Ultra Chemistry Vol. 9(1), 13-22 (2013).

Role of Molecular Redundancy in modeling carbonic anhydrase (CA-IV) inhibition

M.L. GANGWAL[1], and SHWETA(Thakur)RATHORE*[2]

[1,2]Department of Chemistry, P.M.B. Gujarati Science College Indore (INDIA)
*E-mail:shweta_rathore02@yahoo.com

(Acceptance Date 25th January, 2013)

Abstract

In a previous article we deal with the QSAR model of water soluble sulfonamides towards carbonic anhydrase(CA-II)isoenzyme.In this paper we discuss the QSAR model of same drugs using molecular redundancy as a molecular descriptor towards carbonic anhydrase – IV(CA-IV)isozyme. The regression analysis has shown that out of the pool of topological indices used, the Molecular Redundancy Index (MRI) is the best one for modeling CA inhibitory properties against the CA-IV isozyme. The results are critically discussed on the basis of statistical parameters.

Key words: QSAR, MRI, CA-IV isozyme.

1. Introduction

Carbonic anhydrases (CAs) are important enzymes, found in red blood cells, gastric mucosa, pancreatic cells and rehal tubes. They are responsible for the introversion of carbonic acid and carbon dioxide to bicarbonate and H_3O^+, playing an important role in several physicopathological processes that include the blood transport of carbon dioxide, the formation of hydrochloric acid in the stomach and elevated pressure of the aqueous humar in eye (glaucoma)[1-5]. It is interesting to mention that many different forms of the carbonic anhydrase (CA) enzyme appear in the mammalian body, each having specific functionality. Diseases caused by problematic acid-base secretion chemistry in the body, particularly in the eye, have been linked to the dysfunctional activities of several types of carbonic anhydrases[27]. Excess secretion of aqueous humor in the eye can cause pressure gradients to occur permanent damaging eye tissue. Employing drugs, which reduce the rate of formation of aqueous humor, can treat diseases such as macular edema and open-angle glaucoma. It is believed that certain CA enzymes contribute to the creation of eye humor through production of bicarbonate ions (27A). Drugs inhibiting the activity of the CA isoenzymes that exist in the eye have been successfully in relieving symptoms of these diseases. The synthesis and testing of a wide

range of new drugs, which could inhibit CA-II secretary activity, is a continual goal in the medicinal community.

Quantitative structural-activity relationship (QSAR) methodology can be helpful in stretching of large library of possible drug, candidates for selectivity and potency. Mathematical models are formed that correlate molecular structure to an activity or property of interest. Molecular structure is encoded through the generation of the descriptions, which are numerical values corresponding to topological, geometric, or electoral structural features. Molecular Redundancy Index rank molecule according to symmetry & to include structural characteristics influencing biological activity.MRI is derived from information theory and molecular graph theory[32] and is defined as

$$MRI = \frac{\sum n_i \log n_i}{N \log N} \qquad I$$

where n_i is the number of atoms of the same kind in the i^{th} atom set, i is the number of different atoms in the molecule & N is the molecular negentropy.

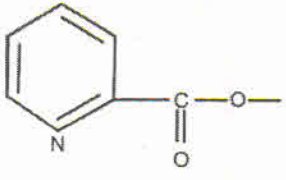
The eq.I shows that calculation of MRI leads to quantification of the information content. It encodes the salient steric properties of the molecules in cases where biological activity is non specific. It ranks them correctly according to nonspecific biological potency and thus, provides mechanistic interpretation of drugs at molecular level based on probability consideration.

2. *Methodology Used :*

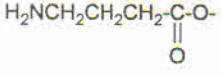
Multiple Linear Regression calculates QSAR equations by performing standard multivariable regression calculations using multiple variables in a single equation. The statistical parameters obtained are very useful to investigate the participation of each of the descriptors for modeling the activity. The modeling will be effectively carried out using softwares: REGRESS-1[36], MARTHA[39], ORIGIN[40], & NCSS. Finally the proposed QSAR models will be cross validated by leave-on-out procedure.

A large set of aromatic and hetroaromatic sulfonamides incorporating (i) picolinoyl, (ii) GABA, (iii) B-alanyl and (iv) other moiety (mentioned below) will be finally created and will be subjected to the aforementioned modeling these moieties are mentioned below:

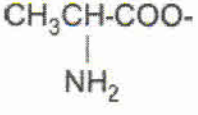
(i) ***Picolinoyl Moiety:***



(ii) ***GABA moiety***:



(iii) ***β -Ananyl moiety***:



(iv) ***Other Moiety (Tails)***:

(A) CF_3SO_2Cl (B) n-$C_4F_9SO_2Cl$

(C)

(D)

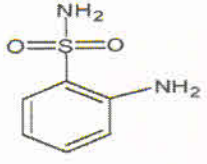
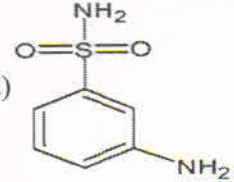
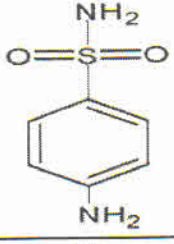
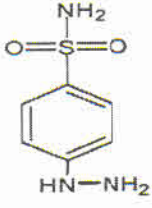
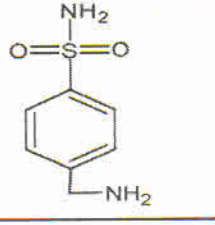
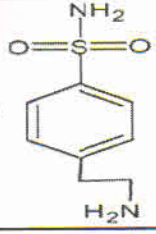
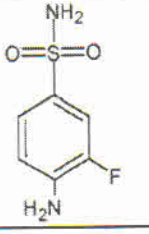
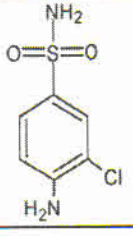
The structural details of the parent benzene sulfonamides are given in Table 1. from these compounds the aforementioned derivatives will be obtained by replacing one of the hydrogen atoms of the NH_2 moiety. In this way a large set of substituted benzene sulfonamides will be generated and the role of molecular redundancy in modeling their inhibition potential will be investigated. The main topological indices which will be used in the present study are mentioned below:

1) Wiener index[19] (W),
2) Szeged index(Sz),[20-24]
3) Randic connectivity[25-27] index (^1X),(^2X), ($^1X^v$),
4) Balaban index(J),[8.10.11]
5) Platts number(P)

3. Results

The results obtained in the present study for modeling CA-II inhibitiors we have mainly used MRI as the main correlating parameter. A series of as many as 125 CA inhibitors are used for this purpose. The modeling is carried out using correlation analysis using the method of least squares (the correlating parameters in terms of A,B,C,D,... etc, &coding C1,C2,C3..). A variety of models were obtained and their statistical significant as well as predictive power were judged. The variety of topological indices calculated for this set using DRAGON software. Finally, the models were validated using valadition technique. several cross-validated parameters were used for this purpose. The most appropriate models were then discussed in Table 4.

Table 1. Structural details of carbonic anhydrase used in present investigation

1) 	2) 	3) 	4) 
5) 	6) 	7) 	8) 

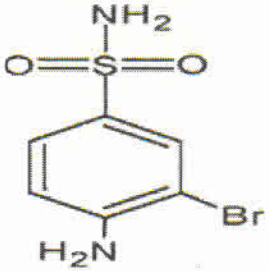
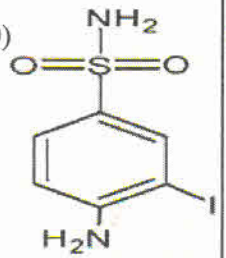
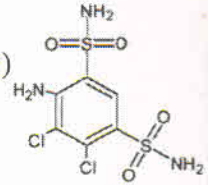
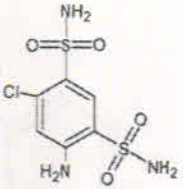
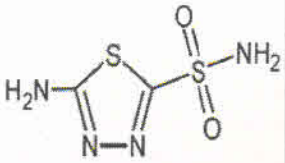
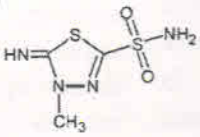
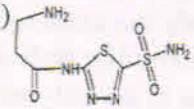
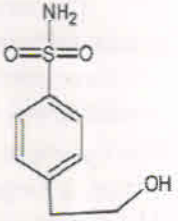
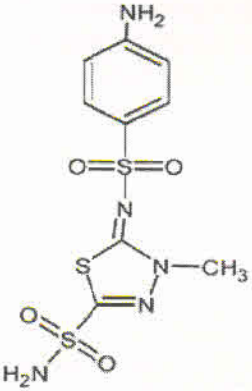
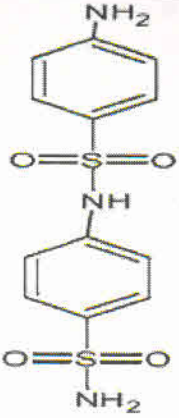
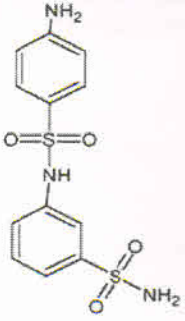
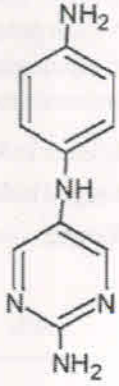
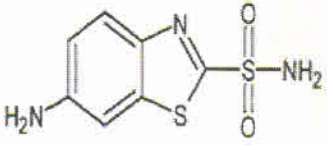
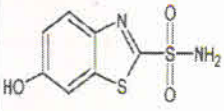
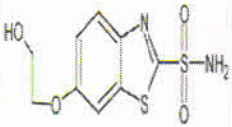
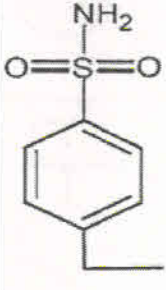
9) 	10) 	11) 	12) 
13) 	14) 	15) 	16) 
17) 	18) 	19) 	20) 
21) 	22) 	23) 	24) 

Table 2.

co. no	CAIV	W	w+p3	p	J	MRI	0χ	1χ	2χ
1	3.1173	144	158	55	2.5451	8.43714	8.483	5.0159	5.234
2	3.3424	148	164	55	2.4607	8.43714	8.483	4.999	5.335
3	3.4771	152	132	52	2.3936	8.43714	8.483	4.999	5.323
4	3.5072	201	216	65	2.3588	8.09151	9.19	5.537	5.492
5	3.4472	201	216	65	2.3588	8.41368	9.19	5.537	5.492
6	3.3892	262	278	78	2.3049	10.2775	9.897	6.037	5.872
7	2.2553	189	205	66	2.5123	7.464112	9.353	5.4097	5.831
8	2.5052	189	205	66	2.5123	7.464112	9.353	5.4097	5.831
9	1.8195	189	205	66	2.5123	7.464112	9.353	5.4097	5.831
10	2.0969	189	205	66	2.5123	7.464112	9.353	5.4097	5.831
11	2.243	458	487	136	2.9912	10.43928	13.594	7.4592	8.758
12	2.2041	399	424	120	2.8525	10.43928	12.724	7.0317	8.371
13	2.7324	113	123	45	2.4489	4.20616	7.776	4.499	4.981
14	2.5502	146	159	54	2.5376	3.12545	8.646	4.9097	5.489
15	2.0969	403	420	105	2.3042	5.67768	11.475	6.9309	6.872
16	2.6532	262	278	78	2.3049	8.18616	9.897	6.037	5.872
17	-0.0969	948	978	210	1.937	10.4608	15.837	9.5932	10.421
18	1.699	1004	1019	209	1.8165	17.2431	15.673	9.6825	10.233
19	1.7243	360	986	210	1.8996	17.2431	15.673	9.6825	10.245
20	2.1875	687	703	155	1.3774	12.9233	10.673	8.5378	6.487
21	1.2798	287	306	91	1.9874	7.35105	11.052	6.4654	7.142
22	1.2304	287	306	91	1.9874	14.38515	11.052	6.4654	7.142
23	1.1761	543	566	141	1.856	7.01838	12.466	8.0034	7.876
24	2.7482	192	207	66	2.4689	6.72149	9.19	5.3123	5.492

Where,

W	: C_3	: A	:	Wiener index.
$W + P_3$	: C_4	: B	:	Reduced Wiener index.
P	: C_5	: C	:	Platts Number
J	: C_6	: D	:	Balaban Index
MRI	: C_7	: E	:	Molecular Redundancy Index
^0X	: C_8	: F	:	Zero order Randic Connecting Index
^1X	: C_9	: G	:	First order Randic Connecting Index
^2X	: C_{10}	: H	:	Second order Randic Connecting Index
CAIV	: Inhibition	: C_2	:	Inhibition potential of CAIV.

Table 3.

Selection Results Section

Model Size	R-Squared R-Squared	Change	Coded Variables
1	0.522540	0.522540	H
2	0.597671	0.075131	EH
3	0.700545	0.102874	DEH
4	0.773201	0.072656	BDEH
5	0.783263	0.010062	BDEFH
6	0.789210	0.005947	ABDEFH
7	0.800634	0.011424	ABCDEGH
8	0.800677	0.000043	ABCDEFGH

In order to know the significant model we have plotted a graph between R^2 and variable count (x axis). These curves become parallel to the X axis when the number of descriptors 4 indicating that the maximum of 4 descriptors can be used for modeling log Ki(hCA IV)

Log Ki(hCA IV)= 2.20534879743786-.926978792238005*C10+ 3.35635836996904E-03*C4+ 1.71608029079951*C6+ .125579569956179*C7

N = 24, R^2 = 0.7732, R^2A = 0.7255, CV = 0.2005, F = 16.194

For furthure examination and statistical preference of the model we have carried out Ridge regression analysis for the parameters involved in the above model.

Table 4.

Ridge Regression Report$^{4.6.7.10}$ (mod-4)

In this report Multicollinearity is a problem so we remove parameter 4and we obtained new ridge report.

Ridge Regression Report$^{6.7.10}$ (mod-4)

Correlation Matrix Section

	C6	C7	C10	C2
C6	1.000000	-0.491229	-0.362135	0.449515
C7	-0.491229	1.000000	0.688087	-0.298502
C10	-0.362135	0.688087	1.000000	-0.722869
C2	0.449515	-0.298502	-0.722869	1.000000

Least Squares Multicollinearity Section

Independent Variable	Variance Inflation	R-Squared Vs Other X's	Tolerance
C6	1.3200	0.2424	0.7576
C7	2.1781	0.5409	0.4591
C10	1.9020	0.4742	0.5258

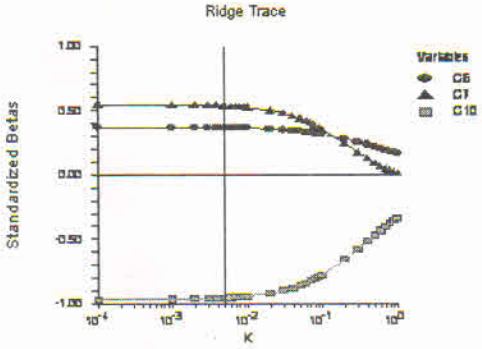
Since all VIF's are less than 10, multicollinearity is not a problem.

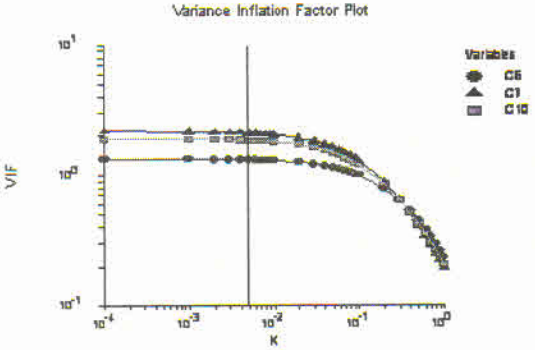
Eigenvalues of Correlations

No.	Eigenvalue	Incremental Percent	Cumulative Percent	Condition Number
1	2.039872	68.00	68.00	1.00
2	0.666909	22.23	90.23	3.06
3	0.293219	9.77	100.00	6.96

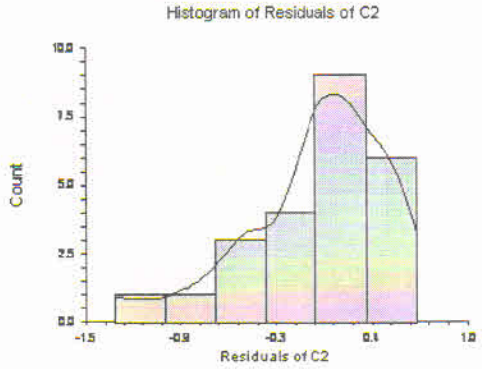
All Condition Numbers less than 100. Multicollinearity is NOT a problem.

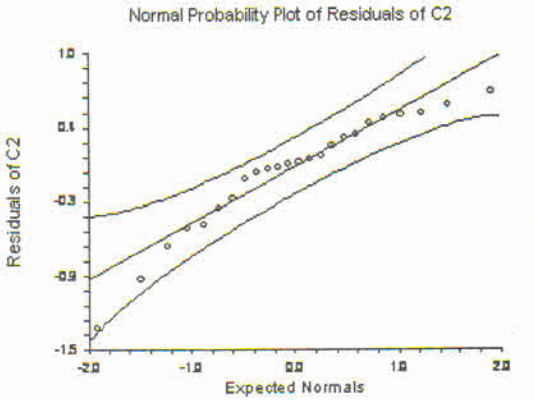
Ridge Trace Section





Residual Plots Section





4. Discussion

We now discuss the results obtained for the CA IV inhibition. Here also the correlating parameters are the same as were used for earlier case of inhibition (CA-II).

	C_2	=	CA IV inhibitory activity.
A	C_3	=	W
B	C_4	=	$W + P_3$
C	C_5	=	P
D	C_6	=	J
E	C_7	=	MRI
F	C_8	=	$^0\chi$
G	C_9	=	$^1\chi$
H	C_{10}	=	$^2\chi$

The variable selection in multiple regression analysis yielded models (3). Variation in R^2 with the number correlating parameters has shown that the maximum variables used should be at least four and that such model will be more appropriate for the exhibition of CA IV inhibition. Also, that all the four models are statistically allowed. We now discuss the results obtained in successive regression analysis[33-55].

(i) *One Variable CA IV Inhibition :*

One variable model containing C_{10} as the correlating parameters is found as below

CA IV inhibition=4.8270–0.3745 (±0.0763) C_{10}

N=24, R^2=0.5225, R^2A=0.5008, CV=0.2704, F=24.077

Note that C_{10} means $^1\chi$ Its negative coefficient indicated that decrease second order branching is favourable for the exhibition of CA IV.

(ii) *Two Variable Modeling for CA IV Inhibition :*

The two variable model containing C_{10} and C_7 as the correlating parameters. That is in additive C_{10}, this model also contain C_7 as the correlating parameter. That means in addition it contain MRI as the correlating parameter and is found as below :

CAIV inhibition=4.8678–0.5091 (±0.0988) C_{10}
+ 0.0955 (±0.0482) C_7

N=24, R^2=0.5977, R^2A=0.5594, CV=0.2541, F=15.598

(iii) *Three Variable Modeling for CA IV Inhibition :*

Further addition of C_6 (Balaban index) yielded a model with still better statistics :

CA IV inhibition=2.3264–0.5004 (±0.0878) C_{10}

+0.9105 (±0.3412) C_6+0.1383 (±0.0455) C_7

N=24, R^2=0.7005, R^2A=0.0554, CV=0.2246, F=15.596

This again shows the favourable role of MRI in the exhibition of CA IV inhibition.

(iv) *Four Variable Modeling for CA IV Inhibition :*

Finally addition of C_4 (W + P_3) gave still better model :

CA IV Inhibition=2.2053–0.9270 (±0.1807) C_{10}

$+0.0034\ (\pm 0.0014)\ C_4 + 1.7161\ (\pm 0.4506)\ C_6$

$+\ 0.1256\ (\pm 0.0411)\ C_7$

$N=24, R^2=0.7732, R^2A=0.7255, CV=0.2005, F=16.194$

The still higher parameter models saffer from the similar defect that in them one or more correlating parameter have standard deviation more than their respective coefficients and that such models are not allowed statistically[11-18].

5. Acknowledgement

The authors are thankful to **Late Dr. P.V. Khadikar** for providing necessary guidance.

References:

1. Supuran, C.T., Scozzoferva, A. *Exp. Opin. Theripatent, 12,* 217 (2002).
2. Supuran, C.T., Casini, A. *Med. Res. Rev. 23,* 146 (2003).
3. Smith, K.S., Ferry, J.G. FEMS *Microbiol Rev. 24,* 335 (2000).
4. Supuran, C.T. Scozafava, A *Cur. Med. Chem. Imm. Endoc. Metdls, Agents, 1,* 61 (2001).
5. Supuran, C.T. Scozafava, A. *Exp. Opin. Ther. Patents, 10,* 575 (2000).
6. Hies, M.A., Masereel, B., Rolin, S., Scozzafeva, A. Calnpeanu, G., Cimpeanu, V., Supuran, C.T. *Bioorg. Med. Chem., 12,* 2717-2726 (2004).
7. Supuran C.T. Scozzafava, A. Menabuori, L., Mincion, F., Brigants, C.T. *Eu. J. Pharm Sci., 8,* 317-328 (1999).
8. Mattioni, B.E., Jurs, P.C.J. *Chem. Inf. Comput. Sci., 42,* 94-102 (2002).
9. Khadikar, P.V., Deeb. O., Jaber, A., Singh, J., Agrawal, V.K. Singh, S., Lakhwani, M., Lett. Drug Design Discovery (In Press)
10. Balaban, A.T., Basak, S., Beteringhe, A., Mills, D., Supuran, C.T. *Mol. Discovery, 8,* 401-412 (2004).
11. Balaban, A. T., Khadikar, P. V., Supuran, C.T., Thakur, A., Thakur, M., *Bioorg. Med. Chem Lett, 15,* 3966-3973 (2005).
12. Singh, J., Lakhwani, M., Khadikar, P.V., Balaban A.T. Clare, B.W., Supuran C.T. Rev. Roum Chem. (In Press).
13. Khadikar, P.V., Clare, B.W., Balabar, A.T. Supuran, K.T. Agrawal, V. K. Singh, S., Joshi, A: Lakhwani, M: *Rev. Roum Chem.* (In Press)
14. Khadikar, P.V. (Unpublished results)
15. Thakur, A., Thakur, M., Khadikar, P.V., Supuran, C.T., *Bioorg. Med. Chem., 12,* 789-793 (2004).
16. Agrawal, V.K., Khadikar, P.V. *Bioorg. Med. Chem. Lett., 13,* 447-453 (2003).
17. Jaiswal, M., Khadikar, P.V., Supuran. T. *Bioorg. Med. Chem, 12,* 2477-2482 (2004).
18. Jaiswal, M., Khadikar, P.V. Scozzafava, A., Supuran, C. T. *Bioorg. Med. Chem. Lett., 14,* 3283-3290 (2004).
19. Agrawal, V.K., Bano, S., Supuran, C.T. Khadikar, P.V. *Eu. J. Med. Chem., 39,* 593-600 (2004).
20. Khadikar, P.V. Mandloi, D.J. *Bioinformoics, India 2,* 86-91 (2004).
21. Khadikar, P.V. Sharma, V., Karmarkar, S., Supuran, C.T. *Bioorg. Med. Chem. Lett, 15,* 931-936 (2005).
22. Khadikar, P.V., Sharma, V., Karmakar, S., Supuran, C.T. *Bioorg. Med. Chem. Lett, 15,* 923-930 (2005).
23. Agrawal, V.K. Singh S., Khadikar, P.V., Supuran, C.T. *Bioorgvled. Chem. Lett., 16,* 2044-2051 (2006).
24. Gray, W.D., Maren, T.H., Sisson, G.M.,

Smith, F.H.J. *Pharmacol. Exp. Jher, 121,* 160-169 (1957).
25. Duffel, M.W., Ing, I.S., Segarra, T.M., Dixon, J.A. Barfknechl, C.F., Schoehwold, R.D. *J. Med.Chem., 29,* 1488-1494 (1986).
26. Verpoorte, J.A., Mehta, Sci. Edsau, J.T. J. *Biol. Chem., 242,* 1421-4229 (1967).
27. Maren, T. H. *J. Glaucoma, 4,* 49-62; Chegwidden, W.R., Carter, N.D. Edwans, Y.H. (Eds). The Carbonic (1995). Annydoases; New Horizons, Birkhanser verlag, Basel.
28. Supuran, C.T., Scozzafava, A., Menabuoni, L., Mincionl F., Briganti, F., Minicionl, C.T. *Eu J. Pharm Sci., 8,* 317-328 (1999).
29. Gao, H. Bajorath, *J. Mol. Diversity, 4,* 115-130 (1999).
30. Clare, B.W., Supuran, C.T. *Bioors. Med. Chem., 13,* 2197-2211 (2005).
31. Bonehe, D. Rouvray D.H. Chemical Topology. Applications and Techniques, Gordon & Breack : New York (2000).
32. Bonchev, D. Rouvray, D.H. Chemical Topology Introduction and Fundamentals, Gorden & Breach; New York (1999).
33. Diudea, M.V., Ivanciace, O. Molecular Topology, Complex, Cluj, (1995).
34. Trinajstis, N. Chemical Graph Theory, CRC Press, Boca Raton (FL), (1993).
35. Balaban, A.T. Chemical Applications Graph Theom, Academic Press: London (1976).
36. DRAGON
37. LUKO-1 & REGRESS-1
38. ACD Labs
39. MARTHA
40. ORIGIN
41. Charterjee, S., Hadi. A.S., Price, B. Regression Analysis by Examples, 3rd ed, Wiley, New York (2000).
42. Diudea, M.V., Florescu. M.S. Khadikar, P.V. Molecular Topology and Its Applications, EFICON, Bucharest, 2006 (Eficol Press, Bucuresti)
43. Mandloi, D. QSAR Study on Anti-microbial Activity of Mannicl Bases, Ph.D. Thesis, D.A.V.V. Indore (2004).
44. Sharma, M. Topological Studies on the effects of Hydroxamic acids on DNA Synthesis by Ehrlion Ascites Turner Cells, Ph.D. Thesis, D.A.V.V. Indore (2005).
45. Thaklral, G.S. Ph.D. Thesis, D.A.V.V. Indore (2006).
46. Wiener, H.J. Am. *Chem. Soc., 69,* 17 / (1947).
47. Gutman, I. *Graph Theory Notes New York 27,* 9-15 (1994).
48. Khadikar, P.V., Deshpande, N.V. Kale P.P. Dobrynin, A, Gutman, I.Domotor, G.J. *Chem. Inf. Comput. Sci., 35,* 547-550 (1995).
49. Khadikar, P.V., Kale, P.P., Deshpande, N.V., Karmarkar, S., Agrawal, V.K. Commun Math. Comput. Chem. (MASCH), 43, 7-15 (2001).
50. Khadikar, P.V. Karmarkar, S. Agrawal, V.K. Singh, S. Shrivastava, A. Lukovits, I. Diudea, M.V. Lett. Drug Design Discovery, 2, 606-624 (2005).
51. Khadikar, P.V. *Nat. Acad. Sci. Letters, 23,* 113-118 (2000).
52. Khadikar, P.V., Karmankar, S. Agrawal, V.K., *J. Chem. Inf. Comput. Sci., 41,* 934-949 (2001).
53. Khadikar, P.V., Deshpande, N.V., Kale, P.P., Karmarkar, S. Agrawal, V.K., *J. Math Chem., 29,* 143-150 (2001).
54. Khadikar, P.V., Diudea, M.V., Singh, J., John, P.E., Shrivastava, A., Singh, S., Karmakar, S., Lakhwani, M., Thakur, P., *Curr. Bioact. Comp., 2,* 19-56 (2006).
55. Khadikar, P.V., Agrawal V.K., Karmarkar, S. *Bioorg. Med. Chem., 10,* 3499-3507 (2002).