June 2008 *Volume 4 Issue 4*

Trade Science Inc.

Organic CHEMISTRY

An Indian Journal

Full Paper

OCAIJ, 4(4), 2008 [299-303]

QSAR study on aromatic nitro compounds: Reduction by $CH_3 \cdot CHOH$

Sharmila Baghel[1], Abhilash Thakur[2]*, Mamta Thakur[3], Tulsi R.Thapak[4], Arpan Bharadwaj[4]

[1]Department of Chemistry, Govt. New Girls Degree College, Indore, (INDIA)

[2]Department of Chemistry, Rishiraj Institute of Technology, Indore, (INDIA)

[3]Department of Chemistry, Ujjain Engineering College, Ujjain, (INDIA)

[4]Department of Chemistry, Madhav Science College, Ujjain, (INDIA)

Tell : 91-9926043140

E-mail : abhilashthakur@yahoo.com

Received: 23rd September, 2007 ; Accepted: 28th September, 2007

ABSTRACT

QSAR study on reduction of aromatic nitro compounds by $CH_3 \cdot CHOH$ has been carried out using heterogeneous set of molecular descriptors, which includes some distance-based topological indices and physicochemical descriptors. The results have shown that refractive index and branching plays a dominating role in this regard. The combination of these two parameters yielded statistically significant model. The results are validated by the cross validation parameters and critically discussed with a variety of statistical parameters. The three dimensional features for the reduction are studied by quantum molecular modeling.

KEYWORDS

QSAR;
Modeling;
Quantum parameter;
Physicochemical property.

INTRODUCTION

Over the last decade Hansch group has studied the effect and significance of the hydrophobic parameters as well as the steric parameters in the modeling of various biological activities of several compounds. In a review based on the role of hydrophobic parameters in QSAR studies, the Hansch group[1] has reported that the use of steric parameter σ yielded an excellent statistics on a set of 13 aromatic nitro compounds. Inspired by the work of Hansch[1], and in the continuation of our earlier works[2-9], we have revisited the work of Hansch[1] to see if we can further develop a significant QSAR model with entirely different set of parameters. In order to achieve this objective, we have used a large set of topological indices viz., Wiener index(W)[10], first order connectivity index(χ)[11], Balaban index(J)[12], Szeged index(Sz)[13], and molecular descriptors like Molar refractivity(MR), Molar volume(MV), Parachor (Pc), Refractive index (η), Surface tension(ST) and Density(D) along with some non conventional physicochemical parameters like Approximate surface area (ASA), Surface area grid(SAG) and hydration energy (HE) together with indicator parameters used for the structural and positional specifications.

In order to obtain a statistically significant model, we have used a maximum R^2 method which was followed by stepwise regression analyses[14-16]. The predictive ability of the model is discussed on the basis of the various cross validation parameters. In order to vali-

date our model further, we have used quantum molecular modeling parameters, and on the basis of these parameters, we have analyzed the structural behavior of these molecules. For the molecular modeling, we have optimized the geometries of molecules using the molecular mechanics method by applying the MM+ force field method.

The basic structures of aromatic nitro compounds used in the present study and various substituents are presented in TABLE 1 along with their observed and calculated activities and the parameters that we have used for the prediction of activity. Basak[17] has reviewed the general topic of techniques for topological and other indices for such structure activity relations and the readers are referred to this excellent review for further details on this topic.

RESULT AND DISCUSSION

The set of 12 aromatic nitro compounds and their adopted activities i.e., the reduction of aromatic nitro compounds by $CH_3C{\cdot}HOH$ have been expressed as logK and are presented in TABEL 1. As may be seen from TABLE 1, a very low-level degeneracy is present in the activity logK. As a result of the occurrence of degeneracy in activity logK, it becomes essential to examine the degeneracy in the molecular descriptors also. A perusal of TABLE 2, which contains the topological indices calculated for Aromatic nitro compounds, shows that high to low degeneracy is observed in the topological indices. Balaban[19,20] has shown that these indices, in spite of their degeneracies, can be used successfully in developing statistically significant QSAR models.

The correlation among the descriptors like topological indices, physicochemical properties, indicator parameters and activities shows (TABLE 1) that except for the Balaban index, all other topological indices are highly mutually correlated, while this is not so with other physicochemical parameters used. Furthermore, the data presented in TABLE 4 (Correlation matrix) express that none of the topological or molecular descriptors shows significant correlation with the activity (logK), except the randic branching index. The univariat model is shown below:-

logK=0.4474(±0.2339)χ+6.3063 **(1)**

TABLE 1: Reduction of 4-X-$C_6H_4NO_2$ by CH_3·CHOH in N_2O-saturated aqueous solution

No.	X	logK(Obs.)	logK(Calc.)	Residual	χ	η
1	NO_2	9.38	9.06	0.32	5.609	1.612
2	CN	9.08	8.73	0.35	5.236	1.579
3	SO_2NH_2	8.90	9.19	-0.29	5.910	1.611
4	CHO	9.26	8.93	0.33	5.236	1.617
5	CF_3	8.63	8.45	0.18	5.910	1.472
6	COMe	9.00	8.78	0.22	5.609	1.558
7	CO_2Me	8.93	8.98	-0.05	6.147	1.553
8	$CONH_2$	8.63	9.06	-0.43	5.609	1.612
9	H	8.52	8.50	0.02	4.305	1.612
10	Me	8.18	8.36	-0.18	4.698	1.553
11	OMe	8.04	8.53	-0.49	5.236	1.542
12	OH*	7.85	8.49	-0.64	4.698	1.612

***Data point not included in eq. (3)**

TABLE 2 : Cross validation parameters

Eq	Compd. used	Parameters used	R^2_A	r_{cv}^2	Press/SSY	Q
1.	12	1	-----	-----	2.73	1.19
2.	12	2	0.2953	-0.36	1.36	1.60
3.	11	2	0.3262	-0.17	1.17	1.96

Press-Predictive residual sum of square; SSY-Sum of the squares of response values, r^2_{cv}- Cross-validated predictive correlation coefficient, R^2_A-adjusted r^2; Q-Pogliani quality factor.

n=12, Se=0.4341, R=0.5176, F=3.659, Q=1.19

Equation (1) signifies the presence of branching and connectivity on benzene moiety but this model alone is unable to provide sufficient structural information required for the modeling activity logK.

To understand the structural requirement of aromatic nitro compounds bi-parametric combinations were tested, among those bi-parametric combination only the combination of Randic branching/connectivity index and the refractive index η gives the significant improvement in the result. Model obtained from above parameters is as:

logK=0.5657(± 0.232) χ+4.584(±2.942) η-1.5682 **(2)**

n= 2, Se=0.4061, R=0.6507, F=3.305, Q=1.60

Model shoes the significant role of branching/connectivity and the refractive index in the modeling reduction activity. Model also demonstrates that the branching/connectivity in the molecule play the dominating role over the refractive index for the reduction activity. Increase in the value of R and significant decrease in the Se also conform the suitability of the model for the prediction of reduction activity. There is a increase in the quality factor for the equation also justify the model.

Over all less value of regression coefficient R indicate towards the further improvement of the predictive

potential of the model, so for the improvement in the model we out lair the compound no **(12)** based on the magnitude of the residue.

Model obtained after the outlair is as:

logK=0.4317(± 0.2086)χ+5.3217 (± 2.5378) η–1.9379 (3)

n=11, Se=0.3469, R=0.679, F=3.421, Q=1.96

Comparison of eq.(2), and (3) shows that the model obtained for the set of **(11)** compounds gives the better statistics and most suitable for the prediction of reduction activity of the compounds. It is obvious that reduction in size of data set increases the regression value, but in present case significant lowering of Se and a improvement in the F-statistics along with the significant improvement in the value of Q(quality factor) from eq.(2) to (3) justify the improvement in statistics and deletion of the compound no **(12)**.

The predictive potential of the models proposed by us is determined using cross-validation method[9-12]. The various cross-validation parameters thus calculated are given in TABLE 2.

At this stage, it is worthy to comment on cross validation parameter R^2_A values. We observed that as we passes from the model obtained for**(12)** compounds (Eq.2) to model obtained for**(11)** compounds eq.(3) there is significant increase in the value of R^2_A. The values increasing from 0.2953 to 0.3262, as we passes from eq.(2) to eq.(3). Such an increase in R^2_A values indicate that the deleted compounds have the unfair share in the modeling of respective activity and also showing exceptional behavior from their parent series. The value of R^2_A will decrease if the deletion of the compounds does not reduce the unexplained variation in the model enough to off set the loss of degree of freedom[14-18] Another important cross-validation parameters is the predictive potential correlation coefficient r_{cv}^2; there is increase in the value of r_{cv}^2 from eq. 2 to 3. This shows the better predictive power of eq.(3).

The significant lowering in the value of cross-validation parameters PRESS/SSY from eq. 1 to 3 exhibits the decrease in the uncertainty of prediction. At the same time increase in the value of Q factor demonstrate the improvement in the predictive power from eq.1 to 3.

Based on the above study and magnitude of the residues for the set a of benzoate derivatives, we have selected five compounds, viz., compounds no. **(5,6,7, 9 and 10)** to correlate their modeling parameters with the activity logK. We have done it to find out which property or feature is highly responsible for the reduction activity of aromatic nitro compounds. The optimized molecular structures of these compounds presented in respectively and their modeling parameters are presented in TABLE 3.

TABLE 3 : Modeling parameters used in the present study

Comp. No.	TE	DpM	RMSg
5	9.62	2.312	0.0875
6	10.12	3.778	0.0965
7	11.02	4.027	0.0992
9	3.58	4.000	0.0798
10	3.33	4.301	0.0973

TE=Total Energy, DpM=Dipolmomen, RMSg=Root mean squire gradiant

Best model obtained from the modeling parameters is:

logK=0.0766 (±0.0252) TE+8.0747 (4)

n =5, Se=0.189, R=0.8692, F=9.273

Model obtained from the modeling parameter exhibits that the reduction activity is directly correlated with the Total energy of the molecule e,i. increase in the total energy increases the reduction activity numerically of the aromatic nitro compounds. It also demonstrates that the conformer having the grater stability, are less active and become less reductive in nature.

EXPERIMENTAL

Biological activity

The biological activity for the set of **(12)** compounds used in the present study, as expressed logK, were taken from the literature[1].

Topological indices

All the topological indices were calculated from the hydrogen suppressed molecular graphs. Since the calculations of these topological indices are well documented in the literature, it is not necessary to duplicate the same here.

Wiener index (W)[10]-Wiener index W=W(G) of G is defined as the half sum of the elements of the distance matrix.

$$W=W(G)=\tfrac{1}{2}\sum_{i=1}\sum_{j=1}(D)_{ij}$$

Full Paper

Where, $(D)_{ij}$ is the ijth element of the distance matrix which denotes the shortest graph-theoretical distance between sites i and j of G.

The connectivity index(χ)[11]- the connectivity index $\chi = \chi(G)$ of G is defined by Randic as

$$\chi=\chi(G)= \sum_{i,j}[d(i).d(j)]^{-0.5}$$

Balaban index (J)[12]-The Balaban index J=J(G) of G is defined as

$$J=M/\mu+1 \sum_{bonds} (d_i.d_j)^{-0.5}$$

Where M is the number of bonds in G, μ is the cyclomatic number of G, and d_i (i=1,2,3,….N; N is the number of vertices in G) is the distance sum.

The cyclomatic number $\mu=\mu(G)$ of a cyclic graph G is equal to the minimum number of edges necessary to be erased from G in order to transform it into the related acyclic graph. In case of monocyclic graph $\mu=1$ otherwise it is calculated by means of the following expression

$$M=M-N+1$$

Szeged index(Sz)[13]-the Szeged index, Sz=Sz(G), is calculated according to the following expression:

$$Sz=Sz(G)= \sum_{edges} n_u \cdot n_v$$

Where n_u is the number of vertices lying closer to one end of the edge e=uv; the meaning of n_v is analogous. Edges equidistance from both the ends of an edges, e=uv are not taken into account.

Physicochemical parameter

In present study molar refractivity (MR), molar volume(MV), parachor(Pc), index of refraction(η), surface tension(ST), density(D), and polarizability (Pol) shown in TABLE3 are tested and calculated from computer software acdlabs (Chem Skech5.0)[21].

Regression analysis

Regression analyses were made using maximum R^2 method[14-16] adopting step-wise regression.

Molecular modeling

Molecular optimization and calculation of modeling parameters were made by using software hyper chem7 (Demo version)[22].

CONCLUSION

On the basis of result and discussion made above conclusion can be drawn that the nature of the reduction activity is partially structure specific for the particular series.

For the present set of compounds reduction activity is dependent on the branching of the molecule along with the refractive parameters.

Also it is concluded that, compounds having higher total energy are less active for the present set of compounds.

ACKNOWLEDGMENT

Authors are thankful to Dr. P.V.Khadikar for his guidance and support to perform the present study.

REFERENCES

[1] C.Ansch, A.Kurup, R.Garg, H.Gao; Chem.Rev., **101**, 619 **(2001)**.
[2] M.Thakur, A.Thakur, P.V.Khadikar, C.T.Supuran; Bioorg.Med.Chem.Lett., **15(1)**, 203 **(2005)**.
[3] A.Thakur, M.Thakur, P.V.Khadikar, C.T.Supurane, P.Sudele; Bioorg.Med.Chem., **12(4)**, 789 **(2004)**.
[4] M.Thakur, A.Thakur, P.V.Khadikar; Bioorg.Med. Chem., **12(4)**, 825 **(2004)**.
[5] A.Thakur, M.Thakur, S.Vishwakarma; Bioorg.Med. Chem., **12(5)**, 1209 **(2004)**.
[6] M.Thakur, A.Thakur, P.Sudele; Indian J.Chem., **B43**, 976 **(2004)**.
[7] M.Thakur, A.Agrawal, A.Thakur, P.V.Khadikar; Bioorg.Med.Chem., **12(9)**, 2287 **(2004)**.
[8] A.Thakur, M.Thakur, N.Kakani, A.Joshi, A.Gupta; ARKIVOC, **10(1)**, 36 **(2004)**.
[9] A.Thakur, M.Thakur, P.V.Khadikar; Bioorg.Med. Chem., **11(23)**, 5203 **(2003)**.
[10] H.Wiener; J.Am.Chem.Soc., **69**, 17 **(1947)**.
[11] M.Randic; J.Am.Chem.Soc., **97**, 6609 **(1975)**.
[12] A.T.Balaban; Chem.Phys.Lett., **89**, 399 **(1982)**.
[13] P.V.Khadikar, P.P.Kale, N.V.Deshpande, S. Karmarkar, V.K.Agrawal; Commun.Math.Comput. Chem., MATCH, **43**, 7 **(2001)**.
[14] S.Chaterjee, A.S.Hadi, B.Price; 'Regression Analysis by Examples', 3rd ed., Wiley VCH, New york, **(2000)**.
[15] H.Waterbeemd, Vande (Ed); 'Methods and Prin-

ciples in Medicinal Chemistry', VCH, Weinleium (Ger.), **12**, **(1995)**.

[16] J.Devillers, W.Karcher (Eds); 'Chemical and Environmental Sciences', Klauer Academic Pub., Dorderch, Netherland, **2**, **(1991)**.

[17] S.C.Basak, C.M.Frane, M.E.Rosen, V.R. Magnuson; Med.Sci.Res., **15**, 605 **(1987)**.

[18] R.G.Parr, W.Yang; 'Density Functional Theory of Atoms and Molecules', Oxford, New York, **(1989)**.

[19] A.T.Balaban; J.Chem.Inf.Comput.Sci., **32**, 23 **(1992)**.

[20] A.T.Balaban, T.S.Balaban; Math.Chem., **8**, 383 **(1991)**.

[21] ACD-Lab software for calculating the referred physicochemical parameters, Chemsketh S.O, WWW.Acdlabs.com.

[22] HyperChem; 7 Demoversion from WWW. Hyperchem.com.