Ultra Chemistry Vol. 8(3), 391-400 (2012).

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

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

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

(Acceptance Date 12th December, 2012)

Abstract

The paper deals with the use of Molecular Redundancy for developing QSAR models to predict inhibition values of aromatic & heteroaromatic sulfonamides towards carbonic anhydrase (CA-II 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-II isozyme & that excellent results are obtained using combination of MRI with distance-based topological indices.

Key words: QSAR, MRI, CA-II 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]. Some of the most sold medicinal drugs for treating heartburn are carbonic anhydrase inhibitors Ring-substituted benzene sulfonamides with SO_2NH_2 groups have been known to inhibit carbonic anhydrase-II and several QSAR studies have been published[6-23]. The active site of human carbonic anhydrass II contains a fourfold - coordinated zinc ion: three nitrogen atoms with His^{19} His^{16} and His^{119} and one oxygen atom contributed by water molecule. At physiological P^H, aromatic or heterocyclic unsubstituted sulfonamide ($R\text{-}SO_2NH_2$), which are known to inhibit CA-II, have an ionized sulfonamide group displaces the water from the zinc coordinated sphere. Substitution of the $R\text{-}SO_2\text{-}NH_2$ hydrogens substanitially decreases or completely destroys the CA-II inhibitory activity[24,25] due to steric hinderence. The aromatic side chains of sulfonamide inhibitors interact with the hydrophobic amino acid residues in the binding site *e.g.* Phe^{131}, Leu^{141}, Val^{143}, and Ala^{45} and stabilize the interaction. Unsubstituted amides *i.e.* $R\text{-}CONH_2$ such as urethane,

phenylcarbamate are a second, albeit much less potent, class of known CA-II inhibitors. In contrast to sulfonamides the compounds are basic and much weaker CA-II inhibitors amidst such as SCN^-, CLO^-_4, I^- are also weak CA-II inhibitors with Ki (binding constant) values[26] of 8-13/um. The present study is concern with the role of molecular redundancy in modeling aromatic and heterocyclic sulfonamides as inhibitors of CA-II.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[27]. 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.

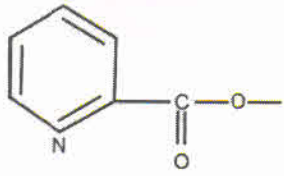
Quantitative structural-activity relationship (QSAR) methodology can be helpful in stretching of large library of possible drug, candidates for selectivity and potency. Mathematical models are formed that correlate molecular structure to an activity or property of interest. Molecular structure is encoded 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.

2. *Methodology Used :*

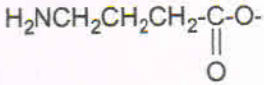
The methodology used in the present investigation is to model the CA-II inhibitory activities of CA inhibitors presented in Table 1, using molecular redundancy as the molecular descriptor. The modeling will be effectively carried out using software's: REGRESS-1 [37], MARTHA[39], ORIGIN[40] and NCSS. Finally the proposed QSAR models will be cross-validated by leave-on-out procedure[41].

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:

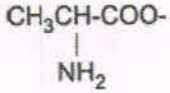
(i) ***Picolinoyl Moiety:***



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



(iii) ***β-Ananyl moiety***:



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

(A) CF_3SO_2Cl (B) $n\text{-}C_4F_9SO_2Cl$

(C)

(D)

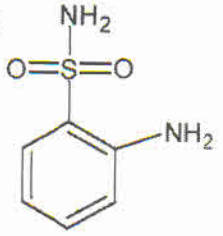
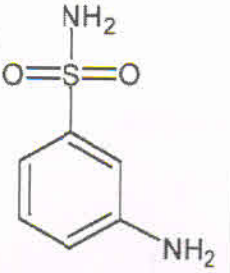
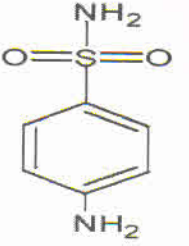
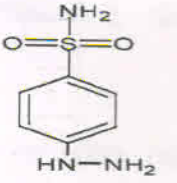
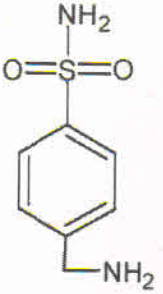
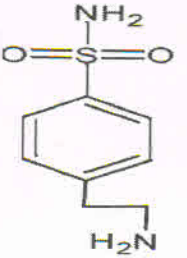
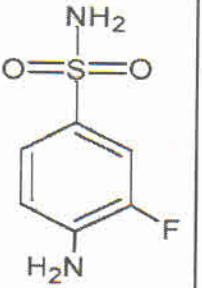
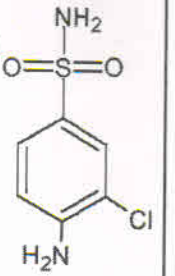
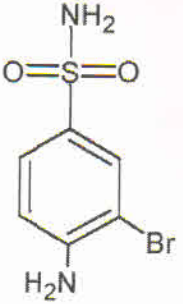
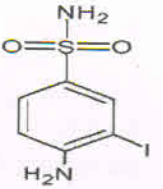
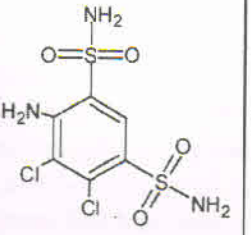
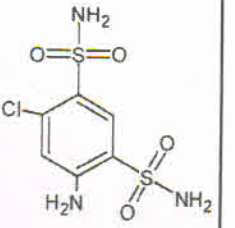
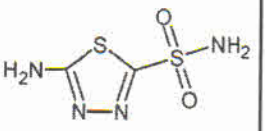
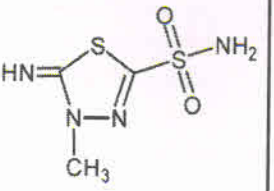
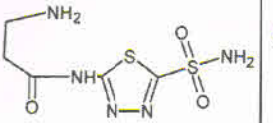
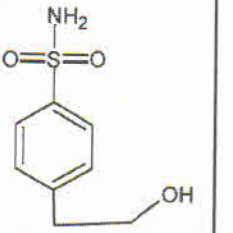
The structural details of the parent benzene sulfonamides are given in Table-1.from these compounds the aforementioned derivatives will be obtained by replacing one of the hydrogen atoms of the NH_2 moiety. In this way a large set of substituted benzene sulfonamides will be generated and the role of 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:

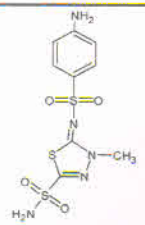
1) Wiener index (W),(19)
2) Szeged index(Sz),(20-24)
3) Randic connectivity index(^1X),(^2X) ,($^1X^v$), (25-27)
4) Balaban index(J),(8)
5) Platts number(P)

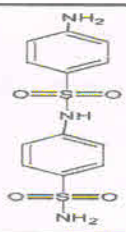
3. Results

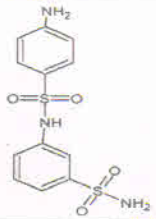
The results obtained in the present study for modeling CA-II inhibitiors we have mainly used MRI as the main correlating parameter. A series of as many as 125 CA inhibitors are used for this purpose. The modeling is carried out using correlation analysis using the method of least squares(the correlating 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.

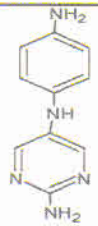
Table 1. Structural details of carbonic anhydrase used in present investigation:

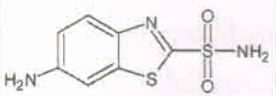
1) 	2) 	3) 	4) 
5) 	6) 	7) 	8) 
9) 	10) 	11) 	12) 
13) 	14) 	15) 	16) 

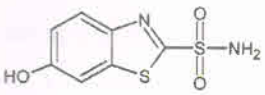
17) 

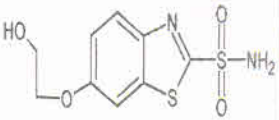
18) 

19) 

20) 

21) 

22) 

23) 

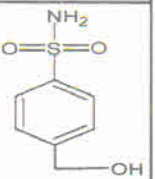
24) 

Table 2.

co no	CAII	W	w+p3	p	J	MRI	0χ	1χ	2χ
1	2.4698	144	158	55	2.5451	8.43714	8.483	5.0159	5.234
2	2.3802	148	164	55	2.4607	8.43714	8.483	4.999	5.335
3	2.4771	152	132	52	2.3936	8.43714	8.483	4.999	5.323
4	2.5052	201	216	65	2.3588	8.09151	9.19	5.537	5.492
5	2.2304	201	216	65	2.3588	8.41368	9.19	5.537	5.492
6	2.2041	262	278	78	2.3049	10.2775	9.897	6.037	5.872
7	1.7782	189	205	66	2.5123	7.464112	9.353	5.4097	5.831
8	2.0414	189	205	66	2.5123	7.464112	9.353	5.4097	5.831
9	1.6021	189	205	66	2.5123	7.464112	9.353	5.4097	5.831
10	1.8451	189	205	66	2.5123	7.464112	9.353	5.4097	5.831
11	1.4472	458	487	136	2.9912	10.43928	13.594	7.4592	8.758
12	1.8751	399	424	120	2.8525	10.43928	12.724	7.0317	8.371
13	1.7782	113	123	45	2.4489	4.20616	7.776	4.499	4.981
14	1.2787	146	159	54	2.5376	3.12545	8.646	4.9097	5.489
15	0.4771	403	420	105	2.3042	5.67768	11.475	6.9309	6.872
16	2.0414	262	278	78	2.3049	8.18616	9.897	6.037	5.872
17	-0.2219	948	978	210	1.937	10.4608	15.837	9.5932	10.421
18	0.7782	1004	1019	209	1.8165	17.2431	15.673	9.6825	10.233
19	0.9542	360	986	210	1.8996	17.2431	15.673	9.6825	10.245
20	1.0792	687	703	155	1.3774	12.9233	10.673	8.5378	6.487
21	0.9542	287	306	91	1.9874	7.35105	11.052	6.4654	7.142
22	0.9031	287	306	91	1.9874	14.38515	11.052	6.4654	7.142
23	0.8451	543	566	141	1.856	7.01838	12.466	8.0034	7.876
24	2.0969	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
CAII	:	Inhibition	:	C_2	: Inhibition potential of CAII.

Table 3. Selection Results Section

Model Size	R-Squared	R-Squared Change	Coded Variables
1	0.564887	0.564887	G
2	0.699541	0.134654	EG
3	0.815365	0.115824	DEH
4	0.823430	0.008064	BDEH
5	0.826349	0.002919	ACDEH
6	0.827475	0.001126	ACDEGH
7	0.828145	0.000670	ACDEFGH
8	0.828803	0.000658	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 3 indicating that the maximum of 3 descriptors can be used for modeling log Ki(hCA II)

Log Ki(hCA II)=.837196803934939-.4155045168518 36*C10+1.0303428751984*C6+.128330259278526*C7

N=24: Rsquare= 0.8154,Adjusted R square =0.7877,F=29.441,CV=0.2119

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(mod.3)

Correlation Matrix Section

C6	C7	C10	C2	
C6	1.000000	-0.491229	-0.362135	0.558511
C7	-0.491229	1.000000	0.688087	-0.303716
C10	-0.362135	0.688087	1.000000	-0.735897
C2	0.558511	-0.303716	-0.735897	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

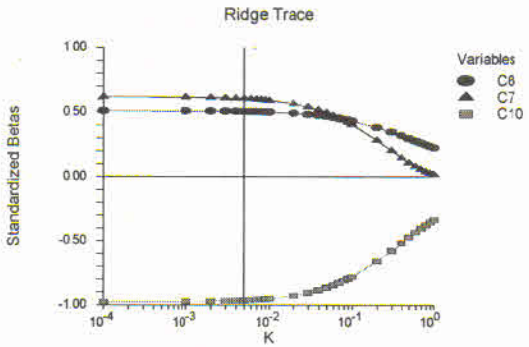
Since all VIF's are less than 10, multicollinearity is not a problem.

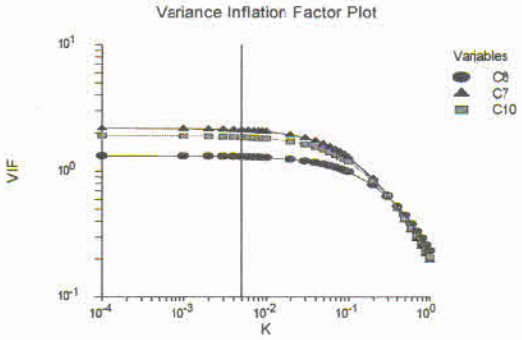
Eigenvalues of Correlations

Incremental No.	Eigenvalue	Cumulative Percent	Condition Percent	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

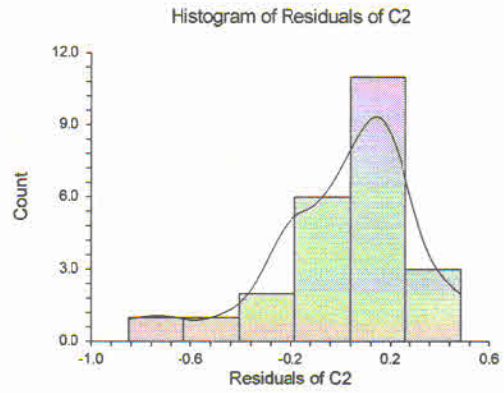
All Condition Numbers less than 100. Multicollinearity is NOT a problem.

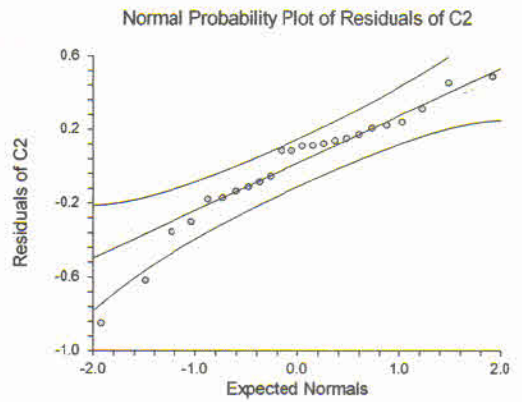
Ridge Trace Section





Residual Plots Section





4. Discussion

We now discuss the carbonic anhydrase inhibition by CAII.The variable section the multiple regression analysis yielded the following results[28-36] :

Model	R^2	Parameters
1.	0.5649	G
2.	0.6995	EG
3.	0.8154	DEH
4.	0.8234	BDEH
5.	0.8263	ACDEH
6.	0.8275	ACDEGH
7.	0.8282	ACDEFGH
8.	0.8288	ABCDEFGH

Here also, the R^2 vs number of parameters as well as the application of rule of thumb indicates that the use of at the most four correlating parameters results into most appropriate model. Therefore, we will discuss one to three variable models for investigation model for CA II inhibition.

(i) One Variable Model for CA II :

CA II inhibition = 3.7596–0.3395 (±0.0635) C_9

N = 24, R^2 = 0.5649, R^2A = 0.5451, CV = 0.3102, F = 28.562

The coefficient of the correlating parameters C_9 *i.e.* first order randic connectivity is negative. Therefore, decrease in second order connectivity increases inhibition by CA II.

(ii) Two Variable Model for CA II :

CA II inhibition=3.8986+ 0.1123 (± 0.0366) C_7

- 0.5192 (±0.0797) C_9

N = 24, R^2 = 0.6995, R^2A = 0.6709, CV = 0.2638, F = 24.447

Here again, the coefficient of C_9 is negative indicating that decrease in first order connectivity is favourable for the CAII inhibition. In addition to C_9, the model also contain C_7 parameters with positive coefficient. It means that addition MRI is favourable for the exhibition of CA II inhibition.

(iii) Three Variable Modeling :

The three variable model yields still better statistics. Here is one more parameter *i.e.* C_{10} is added, the addition of which increase R^2 from 0.6995 to 0.8154. such an improvement is due to added parameter C_{10}. Thus the model containing as C_{10}, C_6 and C_7 is form as :

CA II inhibition=0.8372–0.4155 (± 0.0563) C_{10}

+ 1.0303 (±0.2237) C_6+0.1283 (±0.0294) C_7

N = 24, R^2 = 0.8154, R^2A = 0.7877, CV = 0.2116, F = 29.441

Further higher parameter models were statistically disallowed due to the fact that one or more correlating parameter have standard direction much more that their respective coefficients.

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, S.i Edsau, J.T. *J. Biol. Chem., 242,* 1421-4229 (1967).
27. Maren, T.H. J. Glaucoma, 4, 49-62(1995); Chegwidden, W.R., Carter, N.D. Edwans, Y.H. (Eds). The Carbonic 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., *Cnrr. Bioact. Comp., 2,* 19-56 (2006).
55. Khadikar, P.V., Agrawal V.K., Karmarkar, S. *Bioorg. Med. Chem., 10,* 3499-3507 (2002).
56. Khadikar, P.V., Joshi S., Bajaj A.V. Mandloi, D. Bioorg *Med. Chem. left., 14,* 1187-119L (2004).
57. Balaban, A.T. *Chem. Phys Lett., 89,* 399-404 (1982).
58. Randic, M.J. *Am. Chem. Soc., 97,* 6609-6615 (1975).
59. Randic M.J. *Mol. Graph Model, 10,* 19-35 (2001).
60. Kier, L.B., Hall, L.H. Molecular Connectivity in Structure Activity Relationship, Wiley: New York (1986).