11,* 5353.
39. V. K. Agrawal and P. V. Khadikar, *Bioorg. Med. Chem. Lett.*, **2003**, *13,* 447.
40. V. K. Agrawal, R. Sharma and P. V. Khadikar, *Bioorg. Med. Chem.*, **2002**, *10*, 2993.
41. A. Saxena and P. V. Khadikar, *Acta Pharm.*, **1999**, 49, 171.
42. A. T. Balaban, *Chem. Phys. Lett.,* **1982,** *89*, 399.
43. A. T. Balaban, *MATCH*, *Commun. Math. Computer Chem.,* **1986,** *21*, 115.
44. A. T. Balaban and O. Ivanciuc, "MATH/CHEM/COMP 1988 Studies in Physical and Theoretical Chemistry Series", A. Graovac (Ed.), No. 63, Elsevier, Amsterdam, 1989, p. 193-211.
45. O. Ivanciuc, T. Ivanciuc and A. T. Balaban, *J. Chem. Inf. Comput. Sci.,* **1998**, *38*, 395-401.
46. H. Wiener, *J. Am. Chem. Soc.,* **1947,** *69*, 17.
47. I. Gutman, *Graph Theory Notes New York,* **1994,** *27, 9.*
48. P. V. Khadikar, N. V. Deshpande, P. P. Kale, A. Dobrynin, I. Gutman and G. Domotor, *J. Chem. Inf. Comput. Sci.,* **1995**, *35*, *547.*
49. P. V. Khadikar, P. P. Kale, N. V. Deshpande, S. Karmarkar and V. K. Agrawal, *MATCH, Commun. Math. Comput. Chem.*, **2001**, *43*, 7.
50. M. Randić, *J. Am. Chem. Soc.,* **1975**, *97*, 6609.
51. J. Devillers and A. T. Balaban, (eds.) "Topological indices and related descriptors in QSAR and QSPR", Gordon and Breach, Williston VT, 2000.
52. 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, Amsterdam, 1983, p. 222.
53. S. Chaterjee, A. S. Hadi and B. Price, "Regression analysis by examples", 3rd ed., Wiley, New York, 2000.
54. H. Van de Waterbeemd, "Chemometric methods in molecular mesign", VCH, Weinheim, 1995.
55. J. Devillers, W. Karcher (Eds.) "Applied multiparametric analysis in SAR and environmental studies*",* Kluwer Academic, Dordrecht, 1991.
56. M. Randić, *Acta Chem. Slov.*, **1998**, *45*, 239.

57. M. Randić, *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 672.
58. M. V. Diudea and P. V. Khadikar, "Molecular topology and its applications", Galgotia Publ., New Delhi, India, (in press).
59. N. Trinajstić, "Chemical Graph Theory", 2nd ed., CRC Press, Boca Raton, Florida, 1992, chapter 10, p. 225.
60. M. V. Diudea (Ed.), "QSPR/QSAR studies by molecular descriptors", Nova Science, Huntington, New York, 2000.
61. R. Todeschini and V. Consonni, "Handbook of molecular descriptors", Wiley-VCH, Weinheim, 2000.
62. A. T. Balaban, A. Chiriac, I. Motoc and Z. Simon, "Steric fit in quantitative structure activity relations", Lecture Notes in Chemistry No. 15, Springer Verlag, Berlin, 1980.
63. A. T. Balaban, I. Motoc, D. Bonchev and O. Mekenyan, Topological indices for structure-activity correlations. In *Steric Effects in Drug Design*, (M. Charton, I. Motoc, eds.), *Topics Curr. Chem*., **1983**, *114*, 21-55, Springer, Berlin.
64. Dragon software for calculation of Balaban-type and other indices, www.disat.unimib.it