Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 23 (2005) 339–345
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Topological models for the prediction of HIV-protease inhibitory activity of tetrahydropyrimidin-2-ones

Viney Lather, Anil K. Madan*

Faculty of Pharmaceutical Sciences, M. D. University, Rohtak-124001, India

Received 7 June 2004; received in revised form 21 October 2004; accepted 2 November 2004
Available online 8 December 2004

Abstract

Relationship between the topological indices and HIV-protease inhibitory activity of tetrahydropyrimidine-2-ones has been investigated. Three topological indices, Wiener's index—a distance based topological descriptor, Zagreb group parameter—an adjacency based topological descriptor and eccentric connectivity index—an adjacency-cum-distance based topological descriptor were used for the present investigations. A dataset comprising of 80 substituted tetrahydropyrimidine-2-one analogues was selected for the present studies. The values of the *Wiener's index*, *Zagreb group parameter* and *eccentric connectivity index* for each of the 80 compounds comprising the dataset were computed using an in-house computer program. The dataset was divided randomly into training and test sets. Resultant data was analyzed and suitable models were developed after identifying the active ranges in the training set. Subsequently, a biological activity was assigned to each of the compound involved in the test set using these models, which was then compared with the reported HIV-protease inhibitory activity. Accuracy of prediction using these models was found to vary from a minimum of ~86% to a maximum of ~88%.

Keywords: Topological indices; Wiener's index; Eccentric connectivity index; Zagreb group parameter; HIV-protease

1. Introduction

The human immunodeficiency virus (HIV) encodes an aspartyl protease which is responsible for the processing of the gag and gag–pol gene products. This processing is required for the production of mature, infectious virions and has been a prime target for the intervention [1–5].

The HIV-protease, encoded in the 5′ end of the pol gene, is expressed as part of the gag–pol protein. This gene encodes a 99-amino acid protein. Homodimers of this protein have the aspartyl protease activity that is typical of retroviral proteases; monomers are enzymatically inactive [6]. The enzyme's targets are amino acid sequences in the gag and gag–pol polyproteins, which must be cleaved before nascent viral particles (virions) can mature [7–9]. Cleavage of the gag polyprotein produces three large proteins (p24, p17 and p7) that contribute to the structure of the virion and RNA packaging, and three smaller proteins (p6, p2 and p1) of uncertain function. Although mammalian cells contain aspartyl proteases, none efficiently cleave the gag polyprotein [10,11]. Three of the HIV-cleavage sites are phenylalanine–proline or tyrosine–proline bonds, which are unusual sites of attack for mammalian proteases [12]. Proteolytic cleavage of the gag polyprotein results in morphologic changes in the virion and condensation of the nucleoprotein core. The protease is packaged into virions, and the cleavage events it catalyzes, occur simultaneously with or soon after the budding of the virion from the surface of an infected cell [13]. Proviral DNA lacking functional protease produces immature viral particles [9].

HIV-protease inhibitors prevent cleavage of the gag and gag–pol protein precursors in acutely and chronically infected cells, arresting maturation and thereby blocking the infectivity of nascent virions. The main antiviral action of HIV-protease inhibitors is thus to prevent subsequent waves of infection; they have no effect on cells already harboring integrated proviral DNA [11]. The approved HIV-protease inhibitors are based on amino acid sequences recognized and cleaved in HIV proteins. Indinavir,

* Corresponding author. Tel.: +91 1262 272535.
E-mail address: madan_ak@yahoo.com (A.K. Madan).

1093-3263/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2004.11.005

nelfinavir, ritonavir, saquinavir, lopinavir, amprenavir and the most recently approved atazanavir are structurally related molecules [14]. Most contain a synthetic analogue of the phenylalanine–proline sequence at positions 167 and 168 of the gag–pol polyprotein that is cleaved by the protease [12]. However such compounds possess poor pharmacokinetics and are complex and expensive to synthesize. To overcome such difficulties, significant efforts have been devoted to the development and SAR study of non-peptide inhibitors such as cyclic ureas and other heterocycles.

The characterization of structure in studying relationship between molecular structure and properties has attracted considerable attention in recent years [15–20]. It is now generally accepted that the physicochemical, biomedicinal, and environmental, as well as the toxicological properties of chemicals are determined by certain aspects of their structure [21–23]. In spite of the early recognition of this principle by Crum-Brown and Fraser, the quantitative relationship between structure and properties of molecules remained elusive for almost a century [24]. The central problem has been the difficulty of quantifying the intuitive concept of ''structure''.

Molecular topology overcomes the inherent problem in structure–activity relationship (SAR) studies to quantify chemical structures by translation of chemical structures into characteristic numerical descriptors [25–30]. The structure of a molecule can be looked upon as the mode of organization of an assembled entity where some parts (e.g. atoms or atomic cores, more correctly) are involved in a binary relationship: any two atoms, X and Y, in a molecule are either bonded or not bonded. Such a binary relationship, depicting the basic connectivity of atoms in molecules, is satisfactorily represented by a graph G = (V, R), where the nonempty vertex set V symbolizes the set of atomic cores and the set R (often called the edge set, E) represents the set of chemical bonds. Graphs can be analytically represented by matrices from which a single topological index (TI) or a set of them can be derived [31]. TIs are sensitive to such structural features as size, shape, bond order, branching, and neighborhood patterns of atoms in the molecule [32]. These indices, if are well chosen, are thus a good characterization of the molecular structure [31]. Although a number of topological indices have been reported, only a handful of them have been successfully employed in SAR studies. *Hosoya's index* [33,34], *Randic's molecular connectivity index*, χ [35,36], *the higher-order connectivity indices*, ${}^{n}\chi$, for the paths of length n defined by Kier and Hall [27], *Balaban's index*, J [37–40], *Wiener's index* [41,42], *Zagreb group parameters*, M_1 and M_2 [43], *eccentric connectivity index* [44–46] are some of the topological indices used in the SAR studies.

In the present study, relationship of *Wiener's index*—a distance-based topological descriptor, *Zagreb group parameter*—an adjacency-based topological descriptor and *eccentric connectivity index*—an adjacency-cum-distance based topological descriptor with the HIV-protease inhibitory activity of tetrahydropyrimidin-2-ones has been investigated by the development of suitable models.

2. Methodology

2.1. Calculations of topological indices

The *Wiener's index* [47–52], a well-known distance-based topological index is defined as half sum of the distances between all the pairs of vertices in a hydrogen-suppressed molecular graph, that is

$$W = \frac{1}{2}\left(\sum_{i=1}^{n} Pi\right) \tag{1}$$

where Pi is the length of the path that contains the least number of edges between vertex i and vertex j in graph G and n is the maximum possible number of i and j.

The *Zagreb group parameter* M_1 proposed by Gutman et al. [53,54] is defined as the sum of squares of degree over all vertices and is represented by following equation:

$$M_1 = \sum_{i=1}^{n} (V_i^2) \tag{2}$$

where V_i is the degree of vertex i in a hydrogen-suppressed molecular structure. The vertex degree V_i for a vertex i is given as the sum of the entries in a row i of adjacency matrix.

The *eccentric connectivity index* [55] denoted by ξ^c is defined as the summation of the product of eccentricity and the degree of each vertex in the hydrogen suppressed molecular graph having n vertices:

$$\xi^c = \sum_{i=1}^{n} (E_i * V_i) \tag{3}$$

where V_i is the degree of vertex i, E_i the eccentricity of the vertex i and n the number of the vertices in graph G. The eccentricity E_i of a vertex i in a graph G is the path length from vertex i to vertex j that is the farthest from i (E_i = max - $d(ij)$; $j \varepsilon$G); the eccentric connectivity index takes into consideration the eccentricity as well as valency of the vertices in a hydrogen-suppressed graph.

2.2. Model development

A dataset comprising of 80 analogues of tetrahydropyrimidin-2-one was selected for the present investigations [56]. The basic structure for these analogues is depicted in Fig. 1 and various substituents enlisted in Table 1. The dataset comprised of both active and inactive compounds.

The values of the *Wiener's index* were computed for each analogue using an in-house computer program. The dataset was divided randomly into two sets. Compounds having odd serial number were designated as test set and those having even number were separated as training set.

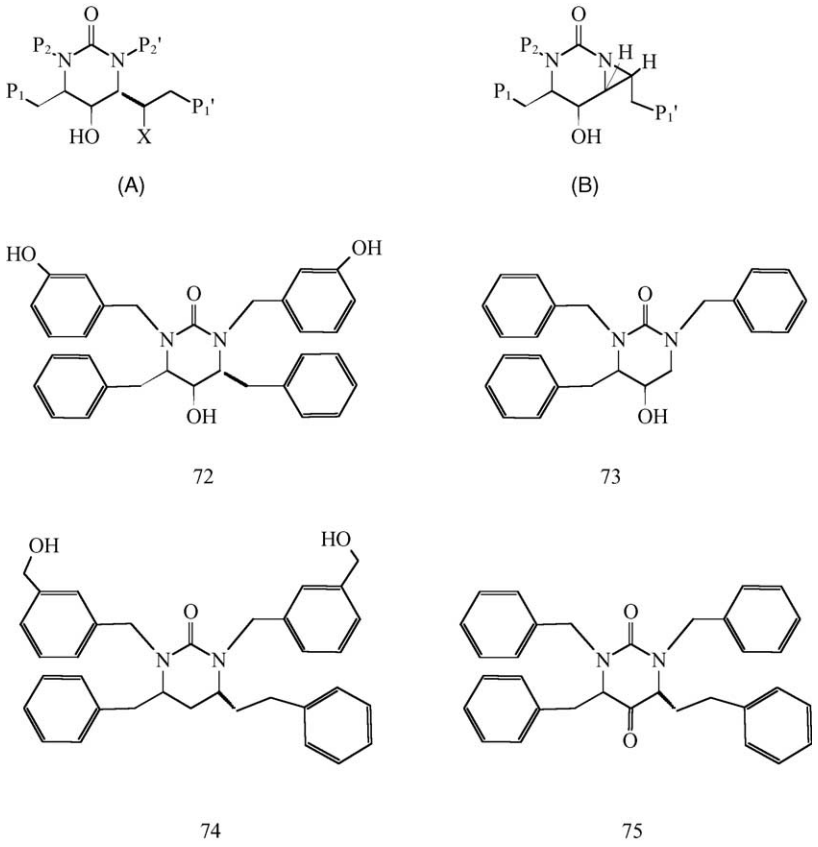

Fig. 1. Basic structures of tetrahydropyrimidine-2-one analogues.

For the selection and evaluation of range specific features, exclusive activity ranges were discovered from the frequency distribution of response level. Resultant data in the training set was analyzed and a suitable model was developed after identification of active range by maximization of the moving average with respect to the active compounds (<35% = inactive, 35–65% = transitional, >65% = active) [57]. Subsequently, each analogue in the test set was assigned a biological activity using this model, which was then compared with the reported HIV-protease inhibitory activity. HIV-protease inhibitory activity was reported quantitatively as K_i values at different concentrations. The analogues possessing K_i values of $\leq$1 nM were considered to be active and analogues possessing K_i values of >1 nM were considered to be inactive for the purpose of 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 derived from the ratio of the total number of compounds correctly to that of the total number of compounds present in both the active and inactive ranges.

Aforementioned procedure was similarly adopted for *Zagreb group parameter* M_1 and *eccentric connectivity index*, ξ^c. The results are summarized in Tables 1–4.

3. Results and discussion

Inhibitors of human immunodeficiency virus (HIV)-encoded protease, combined with nucleoside analogues with antiretroviral activity, cause profound and sustained suppression of viral replication, reduce morbidity, and prolong life in patients with HIV infection. Recent guidelines recommend that initial treatment of all HIV-infected patients include the administration of an HIV-protease inhibitor [58]. This enzyme is therefore a major target for the structure based inhibitor design.

Efficient discovery and creation of novel drug molecules depend on the ability to explore and quantify the relationships between molecular structure and function—particularly the biological activity. The inherent problem in the development of a suitable correlation between chemical structures and biological activity can be attributed to the non-quantitative nature of chemical structures. Graph theory was successfully employed through the translation of chemical structures into characteristic numerical descriptors by resorting to graph invariants [26,29].

In the present study, relationship of *Wiener's index*—a distance-based topological descriptor, *Zagreb group parameter*—an adjacency-based topological descriptor and *eccentric connectivity index*—an adjacency-cum-distance based topological descriptor with the HIV-protease

Table 1
Index values and activity of training set of HIV-protease inhibitors tetrahydropyrimidine-2-ones

Compound no.	P_2	P_2'	X	W	M_1	ξ[a]	Reported activity
A_2	3-Cyanobenzyl	3-Cyanobenzyl	H	5422	212	1018	–
A_4	3-Cyano-4-fluorobenzyl	3-Cyano-4-fluorobenzyl	H	6216	224	1079	–
A_6	3-Hydroxymethylbenzyl	3-Hydroxymethylbenzyl	H	5470	212	1023	+
A_8	3-(Carboxamido)benzyl	3-(Carboxamido)benzyl	H	6620	224	1079	+
A_{10}	3-(Carboxamido)benzyl[b]	3-(Carboxamido)benzyl[b]	H	7872	248	1245	+
A_{12}	3-(Carboxamido oxime)benzyl	3-(Carboxamido oxime)benzyl	H	7086	232	1207	+
A_{14}	3-(Carboxamido oxime)benzyl[b]	3-(Carboxamido oxime)benzyl[b]	H	8900	256	1361	+
A_{16}	3-Aminobenzyl	3-Aminobenzyl	H	4704	204	919	–
A_{18}	4-Amino-3-fluorobenzyl	4-Amino-3-fluorobenzyl	H	5462	216	1023	–
A_{20}	3-(Pyrazol-3-yl)benzyl	3-(Pyrazol-3-yl)benzyl	H	7982	252	1333	–
A_{22}	Indazol-6-yl-methyl	Indazol-6-yl-methyl	H	6202	236	1137	+
A_{24}	3-Aminoindazol-5-yl-methyl	3-Aminoindazol-5-yl-methyl	H	6933	248	1187	+
A_{26}	3-(5-Methyl-2-pyridylcarboxamido)-benzyl	3-(5-Methyl-2-pyridylcarboxamido)-Benzyl	H	14338	300	2209	+
A_{28}	Cyclopropylmethyl	Cyclopropylmethyl	H	2594	168	699	–
A_{30}	Cyclopropylmethyl	Cyclopropylmethyl	OH	2726	174	712	–
A_{32}	Benzyl	H	OH	2686	160	718	–
A_{34}	Benzyl	Benzyl	Br	4348	198	892	–
A_{36}	2-Naphthylmethyl	2-Naphthylmethyl	H	7062	244	1269	–
A_{38}	2-Naphthylmethyl	2-Naphthylmethyl	F	7352	250	1292	–
A_{40}	2-Naphthylmethyl	2-Naphthylmethyl	[c]	7062	244	1269	–
A_{42}	3-(*N*-methyl-amino)benzyl	3-(*N*-methyl-amino)benzyl	Br	5718	218	1044	–
A_{44}	3-Hydroxybenzyl	3-Hydroxybenzyl	F	4980	210	944	–
A_{46}	3-Hydroxymethylbenzyl	3-Hydroxymethylbenzyl	Br	5718	218	1044	–
A_{48}	3-Cyanobenzyl	H	H	3012	164	779	–
A_{50}	3-Hydroxybenzyl	H	H	2829	155	809	–
A_{52}	3-(Carboxamido)benzyl	H	H	3276	170	808	–
A_{54}	3-(Carboxamido oxime)benzyl	H	H	3573	174	865	+
A_{56}	3-Aminoindazol-5-yl-methyl	H	H	3534	182	867	+
A_{58}	H	Indazol-5-yl-methyl	H	3247	176	792	–
A_{60}	Benzyl	3-Cyano-4-fluorobenzyl	H	4794	208	928	–
A_{62}	3-(Carboxamido-oxime)benzyl[a]	Cyclopropylmethyl	H	4492	200	969	+
A_{64}	3-Aminoindazol-5-yl-methyl	Benzyl	H	5458	220	1041	+
A_{66}	3-Aminobenzyl	Indazol-5-yl-methyl	H	5447	220	1029	+
A_{68}	3-Aminobenzyl	3-Cyanobenzyl	H	5103	208	974	–
A_{70}	3-Aminobenzyl	3-(Carboxamido)benzyl	H	5492	214	1025	+
72[d]	–	–	–	4346	200	868	–
74[d]	–	–	–	5249	206	1006	–
B_{76}	H	–	–	1260	126	499	–
B_{78}	3-Cyanobenzyl	–	–	2908	174	753	–
B_{80}	3-(Carboxamido oxime)benzyl	–	–	3453	184	841	–

+, Active compound; −, inactive compound.
[a] The P_1/P_1' phenyl groups contain a 4-fluoro substituent.
[b] The P_1/P_1' phenyl groups contain a 3,4-difluoro substituent.
[c] The P_1' phenyl group contains a trans-styrene substituent.
[d] Structures shown in Fig. 1.

Table 2
Proposed model based on training set of tetrahydropyrimidine-2-ones

Model index	Nature of range in the proposed model	Index value
W	Inactive	<5447
	Active	5447–6933
	Transitional	>7062
M_1	Inactive	<212
	Transitional	212–218
	Active	220-300
ξ^c	Inactive	<1025
	Active	1025–1245
	Transitional	>1245

W = Wiener's index; M_1 = Zagreb index; ξ^c = eccentric connectivity index.

inhibitory activity of tetrahydropyrimidin-2-ones has been investigated. The models (Table 2) were developed using the training set of the compounds (Table 1) and evaluated using the test set (Table 3). The accuracy of prediction (Table 4) is based upon the compounds in the test set only. The recent work of Katritzky et al. on tetrahydropyrimidin-2-ones proposed various QSAR models [59]. Also Nair et al. performed a computational chemistry study on tetrahydropyrimidine-2-ones and correlated the inhibitor–enzyme complexation energies, inhibitor solvation energies and both polar and nonpolar buried surface areas with the observed values of the binding affinity [60]. The methodology used in the present studies is entirely a

Table 3
Cross-validation test with regard to HIV-protease inhibitory activity of tetrahydropyrimidine-2-ones using proposed models based upon Wiener's index, Zagreb group parameter and eccentric connectivity index with HIV-protease inhibitory activity

Compound no.	P_2	P_2'	X	W	M_1	ξ^c	HIV-protease inhibitory activity			
							Predicted			Reported
							W	M_1	ξ^c	
A_1	H	H	H	1292	116	489	–	–	–	–
A_3	3-Cyanobenzyl[a]	3-Cyanobenzyl[a]	H	6244	224	1107	+	+	+	–
A_5	3-Acetylbenzyl	3-Acetylbenzyl	H	6220	224	1079	+	+	+	+
A_7	3-Carboxybenzyl	3-Carboxybenzyl	H	6220	224	1079	+	+	+	+
A_9	3-(Carboxamido)benzyl[a]	3-(Carboxamido)benzyl[a]	H	7088	236	1191	±	+	+	+
A_{11}	3-(Carboxamido)-4-fluorobenzyl	3-(Carboxamido)-4-fluorobenzyl	H	7002	236	1135	±	+	+	–
A_{13}	3-(Carboxamido oxime)benzyl[a]	3-(Carboxamido oxime)benzyl[a]	H	8016	244	1301	±	+	±	+
A_{15}	3-(Carboxamido oxime)-4-fluorobenzyl	3-(Carboxamido oxime)-4-fluorobenzyl	H	7908	244	1265	±	+	±	+
A_{17}	3-Amino-4-fluorobenzyl	3-Amino-4-fluorobenzyl	H	5462	216	1023	+	±	–	+
A_{19}	3-(*N*-Methyl-amino)benzyl	3-(*N*-methyl-amino)benzyl	H	5470	212	1023	+	±	–	–
A_{21}	Indazol-5-yl-methyl[a]	Indazol-5-yl-methyl[a]	H	6202	236	1137	+	+	+	+
A_{23}	(3-Methylindazol-5-yl)-methyl	(3-Methylindazol-5-yl)-methyl	H	6933	248	1187	+	+	+	+
A_{25}	3-Aminobenzisoxazol-5-yl-methyl	3-Aminobenzisoxazol-5-yl-methyl	H	6933	248	1187	+	+	+	+
A_{27}	3-(*N*-2-thiazolylcarboxamido)-benzyl	3-(*N*-2-thiazolylcarboxamido)-benzyl	H	11476	280	1787	±	+	±	+
A_{29}	Cyclopropylmethyl	Cyclopropylmethyl	F	2726	174	712	–	–	–	–
A_{31}	Benzyl	H	H	2519	154	699	–	–	–	–
A_{33}	Benzyl	Benzyl	H	4041	192	869	–	–	–	–
A_{35}	Benzyl	Benzyl	c	4041	192	869	–	–	–	–
A_{37}	2-Naphthylmethyl	2-Naphthylmethyl	N_3	8071	258	1344	±	+	±	–
A_{39}	2-Naphthylmethyl	2-Naphthylmethyl	Br	7352	250	1292	±	+	±	–
A_{41}	3-(*N*-methyl-amino)benzyl	3-(*N*-methyl-amino)benzyl	H	5470	212	1023	+	±	–	–
A_{43}	3-Hydroxybenzyl	3-Hydroxybenzyl	H	4752	204	925	–	–	–	+
A_{45}	3-Hydroxymethylbenzyl	3-Hydroxymethylbenzyl	H	5470	212	1023	+	±	–	+
A_{47}	Benzyl	H	H	2519	154	699	–	–	–	–
A_{49}	3-Cyano-4-fluorobenzyl	H	H	3274	170	808	–	–	–	–
A_{51}	3-Aminobenzyl	H	H	2829	155	809	–	–	–	–
A_{53}	3-(Carboxamido)-4-fluorobenzyl	H	H	3438	176	813	–	–	–	–
A_{55}	3-(Carboxamido oxime)-4-fluorobenzyl	H	H	3844	180	894	–	–	–	–
A_{57}	Indazol-5-yl-methyl	H	H	3267	176	838	–	–	–	–
A_{59}	H	3-Aminobenzyl	H	2734	160	688	–	–	–	–
A_{61}	3-Cyano-4-fluorobenzyl	Benzyl	H	5054	208	982	–	–	–	–
A_{63}	Benzyl	3-Aminoindazol-5-yl-mehtyl	H	5450	220	1029	+	+	+	+
A_{65}	Indazol-5-yl-mehtyl	3-Aminobenzyl	H	5451	220	1037	+	+	+	+
A_{67}	3-Aminobenzyl	3-Methylindazol-5-yl-mehtyl	H	5803	226	1056	+	+	+	+
A_{69}	3-Aminobenzyl	3-Aminoindazol-5-yl-methyl	H	5803	226	1056	+	+	+	+
A_{71}	3-Aminobenzyl	3-(Carboxamido oxime)benzyl	H	5850	218	1052	+	±	+	+
73[b]	–	–	–	2228	150	626	–	–	–	–
75[b]	–	–	–	4117	192	888	–	–	–	–
B_{77}	Benzyl	–	–	2431	164	676	–	–	–	–
B_{79}	3-(Carboxamido)benzyl	–	–	3164	180	780	–	–	–	–

+, Active compound; –, inactive compound; ±, compound in the transitional range where activity could not be specifically assigned. The P_1/P_1' phenyl groups contain a 3,4-difluoro substituent, the P_1' phenyl group contains a trans-styrene substituent.

[a] The P_1/P_1' phenyl groups contain a 4-fluoro substituent.

[b] Structures shown in Fig. 1.

different one and aims at the development of suitable models for providing lead molecules through exploitation of the active ranges in the proposed models based on topological indices.

Retrofit analysis of the data in Tables 3 and 4 reveals the following information with regard to different models used in the present study.

Model based upon *Wiener's index*:

- A total of 29 out of the 33 test compounds were classified correctly in both the active and inactive ranges. The overall accuracy of prediction was found to be 87.8%, with respect to HIV-protease inhibitory activity.

Table 4
Evaluation of proposed models using test set of tetrahydropyrimidine-2-ones

Model index	Nature of range in the proposed model	Index value	Number of compounds falling in the range		Percent accuracy	Average K_i (nM)	
			Total	Correct		Total	Correct
W	Inactive	<5450	18	17	94.4	568.2	601.6
	Active	5450–6933	15	12	80	1.85	0.32
	Transitional	>6933	07	N.A.	N.A.	N.A.	N.A.
M_1	Inactive	<212	18	17	94.4	568.2	601.6
	Transitional	212-218	05	N.A.	N.A.	N.A.	N.A.
	Active	220–280	17	13	76.4	26.4	0.28
ξ^c	Inactive	<1029	22	19	86.3	465.4	538.8
	Active	1029–1191	13	11	84.6	1.46	0.33
	Transitional	>1191	05	N.A.	N.A.	N.A.	N.A.

N.A.: not applicable.

- The active range had a *Wiener's index* value of 5450–6933. Twelve out of the 15 analogues in the active range exhibited the HIV-protease inhibitory activity.
- The average K_i value was found to be 0.32 nM for the correctly predicted compounds, indicating the presence of highly active compounds in the active range.

Model based upon Zagreb group parameter:

- A total of 30 out of the 35 test compounds were classified correctly in both the active and inactive ranges. The overall accuracy of prediction was found to be 85.7%, with respect to HIV-protease inhibitory activity.
- The active range had a *Zagreb group parameter* value of 220–280. Thirteen out of the 17 analogues in the active range exhibited the HIV-protease inhibitory activity.
- The average K_i value was found to be 0.28 nM for correctly predicted compounds, indicating the presence of highly active compounds in the active range.

Model based upon eccentric connectivity index:

- A total of 30 out of the 35 compounds were classified correctly in both the active and inactive ranges. The overall accuracy of prediction was found to be 85.7%, with respect to HIV-protease inhibitory activity.
- The active range had an *eccentric connectivity index* value of 1029–1191. Eleven out of the 13 analogues in the active range exhibited HIV-protease inhibitory activity.
- The average K_i value was found to be 0.33 nM for correctly predicted compounds, indicating the presence of highly active compounds in the active range.

4. Conclusion

Investigations reveal significant correlations between the topological indices used in the present study and HIV-protease inhibitory activity of tetrahydropyrimidin-2-ones. The overall degree of prediction was found to be ∼88% in case of *Wiener's index*, ∼86% in case of *Zagreb group parameter*, and ∼86% in case of *eccentric connectivity index*. Prediction with the *Wiener's index* was better when compared to *eccentric connectivity index* and *Zagreb group parameter.*

High degree of predictability of the proposed models offer a vast potential for providing lead structures for the development of potent HIV-protease inhibitors.

References

[1] S. Thaisrivaongs, HIV protease inhibitors, Annu. Rep. Med. Chem. 29 (1994) 133–144.
[2] J.P. Vacca, Design of tight-binding human immunodeficiency virus type 1-protease inhibitors, Meth. Enzymol. 241 (1994) 311–334.
[3] P.L. Darke, J.R. Huff, HIV protease as an inhibitor target for the treatment of AIDS, Adv. Pharmacol. 25 (1995) 399–454.
[4] K.T. Chong, Recent advances in HIV-1 protease inhibitors, Exp. Opin. Invest. Drugs 5 (1996) 115–124.
[5] G.V. De Lucca, S. Erickson-Viitanen, P.Y.S. Lam, Cyclic HIV protease inhibitors capable of displacing the active-site structural water molecule, Drug Discovery Today 2 (1997) 6–18.
[6] L.H. Pearl, W.R. Taylor, A structural model for the retroviral proteases, Nature 329 (1987) 351–354.
[7] R.A. Kramer, M.D. Schaber, A.M. Skalka, K. Ganguly, F. Wong-Staal, E.P. Reddy, HTLV-III *gag* protein is processed in yeast cells by the virus *pol*-protease, Science 231 (1986) 1580–1584.
[8] M.C. Graves, J.J. Lim, E.P. Heimer, R.A. Kramer, An 11-kDa form of human immunodeficiency virus protease expressed in *Escherichia coli* is sufficient for enzymatic activity, Proc. Natl. Acad. Sci. USA 85 (1988) 2449–2453.
[9] S.F.J. Le Grice, J. Mills, J. Mous, Active site mutagenesis of the AIDS virus protease and its alleviation by trans complementation, EMBO J. 7 (1988) 2547–2553.
[10] C. Flexner, S.S. Broyles, P. Earl, S. Chakrabarti, B. Moss, Characterization of human immunodeficiency virus gag/pol gene products expressed by recombinant vaccinia viruses, Virology 166 (1988) 339–349.
[11] V. Karacostas, K. Nagashima, M.A. Gonda, B. Moss, Human immunodeficiency virus-like particles produced by a vaccinia virus expression vector, Proc. Natl. Acad. Sci. USA 86 (1989) 8964–8967.
[12] C. Debouck, The HIV-1 protease as a therapeutic target for AIDS, AIDS Res. Hum. Retroviruses 8 (1992) 153–164.
[13] H.A. Overton, D.J. McMillan, S.J. Gridley, S. Redshaw, J.S. Mills, Effect of two novel inhibitors of the human immunodeficiency virus protease on the maturation of the HIV *gag* and *gag–pol* polyproteins, Virology 179 (1990) 508–511.

[14] S. Raffanti, D.W. Haas, Antimicrobial agents: antiretroviral agents., in: J.G. Hardman, L.E. Limbird (Eds.), The Pharmacological Basis of Therapeutics, vol. 10, McGraw-Hill, USA, 2001, p. 1364.
[15] D.H. Rouvray, The challenge of characterizing branching in molecular species, Discrete Appl. Math. 19 (1988) 317–338.
[16] M. Randic, Graph theoretical analysis of diamagnetic susceptibilities of alkanes and alkenes, Chem. Phys. Lett. 53 (1978) 602–605.
[17] M. Randic, Survey of structural regularities in molecular properties. I. Carbon-13 shifts in alkanes, Int. J. Quant. Chem. 23 (1983) 1707–1722.
[18] M. Randic, On molecular identification numbers, J. Chem. Inf. Comput. Sci. 24 (1984) 164–175.
[19] S.C. Basak, V.R. Magnuson, G.J. Niemi, R.R. Regal, Determining structural similarity of chemicals using graph-theoretic indices, Discrete Appl. Math. 19 (1988) 17–44.
[20] P.A. Philip, T.S. Balaban, A.T. Balaban, A new approach for devising local graph invariants: derived topological indices with low degeneracy and good correlation ability, J. Math. Chem. 1 (1987) 61–83.
[21] C. Hansch, Recent advances in biomedicinal QSAR, in: N.B. Chapman, J. Shorter (Eds.), Correlation Analysis in Chemistry, Plenum Press, New York, 1978, pp. 397–438.
[22] Y.C. Martin, Quantitative Drug Design: A Critical Introduction, Marcel Dekker, New York, 1978.
[23] M. Tichy, QSAR in Toxicology and Xenobiochemistry, Elsevier, Amsterdam, 1985.
[24] A. Crum-Brown, T. Fraser, Trans. Roy. Soc. Edin. 25 (151) (1868/1869) 693–739.
[25] N. Trinajstic, Chemical Graph Theory, CRC Press, Boca Raton, FL, 1983.
[26] A.T. Balaban, J. Motoc, D. Bonchev, O. Mekennyan, Topological indices for structure-activity correlations, Top. Curr. Chem. 114 (1983) 21–55.
[27] L.B. Kier, L.H. Hall, Molecular Connectivity in Structure–Activity Analysis, Research Studies Press, Letchworth, UK, 1986.
[28] A.R. Katritzky, E.V. Gordeeva, Traditional topological indices versus electronic, geometric and combined molecular descriptors in QSAR/QSPR research, J. Chem. Inf. Comput. Sci. 33 (1993) 835–857.
[29] S.C. Basak, S. Bertlsen, G.D. Grunwald, Application of graph theoretical parameters in quantifying molecular similarity and structure–activity relationships, J. Chem. Inf. Comput. Sci. 34 (1994) 270–276.
[30] E. Estrada, A. Ramirez, Edge adjacency relationships and molecular topographic descriptors: definition and applications, J. Chem. Inf. Comput. Sci. 36 (1996) 837–843.
[31] R. Garcia-Domenech, J.V. De Julian-Ortiz, M.J. Duart, J.M. Garcia-Torrecillas, G.M. Anton-Fos, I. Rios-Santamarina, C. De Gregorio-Alapont, J. Galvez, Search of topological pattern to evaluate the toxicity of heterogeneous compounds, SAR QSAR Environ. Res. 12 (2001) 237–254.
[32] B.D. Gute, S.C. Basak, Predicting acute toxicity of benzene derivatives using theoretical molecular descriptors: a hierarchical QSAR approach, SAR QSAR Environ. Res. 7 (1997) 117–131.
[33] H. Hosoya, A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons, Bull. Chem. Soc. Jpn. 44 (1971) 2332–2339.
[34] H. Hosoya, Topological index as strong sorting device for coding chemical structure, J. Chem. Doc. 12 (1972) 181–183.
[35] M. Randic, On characterization of molecular branching, J. Am. Chem. Soc. 97 (1975) 6609–6615.
[36] S. Gupta, M. Singh, A.K. Madan, Application of graph theory: relationship of molecular connectivity index and atomic molecular connectivity index with anti-HSV activity, J. Mol. Struct. (Theochem.) 571 (2001) 147–152.
[37] A.T. Balaban, A. Chiriac, I. Motoc, Z. Simon, Steric fit in QSAR, Lect. Notes Chem. 15 (1980) 22–27.
[38] A.T. Balaban, Highly discriminating distance-based topological indices, Chem. Phys. Lett. 89 (1982) 399–404.
[39] A.T. Balaban, Applications of graph theory in chemistry, J. Chem. Inf. Comput. Sci. 25 (1985) 334–343.
[40] A.T. Balaban, P. Filip, Computer programme for topological index *J* (average distance sum connectivity), J. Math. Chem. 16 (1984) 163–190.
[41] S. Sardana, A.K. Madan, Application of graph theory: relationship of molecular connectivity index. Wiener's index and eccentric connectivity index with diuretic activity, Commun. Math. Chem. (MATCH) 43 (2001) 85–98.
[42] S. Sardana, A.K. Madan, Topological models for prediction of antihypertensive activity of substituted benzylimidazoles, J. Mol. Struct. (Theochem.) 638 (2003) 41–49.
[43] S. Sardana, A.K. Madan, Application of graph theory: relationship of antimycobacterial activity of quinolone derivatives with eccentric connectivity index and Zagreb group parameters, Commun. Math. Chem. (MATCH) 45 (2002) 35–53.
[44] S. Sardana, A.K. Madan, Predicting anti-HIV activity of TIBO derivatives: computational approach using a novel topological descriptor, J. Mol. Model. 8 (2002) 258–265.
[45] S. Gupta, M. Singh, A.K. Madan, Application of graph theory: relationship of eccentric connectivity index and Wiener's index with anti-inflammatory activity, J. Math. Anal. Appl. 266 (2002) 259–268.
[46] S. Sardana, A.K. Madan, Predicting anticonvulsant activity of benzamides/benzylamine: computational approach using topological descriptors, J. Comput. Aid. Mol. Des. 16 (2002) 545–550.
[47] H. Wiener, Correlations of heats of isomerization and differences in heats of vaporization of isomers among paraffin hydrocarbons, J. Am. Chem. Soc. 69 (1947) 2636–2638.
[48] H. Wiener, Influence of interatomic forces on paraffin properties, J. Chem. Phys. 15 (1947) 766.
[49] H. Wiener, Vapour-pressure–temperature relations among the branches paraffin hydrocarbons, J. Chem. Phys. 15 (1948) 425–430.
[50] H. Wiener, Relationship of physical properties of isomeric alkanes to molecular structure, surface tension, specific dispersion and critical solution temperature in aniline, J. Phys. Colloid. Chem. 52 (1948) 1082–1089.
[51] M. Randic, X. Guo, T. Oxely, H. Krishnapriyan, Wiener matrix: source of novel graph invariants, J. Chem. Inf. Comput. Sci. 33 (1993) 709–716.
[52] M. Randic, X. Guo, T. Oxely, H. Krishnapriyan, L. Naylor, Wiener matrix invariants, J. Chem. Inf. Comput. Sci. 34 (1994) 361–367.
[53] I. Gutman, B. Russic, N. Trinajstic, C.X. Wicox Jr., Graph theory and molecular orbitals. XII. Acyclic polyenes, J. Chem. Phys. 62 (1975) 3399–3405.
[54] I. Gutman, N. Trinajstic, Graph theory and molecular orbitals III, Chem. Phys. Lett. 17 (1972) 535–538.
[55] V. Sharma, R. Goswami, A.K. Madan, Eccentric connectivity index: a novel highly discriminating topological descriptor for structure property and structure–activity studies, J. Chem. Inf. Comput. Sci. 37 (1997) 273–282.
[56] G.V. De Lucca, J. Liang, I. De Lucca, Stereospecific synthesis, structure–activity relationship, and oral bioavailability of tetrahydropyrimidin-2-one HIV protease inhibitors, J. Med. Chem. 42 (1999) 135–152.
[57] S. Gupta, M. Singh, A.K. Madan, Predicting anti-HIV activity: computational approach using novel topological descriptor, J. Comput. Aid. Mol. Des. 15 (2001) 671–678.
[58] C. Flexner, HIV-protease inhibitors, N. Eng. J. Med. 338 (1998) 1281–1292.
[59] A.R. Katritzky, A. Oliferenko, A. Lomakab, M. Karelsonb, Six-membered cyclic ureas as HIV-1 protease inhibitors: a QSAR study based on CODESSA PRO approach, Bioorg. Med. Chem. Lett. 12 (2002) 3453–3457.
[60] A.C. Nair, P. Jayatilleke, X. Wang, S. Miertus, W.J. Welsh, Computational studies on tetrahydropyrimidin-2-one HIV-1 protease inhibitors: improving three-dimensional quantitative structure–activity relationship. Comparative molecular field analysis models by inclusion of calculated inhibitor and receptor based properties, J. Med. Chem. 45 (2002) 973–983.