Development of QSAR Model of Substituted Benzene Sulphonamide using Multiple Regression Analysis

R.G. VARMA and RASHMI WAGDE

*Department of Chemistry, P.M.B. Gujarati Science College, Indore, India.

ABSTRACT

In continuation of our earlier work in this paper we studied 50 substituted Benzenesulphonamide using substituent nanofluorobutyl sulphonyl chloride ($C_4F_9SO_2Cl$) and pentafluoro benzene sulphonyl chloride ($C_6F_5SO_2Cl$) Accordingly we have development QSAR model of studied compounds. These models were derived using the parameters Balaban Index, Balaban-type index,(Jhetz, Jhetm, Jhetv, Jhete, Jhetp) Balaban related index (F,G.) and Randic connectivity index (χ1). The best suitable model is predicted on the basis of Maximum-R2 (R-squared).

**Key words:** Topological indices, Balaban and Balaban related indices.

INTRODUCTION

Carbonic Anhydrase-(II) are important enzymes found in red blood cells, gastric mucosa, pancreatic cells and renal tubes. The catalytic action of CA(II) is fundamental for respiration and transportation of CO_2 between metabolizing tissues and organs, pH regulation and homeostasis, CO_2 fixation (for algae and green plants), several metabolic biosynthetic pathways, such as glucogenesis, lipid genesis and ureagenesis, bone resorption, calcification and (in vertebrates).

CA-(II) inhibitors have found a wide range of application such as diuretics, antiepileptics, as agents for the treatment of glaucoma and modulators of cancer chemotherapy.

Carbonic Anhydrase is a ubiquitous and physiologically highly relevant isoform. It is highly efficient catalyst for the reversible hydration of carbon dioxide:

$$CO_2 + H_2O \rightarrow HCO_3^- + H^+$$
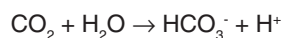
At physiological pH, aromatic and heterocyclic unsubstituted sulphonamides ($R-SO_2NH_2$),which are known to inhibit CA-II have an ionisable sulphonamide group (pK_a-6-10). Upon binding the sulphonamide group displaces the water molecule/hydroxide substantially decreases the activity due to steric hindrance[3,4]. The aromatic side chains of sulphonamide interacts with hydrophobic amino acid residues in the binding site e.g. Phe 131, Leu 141, Val 143, Ala 145 and stabilize the reaction. Unsubstituted amides such i.e. $R-CO_2NH_2$ such as urethane, phenyl carbamate are a second albeit much less potent, class of known CA-II inhibitor.

In contrast to sulphonamides the compounds are basic and much weaker. CA-II inhibitor amidst such as SCN^-, ClO_4^-, I- are also weak inhibitor with K (binding constant) values of 8-13/um(5)

Methodology

A QSAR , model is developed on the basis of topological indices. The values of topological indices were calculated by applying Graph Theory

to the chemical structure. A graph G) G(V,E) is an ordered pair consisting of two sets V) V(G) and E) E(G). Elements of the Set V(G) are called vertices and elements of the set E(G) , involving the binary relation between the vertices, are called edges. By removing all hydrogen atoms from the chemical formula of a chemical compound containing covalent bonds, one obtains the hydrogen-depleted (or hydrogen-suppressed) molecular graph of that compound. In

**Table 1 : Basic chemical structures of Carbonic Anhydrase**
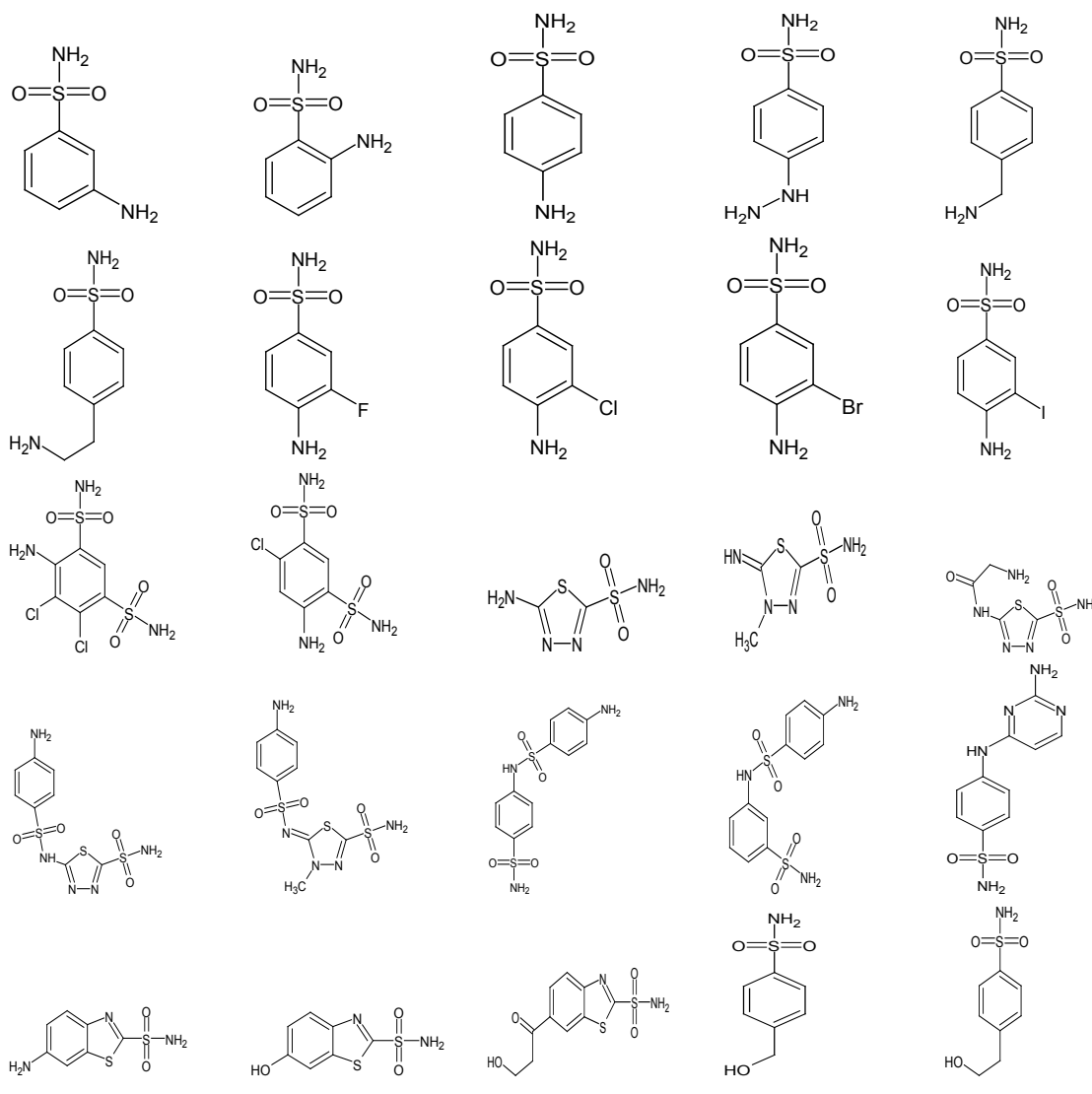
Substituent:

(A) Nanofluoro butylsulphonylchloride ($C_4F_9SO_2Cl$).
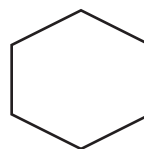
(B) Pentafluoro benzene sulphonyl chloride ($C_6F_5SO_2Cl$).

Graph theory, the atoms are considered as vertices and the bonds as "edges" Vertices are represented by " and edges by "-". In this study, the expressions "molecular graph" and "." "molecule", "vertex" and "atom", "edge" and "bond" are used interchangeably. This can be well explained by following example:
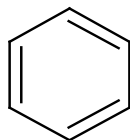
Benzene

Molecular graph of Benzene

Table 2(a): Observation

Model	Parameter	R2	Adj.R2	Std.Error	Q
1	Jhete	0.1199	0.0816	0.9963936	0.347519
2	J.Jhetm	0.6149	0.5799	0.6738863	1.1636319
3	J,Jhetm, Jhete	0.7203	0.6803	0.5878794	1.14436718
4	J,Jhetm, Jhete,F	0.7679	0.7215	0.5486671	1.5971416
5	J, Jhetm, Jhete, G,X1	0.7942	0.7401	0.530107	1.6811304
6	J, Jhetm, Jhete, F,G,X1	0.8593	0.8124	0.4503371	2.0584233
7	J, Jhetz, Jhetm, Jhete, F,G,X1	0.8636	0.8074	0.4562741	2.367161
8	J,Jhetz, Jhetm, Jhete,Jhetp F,G,X1	0.8645	0.7967	0.4687735	1.9834417
9	J, Jhetz, Jhetm,Jhetv, Jhryr, Jhetp, F,G,X1	0.8647	0.7835	0.4837536	1.922244

Table 2(b):

Model	Parameter	R2	Adj.R2	Std.Error	Q
1	Jhetv	0.1098	0.0756	0.893623	0.3515638
2	Jhrtz,.Jhete	0.5331	0.4906	0.6280298	1.162583
3	Jhetz,,Jhetv, Jhete	0.5910	0.5326	0.60159	1.277889
4	Jhetz,Jhetv, Jhete,F	0.6288	0.5546	0.5872538	1.3503005
5	Jhetz, Jhetv, Jhete, F,X1	0.6495	0.5572	0.5855453	1.3763506
6	J, Jhetm, Jhete, F,G,X1	0.6597	0.5462	0.3513644	2.3116149
7	J, Jhetz, Jhetm, Jhetv,Jhetv,Jhete, F,X1	0.6618	0.5226	0.6079906	1.3380321
8	J,Jhetz, Jhetm, Jhetv,Jhete,Jhetp F,X1	0.6636	0.4954	0.6250982	1.3031816
9	J, Jhetz, Jhetm,Jhetv, Jhete, Jhetp, F,G,X1	0.6639	0.4622	0.645291	0.0017628

The structural details of parent Benzene sulphonamide and the substituents are given in Table-1.

The methodology used in present investigation is to model the CA-II inhibitory activities of CA-II inhibitors collected from the literature (6-8) . The topological indices are calculated using softwares : Dragon (9) and ACD Labs(10) used for structure optimization

The molecular descriptors evaluated in present work are: Balabanindex (J)(11) , Balaban-type index :"Z" weighted (Jhetz), mass-weighted (Jhetm), vanderWalls weighted (Jhetv), electronegativity-weighted (Jhete) and polarisability-weighted(Jhetp)

Balaban related index (F&G)(12) and the First-order connectivity index or Randic connectivity index(X1)[13].

Balaban index

Balaban Index J (average distance-sum-connectivity index) is defined by[12]:

$$J = J(G) = \frac{M}{\mu+1}\Sigma d_i d_j^{-1/2}$$

Where M is the number of edges in a graph G,μ is the cyclomatic number of G. di's (i=1,2,3,4.. is the number of vertices in G) are the distance sums of atoms in G.

Balaban index is extended for the calculation of atom and bond parameters for heteroatom-containing molecules used in computing graph theoretic invariants. This generates a series of Balaban-type index. The different parameters are "Z" weighted (Jhetz), mass-weighted (Jhete), van der Waals weighted (Jhetm), electronegativity-weighted (Jhete) and polarisability-weighted Balaban index. In this approach, the atom and bond weights are computed on the basis of relative atom and bond parameters, using carbon as standard. The new definition of the atom and bond parameters leads to a periodic variation versus the atomic number (Z), with a more natural variation when compared with the parameters defined only by Z. This approach is used to define and compute topological indices based on graph distance. The most used molecular graph descriptors in establishing quantitative structure- property relationships (QSPRS) and quantitative structure-activity relationships (QSAR's) are topological indices[14-19].

Balaban related index

For those properties where not only the "shape" but also the size of the graph influence the property/activity, two indices related to "J" have been developed. The first of these index is F, defined as:

$$F = E\,\Sigma all\,edges\,(s_i \times s_j^{-0.5} = J(R+1)$$

This index is able to separate graph size, cyclicity and branching[20].

The other J-related index "G" is more convenient for correlations. It is defined as[20]:

$$G = \left[\frac{n^2 E}{n+R+1}\right]\Sigma_{all\,edges}\,(s_i \times s_j)^{-0.5} = \frac{n^2 J(R+1)}{n+R+1} = \frac{n^2 F}{n+R+1}$$

First –order connectivity index ($\chi 1$) : the connectivity-index graph ($\chi 1$) of a graph G is defined by Randic as under[21].

$$\chi = \chi(G) = \underset{i}{\Sigma}\,\underset{j}{\Sigma}\,\delta i\,\delta j]^{-0.5}$$

Drug-designing will be performed using correlation analysis employing NCSS[22] software.

The methodology used being the method of maximum-R^2 employing multiple regression analysis[23] R^2 (R- squared) is the correlation coefficient between experimental values and model prediction on the test set. Maximum R2 method is used to identify prediction of model. According to Tropsha's group[24-26] a QSAR model is considered predictive, if the following condition is satisfied :

$$R^2{}_{pred} > 0.6$$

$$\frac{R^2 - R_0^2}{R^2} < 0.1\;0R\;(R^2 - R_0^2) < 0.1$$

$$0.85 \leq k \leq 1.15\,or\,0.85 \leq k' \leq 1.15$$

R^2 is the correlation coefficient between experimental values and model prediction on the test set. Mathematical definitions of R20 R2 and K1 are presented clearly in literature (26) and are not described here for brevity.

RESULTS AND DISCUSSION

Based on the information contained in Table-2(A), we conclude that model 2 to 9 gives R2>0.6 and in Table-2(B) model 4 to 9 shows significant R2.On comparing observations in Table-2(A) and (B), Model 9 in Table-2(A) shows maximum-R2 i.e.0.8647. This model should be considered as best suitable model for inhibition of Carbonic Anhydrase. The Model is as follows:

Estimated Model

-9.48825311691409+.897909585675978*c10+3.07841644442707*c2+
.291542800313685*C3-285691416656861*C4+3.9

9506176564937E-02*C5+
2.25540907627915*C6+.317756734972618*C7+.4
22514045832399*C8-
5.33886464481961E-02*C9

All these observations indicates that substituent nanofluoro buty1 sulphony1 chloride($C_4F_9SO_2Cl$).

REFERENCES

1. Varma, R.G., Wagde, R. *J. of Ultrascientist of Physical Sc.,* **25**(1): (2013).
2. Supuran, C.T., Scozzafava, *A.Exp. Opin. Theripeutent.,* **12**: 217 (2002)
3. Thakur,A.,Thakur,M., Khadikar,P.V., Supuran,C.T., *Bioorg.Med.Chem.* **12**: 789-793 (2004).
4. Agrawal,V.K.,Khadikar, P.V., *Bioorg.Med Chem.Lett.,* **13**: 447-453 (2003)
5. Verpoorte,J.A.,Mehta,S.I.,Edsau, *J.T.Biol. Chem.,* **244**: 1421-1429 (1967).
6. Supuran, C.T., Scozzafava, A., Menabuoni L., Mincionl.F., Brigantti., Minicionl.C.T., *Eu. J. Pharm. Sci.* **8**: 317-328 (1999)
7. Gao,H.,Bajorath, *J. Mol.Diversity,* **4**: 115-130 (1999).
8. Clare, B.W., Supuran, C.T., *Bioors.Med. Chem.***13**: 2197-2211 (2005).
9. DRAGON
10. ACD Labs.
11. Balaban, A.T., Basak,S. Beteringhe,A., Mills,D.; Supuran.C.T., *Mol. Discovery,* **8**: 401-412 (2004)
12. Ovidiu Ivanciuc, Teodora Ivanciuc, and Alexandru T. Balaban, *J.Chem.Inf.Compit. Sci.,* **38**: 395-401 (1998).
13. Mattioni,B.E.,Jurs,P.C.*J.Chem.inf.Comput. Sci.,* **42**: 94-102 (2005).
14. (a) Balaban, A.T. Applications of Graph Theory in Chemisty. *J. Chem. Inf. Comput, Sci.* **25**: 334-343 (1985).
    (b) Balaban, A.T. Using Real Numbers as Vertex Invariants for third Generation Topological Indexes. *J. Chem. Inf. Comput. Sci.* **32**: 23-28 (1992).
    (c) Balaban, A. T. Lowering the Intra-and Intermolecular Degeneracy of Topological Invariants. *Croat. Chem. Acta* **66**: 447-458 (1993).

(d) Balaban, A.T. Local versus Global (i.e. Atomic versus Molecular) Numerical Modeling of Molecular Graphs. *J. Chem. Inf. Comput. Sci*,. **34**: 398-402 (1994) (e) Balaban, A.T. Real Number Local (Atomic) Invariants and Global (Molecular) Topological Indices. *Re V. Roum. Chjim.,* **39**: 245-257 (1994).
(f) Balaban, A. T. Local (Atomic) and Global (Molecular) Graph-Theoretical Descriptors. SAR QSAR En Viron. Res., **3**: 81-95 (1995).
15. Trinajstic, N. Chemical Graph Theory, 2nd ed.; CRC: Boca Raton, FL. (1992)
16. Hosoya, H. A Newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons. *Bull. Chem.. Soc.Jpn.* **44**: 2332-2339 (1971).
17. Randic, M. On Characterizatoin of Molecular Branching. *J. Am. Chem.. Soc.* **97**: 6609-6615 (1971).
18. (a) Kier, L.B. , Hall , L. H. Molecular Connectivity in chemistry and Drug Research; Academic: New York (1976).
    (b) Kier, L.B., Hall, L.H. Molecular Connectivity in Structure-Activity Analysis; Research Studies: Letchworth (1986).
19. Schultz, H.P. Topological Organic chemistry. 1. Graph Theory and Topological Indices of Alkanes. *J. Chem. Inf. Comput. Sci.* **29**: 227-228 (1989).
20. Balaban, A.T., Khadikar, P.V. , Supuran, C.T., thakur, A, thakur, M., *Bioorg. Med. Chem. Lett.* **15**: 3966-3973 (2005)
21. Randic, M, *J. Am. Chem.Soc.* **97**: 6609-6615 (1975).
22. NCSS.
23. Singh, J., Shaik, B., Singh., S., Agrawal, V.K., Khadikar, P.V., Deeb O., Supuran, C.T., Chemical biology and drug design (2005).
24. Shen, M., Beguin, C., Golbraikh, A., Stables.

  J. Stables, J. Kohn, H. and Tropsha, *A. J. Med. Chem.* **47**: 2356-2364 (2004).

25. Tropsha, A. Gramatica, P. Gombar, V.K.*, QSAR & Comb. Science*. **22**: 69-77 (2003).

26. Golbraikh, A. and Tropsha, *A. J. Mol. Graph Mod.,* **20**: 269-276 (2002).

27. P. Tiwari and J.P. Mishra, *Orient. J. Chem*., **26**(3): 1113-1117 (2010).