Scientific Paper

Topochemical Model for Prediction of Corticotropin Releasing Factor Antagonizing Activity of *N*-Phenylphenylglycines

Sanjay Bajaj,[a] Surinder Singh Sambi,[a] and Anil Kumar Madan[b]*

[a] *School of Chemical Technology, G. G. S. Indraprastha University, Delhi-110006, India*
[b] *Faculty of Pharmaceutical Sciences, M. D. University, Rohtak – 124001, India*
E-mail: madan_ak@yahoo.com

Received 11-02-2005

Abstract

In the present study, the relationship between the *Zagreb topochemical index* M_1^c and the corticotropin releasing factor receptor antagonizing activity of *N*-phenylphenylglycine analogs has been investigated. The values of Zagreb topochemical index M_1^c of all the analogs involved in the data set were calculated using an in-house computer program. The resulting data was analyzed and a suitable model was developed after identification of the active range. Subsequently, a biological activity was assigned to each of the compounds involved in the dataset which was then compared with the reported biological activity. Accuracy of prediction was found to be 82.3% using the said model. High predictability of the model offers vast potential for providing lead structures for development of potent CRF receptor antagonist.

Key words: topochemical model, Zagreb topochemical index, phenylphenylglycines, CRF receptor antagonist, topological

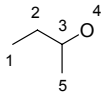
Introduction

A primary goal of any drug design strategy is to predict the activity of new compounds,[1] and one way of doing this is to exploit the information contained in the structures of the molecules. Molecular topology has been extensively used in predicting the physical as well as biological properties of various types of drugs and toxic agents.[2] In order to correlate property or activity of a molecule with its topology, one must first convert by an algorithm the information contained in the molecular graph to a numerical characteristic. A scalar numerical descriptor uniquely determined by a molecular graph is named a topological (graph-theoretic) index.[3,4] A number of topological and topochemical indices have received great attention due to their applications in structure-activity relationship studies and drug research.[5–9] Amongst the most important ones are molecular connectivity indices,[10,11] Wiener's index,[12,13] Balaban's indices,[14,15] Hosoya index,[16] Zagreb indices M_1 and M_2,[17–20] eccentric connectivity index[21–23] and eccentric adjacency index.[24] Topochemical indices that have been reported and used for structure-activity relationship studies include molecular connectivity topochemical index,[25] eccentric adjacency topochemical index,[26] Wiener's topochemical index[27,28] and superadjacency topochemical index[29,30] etc.

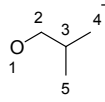
Corticotropin releasing factor (CRF; also known as corticotropin releasing hormone) is the primary physiological regulator of the hypothalamic-pituitary-adrenal axis, presiding over a large number of neuronal, endocrine, and immune processes.[31] While corticotropin releasing hormone (CRH) has been implicated in a variety of brain disorders such as ischemic injury, the molecular mechanism by which CRH elicits its activities is largely unclear.[32] Studies of CRF agonists have shown that abnormal hormonal levels may be important in neuropsychiatric disorders such as anxiety and depression, substance abuse, eating disorders, premature parturition and gastrointestinal maladies.[31]

In the present investigation, refined Zagreb index M_1 termed as Zagreb topochemical index M_1^c has been used for development of model for prediction of corticotropin releasing factor (CRF_1) antagonizing activity of *N*-phenylphenylglycine analogs. Zagreb indices are widely used in QSPR and QSAR. They are also included in a number of programs used for the routine computation of topological indices, such as POLLY, OASIS, DRAGON, Cerius, TAM, DISSIM, etc. One of the major limitations of Zagreb indices is that they do not consider the presence of heteroatom in a molecule. Therefore these indices have recently been refined and the refined Zagreb indices have been termed as Zagreb topochemical indices. These

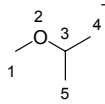
Chemical Adjacency Matrix $M_1^c(G)=\sum_{i=1}^{n}\left(d^c(i)\right)^2$



	1	2	3	4	5	d^c_i	
1	0	1	0	0	0	1	
2	1	0	1	0	0	2	
3	0	1	0	1.33	1	3.33	
4	0	0	1	0	0	1	
5	0	0	1	0	0	1	= 18.109



	1	2	3	4	5	d^c_i	
1	0	1	0	0	0	1	
2	1.33	0	1	0	0	2.33	
3	0	1	0	1	1	3	
4	0	0	1	0	0	1	
5	0	0	1	0	0	1	= 17.443



	1	2	3	4	5	d^c_i	
1	0	1.33	0	0	0	1.33	
2	1	0	1	0	0	2	
3	0	1.33	0	1	1	3.33	
4	0	0	1	0	0	1	
5	0	0	1	0	0	1	= 18.886

Figure 1. Calculation of Zagreb topochemical index M_1^c for three five vertex similarly branched structures containing only one oxygen as heteroatom.

indices are sensitive to both the presence and relative position of heteroatom(s). These indices are denoted by M_1^c and M_2^c.[33]

Zagreb Topochemical Index M_1^c: Zagreb Topochemical Index M_1^c is a modification of the Zagreb group parameter or Zagreb index M_1, which was introduced by Gutman and Trinajstic.[17] This index has recently been refined and has been termed as Zagreb topochemical index M_1^c. This index is sensitive to both the presence as well as relative position of heteroatom(s) in a molecule and has been reported to have much lower degeneracy in comparison to the original index.[33] It is defined as the summation of the squares of chemical degrees over all the vertices in hydrogen suppressed molecular graph. It is expressed by the equation:

$$M_1^c(G)=\sum_{i=1}^{n}\left(d^c(i)\right)^2$$

For hydrogen suppressed molecular graph (G), v_1, $v_2 ... v_n$ are vertices, n is the number of vertices and the number of first neighbors of a vertex v_i is the chemical degree of this vertex and is denoted by $d^c(i)$.

The Zagreb topochemical index M_1^c can be easily calculated from the chemical adjacency matrix of hydrogen suppressed molecular structure. Chemical degree for a vertex i is the sum of entries in a row i of chemical adjacency matrix. When the adjacency matrix is weighted corresponding to the heteroatom present within the molecule, the matrix may be termed as chemical adjacency matrix and the degree of a vertex obtained from such matrix is called chemical degree of that vertex. This matrix is obtained by substituting, row elements corresponding to heteroatom(s), with relative atomic weight with respect to carbon atom. Thus, in this matrix the non-zero row elements of an adjacency matrix represent chemical adjacency between the corresponding vertices in a molecular graph.[33] Calculation of M_1^c for three five vertex similarly branched structures containing only one oxygen as heteroatom has been exemplified in Figure 1.

Model Design and Analysis

A data set comprising of 39 *N*-phenylphenylglycine analogs[31] was selected for the present investigations. The original dataset comprised of 40 compounds, but two compounds had same structure as well as activity and hence one of those was not included in the investigation. The basic structure for these analogues is depicted in Figure 2 and various substituents enlisted in Table 1.

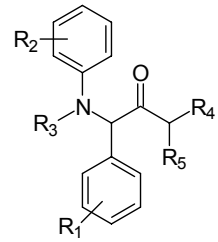
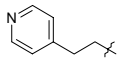
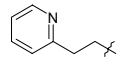
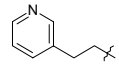
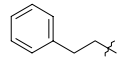
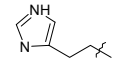
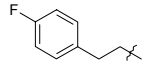
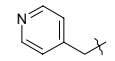
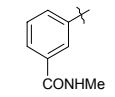
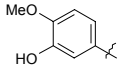
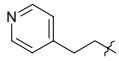
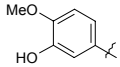
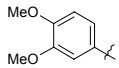
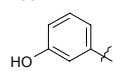
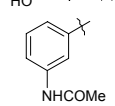


Figure 2. Basic structure of *N*-phenylphenylglycine analogs.

The values of the *Zagreb topochemical index* M_1^c were computed for all the analogues involved in the data set using an in-house computer program. The resulting data was analyzed and suitable model was developed after identification of the active range by maximization of the moving average with respect to the active compounds (<35% = inactive, 35–65% = transitional, ≥65% = active).[24, 29] Subsequently, each analogue was assigned a biological activity using this model which was then compared with the reported CRF_1 antagonizing activity. The activity was reported[31] in terms of binding affinity expressed as K_i (nM). The analogues possessing K_i (nM) values of <1000 were considered to be active and analogues possessing K_i (nM) values ≥1000 were considered to be inactive for the purpose of the present study. The percentage degree of prediction of a particular range was derived from the ratio of the number of compounds predicted correctly to the total number of compounds present in that range. The overall degree of prediction was calculated from

Table 1. Relationship between Zagreb topochemical index M^c_1 and CRF_1 antagonizing activity (CRF_1A) of *N*-phenylphenylglycine analogs.

Comp. No.	Substituent			M^c_1	CRF_1A Activity	
	R_1	R_2	R_4		Assigned	Reported[31]
1	2-Br	2-Cl, 4-CH_3		227.259	–	–
2	2-Cl	2-Cl, 4-CH_3	- do -	176.724	–	–
3	2-F	2-Cl, 4-CH_3	- do -	164.98	–	–
4	2-$CHCH_2$	2-Cl, 4-CH_3	- do -	165.142	–	+
5	2-CH_3	2-Cl, 4-CH_3	- do -	161.142	–	+
6	2-$(CH_3)_2$	2-Cl, 4-CH_3	- do -	171.142	–	–
7	2-OCH_3	2-Cl, 4-CH_3	- do -	168.028	–	–
8	3-Cl	2-Cl, 4-CH_3	- do -	176.724	–	–
9	3-CH_3	2-Cl, 4-CH_3	- do -	161.142	–	–
10	3-OCH_3	2-Cl, 4-CH_3	- do -	168.028	–	–
11	4-Cl	2-Cl, 4-CH_3	- do -	176.724	–	–
12	4-CH_3	2-Cl, 4-CH_3	- do -	161.142	–	–
13	2,5-di-OCH_3	2-Cl, 4-CH_3	- do -	180.914	±	–
14	2,4-di-OCH_3	2-Cl, 4-CH_3	- do -	180.914	±	–
15	2,6-di-OCH_3	2-Cl, 4-CH_3	- do -	180.914	±	–
16	H	2-Cl, 4-CH_3	- do -	155.142	–	–
17	2-Br	2-Cl	- do -	221.259	–	–
18[a]	2-Br	2-Cl	- do -	227.621	–	–
19	2-Br	2,4-di-Cl	- do -	242.841	–	+
20	2-Br	2-Cl, 4-CH_3		227.593	–	–
21	2-Br	2-Cl, 4-CH_3		227.259	–	–
22	2-Br	2-Cl, 4-CH_3		225.867	–	–
23	2-Br	2-Cl, 4-CH_3		225.041	–	–
24	2-Br	2-Cl, 4-CH_3		235.705	–	–
25	2-Br	2-Cl, 4-CH_3		223.259	–	–
26	2-Br	2-Cl, 4-CH_3		241.479	–	–
27	2-Br	2-Cl, 4-CH_3		238.862	–	–
28	2-CH_2CH_3	2,4-di-Cl		180.724	±	+
29	2,6-di-CH_3	2,4-di-Cl	- do -	182.724	+	+
30	2-CH_3	2,4-di-Cl	- do -	176.724	–	+
31	2-CH_3, 4-F	2,4-di-Cl	- do -	186.562	+	+
32	2-CF_3	2,4-di-Cl	- do -	211.775	–	–
33	2-CH_3	2-F, 4-CH_3	- do -	149.398	–	+
34	2-CH_3	2-CF_3, 4-Cl	- do -	196.193	+	+
35	2-CH_3	2-CF_3, 4-CH_3	- do -	180.611	±	+
36	2-CH_3	2,4-di-Cl		188.327	+	+
37	2-CH_3	2,4-di-Cl		193.104	+	–
38	2-CH_3	2,4-di-Cl		175.441	–	–
39	2-Br	2,4-di-Cl		191.612	+	+

+ active compounds (compounds having *Ki* (nM) < 1000), – inactive compounds, ± compounds in transitional range, R_3 = H unless otherwise specified, a - R_3 = -CH_3, R_5 = H, - do - = same as above.

the ratio of the total number of compounds predicted correctly to that of the total number of compounds present in both the active and inactive ranges. The results are summarized in Tables 1 and 2.

Discussion and Conclusion

Role of CRH and related peptides has been elucidated in inflammatory and allergic disorders, neurological diseases and in pre-term labor.[34] Investigations of CRH type I receptor non-peptide antagonists suggest therapeutic potential for treatment of these and other neuropsychiatric diseases. These antagonists may also be effective in treating more common somatic diseases. Patients with obesity and metabolic syndrome who often have subtle but chronic hypothalamic-pituitary-adrenal hyperactivity, which may reflect central dysregulation of CRH and consequently glucocorticoid hyper-secretion, could possibly be treated by administration of CRH receptor type I antagonists.[35]

In recent years, a large number of topological indices of diverse nature have been proposed but only handful of them have been successfully employed in SAR studies. One of the limitations of the topological indices is their degeneracy. Topochemical versions of the topological indices, having applicability in SAR, are being developed to overcome this limitation. Unlike, the topological indices, these topochemical indices take into consideration not only the presence but also the relative position of the heteroatom(s) thereby reducing the degeneracy to a large extent. Zagreb topochemical index M_1^c is one such topochemical index which is a refined form of Zagreb index M_1. Both topological and topochemical indices have been widely used for the development of models for prediction of biological activity of diverse series of compounds.

Though all the analogues in the dataset reportedly[31] possess varying degree of biological activity but analogues possessing K_i (nM) values of <1000 were considered to be active and analogues possessing K_i (nM) values ≥ 1000 were considered to be inactive for the purpose of the present study. The proposed topochemical model is based upon recently introduced Zagreb topochemical index M_1^c. Retrofit analysis of the data presented in the Tables 1 and 2 reveals the following information with regard to model based upon Zagreb topochemical index M_1^c.

- Biological activity was assigned to 34 out of the total 39 compounds. Out of these 28 (82.30%) were predicted correctly with respect to CRF_1 antagonizing activity.
- The lower inactive range had Zagreb topochemical index M_1^c values of less than 176.73 and the upper inactive range had Zagreb topochemical index M_1^c values greater than 196.20. 23 out of 28 compounds (82.1%) in both the inactive ranges were predicted correctly. The average K_i (nM) of all the correctly predicted compounds in both the inactive ranges was 7561.83.
- The active range had Zagreb topochemical index M_1^c values from 182.72 to 196.20. Biological activity of 5 out of the 6 compounds (83.3%) in the active range was predicted correctly. The average K_i (nM) of the correctly predicted compounds in the active range was only 444.0 indicating high potency of this range.
- The lower inactive range was ideally separated from the active range by transitional range. The transitional range had Zagreb topochemical index M_1^c values from 176.73 to 182.71.
- For estimation of K_i (nM), the following equation was developed

$$K_i\,(\mathrm{nM})^{\mathrm{Cal}} = 1.0285\,(M_1^c)^2 - 402.41(M_1^c) + 40149.1$$

In the proposed model, the lower inactive range is ideally separated from the active range by a transitional range, which indicates the gradual shift in biological activity as one proceeds from the active to lower inactive range and *vice versa*. Predictability of 83% in the active range itself is highly beneficial because this range is responsible for providing lead structures. Analysis of the structures of the compounds in the active range indicates a general trend for activity, i.e. (i) 2-methyl as substituent R_1, (ii) 4-chloro as substituent R_2 and (iii) compounds having ring A as substituent R_4.

High predictability of the proposed model offers vast potential for providing lead structures for development of potent CRF_1 receptor antagonist.

Table 2. Proposed model for CRF_1 antagonizing activity of *N*-phenylphenylglycines.

Model Index	Nature of Range in Proposed Model	Index Value	Number of Compounds falling in the range		Percent Accuracy	Average K_i (nM)	
			Total	Correct		Total	Correct
M_1^c	Lower Inactive	< 176.73	15	11	73.3	6274.9	8331.0
	Transitional	176.73 – 182.71	5	N.A.[a]	N.A.	6194.2	N.A.
	Active	182.72 – 196.20	6	5	83.3	2036.7	444.0
	Upper Inactive	> 196.20	13	12	92.3	6285.8	6792.7

[a] Not applicable.

References

1. A. R. Ortiz, A. T. Pisabarro, G. Federico, R. C. Wade, *J. Med. Chem*. **1995**, *38*, 2681–2691.
2. C. Roychaudhary, S. C. Basak, A. B. Roy, J. J. Ghosh, *Indian Drugs* **1980**, Dec. 97–102.
3. I. Gutman, O. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer-Verlag: Berlin 1986.
4. N. Trinajstic, *Chemical Graph theory*, 2nd Ed.; CRC Press: Boca Raton, FL, 1992.
5. D. H. Rouvray, In *Chemical Applications of Graph Theory*, ed. A.T. Balaban, Academic Press, London, 1976, pp180–181.
6. L. B. Kier, L. H. Hall, *Molecular Connectivity in Structure-Activity Analysis*, Research Study Press: Letchworth. 1986.
7. P. J. Hansen, P. C. Jurs, *J. Chem. Edu.* **1988**, *65*, 574–580.
8. A. R. Katritzky, E. V. Gordeeva, *J. Chem. Inf. Comput. Sci*. **1993**, *33*, 835–857.
9. E. Estrada, A. Ramfrez, *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 837–843.
10. M. Randic, *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615.
11. L. B. Kier, L. H. Hall, W. J. Murray, R. Milan, *J. Pharm. Sci.* **1975**, *64*, 1971–1974.
12. H. Wiener, *J. Am. Chem. Soc*. **1947**, *69*, 17–20.
13. H. Wiener, *J. Am. Chem. Soc*. **1947**, *69*, 2636–2638.
14. A. T. Balaban, *Pure & Appl. Chem*., **1983**, *55*, 199–206.
15. A. T. Balaban, *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 334–343.
16. H. Hosoya, *Bull. Chem. Soc. Jpn*. **1971**, *44*, 2332–2337.
17. I. Gutman, N. Trinajstic, *Chem. Phys. Lett*. **1973**, *17*, 535–545.
18. S. Nikolic, G. Kovacevic, A. Millicevic, N. Trinajstic, *Croat. Chem. Acta*. **2003**, *76*, 113–124.
19. S. Nikolic, I. M. Tolic, N. Trinajstic, I. Baucic, *Croat. Chem. Acta.* **2000**, *73*, 909–921.
20. S. Bajaj, S. S. Sambi, A. K. Madan, *Bioorg. Med. Chem*. **2004**, *12*, 3695–3705.
21. V. Sharma, R. Goswami, A. K. Madan, *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 273–282.
22. S. Sardana, A. K. Madan, *J. Mol. Model.* **2002**, *8*, 258–265.
23. S. Sardana, A. K. Madan, *MATCH Commun Math. Comput. Chem.* **2001**, *43*, 85–98.
24. S. Gupta, M. Singh, A. K. Madan, *J. Computer-Aided Mol. Design* **2001**, *15*, 671–678.
25. H. Dureja, A. K. Madan, *J. Mol. Mod.* – in press.
26. S. Gupta, M. Singh, A. K. Madan, *Ind. J. Chem*. **2003**, *42A*, 1414–1425.
27. S. Bajaj, S. S. Sambi, A. K. Madan, *J. Mol. Str. (Theochem.)* **2004**, *684*, 197–203.
28. S. Bajaj, S. S. Sambi, A. K. Madan, *MATCH Commun Math. Comput. Chem.* Accepted for publication.
29. S. Bajaj, S. S. Sambi, A. K. Madan, *QSAR and Comb. Sci*. **2004**, *23*, 506–514.
30. S. Bajaj, S. S. Sambi, A. K. Madan, *Bioorg Med. Chem. Lett.* **2005**, *15*, 467–469.
31. M. Valentina, J. Penzotti, D. M. Wilson, A. P. Termin, L. Mao, C. M. Crane, F. Hassman, T. Wang, H. Wong, K. J. Miller, S. Grossman, D. J. Grootenhuis, *J. Med. Chem.* **2004**, *47*, 2426–2429.
32. W. Wang, M. Solc, P. Ji, K. E. Dow *Neurosci Lett.* **2004,** *23*, *371*(2–3), 133–137.
33. S. Bajaj, S. S. Sambi, A. K. Madan, *Croat. Chem. Acta.* **2005**, *78*, 165–174.
34. D. K. Grammatopoulos, G. P. Chrousos, *Trends Endocrinol. Metab.* **2002**, *13*, 436–444.
35. C. Contoreggi**,** K. C. Rice, G. Chrousos, G. *Neuroendocrinology* **2004**, *80*, 111–123.

Povzetek

Raziskali smo zvezo med Zagrebškim topokemijskim indeksom M_1^c in CRF receptorjem (receptorjem sprostitvenega dejavnika kortikotropina), ki zavira aktivnost analogov N-fenil-fenil glicina.Vrednosti tega indeksa za vse analoge smo izračunali s pomočjo lastnega računalniškega programa. Dobljene podatke smo analizirali in po ugotovitvi območja aktivnosti postavili ustrezen model. Na osnovi podatkov smo razvili primeren matematični model, v katerem je bila upoštevana biološka aktivnost vseh komponent vključenih v podatke, rezultate izračunov pa smo primerjali z objavljenimi aktivnostmi. Zanesljivost napovedi na osnovi modela je 82,3%, kar predstavlja velik potencial za razvoj spojin vodnic za učinkovite antagoniste CRF receptorjev.