J Mol Model (2004) 10:399–407
DOI 10.1007/s00894-004-0215-8

ORIGINAL PAPER

Vipin Kumar · Satish Sardana · Anil Kumar Madan

Predicting anti-HIV activity of 2,3-diaryl-1,3-thiazolidin-4-ones: computational approach using reformed *eccentric connectivity index*

Received: 25 April 2004 / Accepted: 8 September 2004 / Published online: 3 November 2004

Abstract The *eccentric connectivity index*, which has recently been employed successfully for the development of numerous mathematical models for the prediction of biological activities of diverse nature, has been reformed to overcome its limitations caused by degeneracy and insensitivity towards heteroatoms. The reformed *eccentric connectivity index*, termed the *eccentric connectivity topochemical index*, overcomes the limitations of the *eccentric connectivity index* by exhibiting very low degeneracy and displaying sensitivity to both the presence and relative position of heteroatoms without compromizing the discriminating power of the *eccentric connectivity index*. The relationship of the *eccentric connectivity topochemical index*, *eccentric connectivity index* and *Wiener's index* with regard to the anti-HIV activity of 2, 3-diaryl-1, 3-thiazolidin-4-one derivatives was subsequently investigated. The values of the *eccentric connectivity topochemical index*, the *eccentric connectivity index* and *Wiener's index* of each of 31 analogues comprizing the data set were computed using in-house computer program. Resultant data was analyzed and suitable models developed after identification of active ranges. Subsequently, each derivative was assigned a biological activity using these models, which was then compared with the reported anti-HIV activity. The accuracy of prediction using these models was found to vary from 81 to 90%. The proposed index offers a vast potential for virtual screening of combinatorial libraries, structure property/activity studies and drug design.

Keywords Topological indices · Eccentric connectivity topochemical index · Wiener's index · Eccentric connectivity index · Anti-HIV activity · 2, 3-Diaryl-1 · 3-thiazolidin-4-one

V. Kumar · A. K. Madan (✉)
Faculty of Pharmaceutical Sciences,
M. D. University,
Rohtak, 124 001, India
e-mail: madan_ak@yahoo.com

S. Sardana
Hindu College of Pharmacy,
Sonepat, 131 001, India

Introduction

High-throughput screening and combinatorial chemistry techniques have evolved from a random selection of chemical compounds to the use of various statistical methods to focus the molecular diversity [1]. Instead of testing millions of randomly chosen compounds, medicinal chemists prefer to use computational chemistry methods that offer a smaller number of rationally selected compounds for biological testing. The virtual screening of combinatorial libraries is widely used to design optimally diverse libraries and focused or targeted libraries. Virtual-screening methods can offer the best set of compounds for a combinatorial synthetic scheme to maximize the chances of finding a drug lead, using both efficient statistical techniques and a set of numerical descriptors of the chemical structure [2]. As molecular databases increase in size, information content, and complexity, the application of complicated descriptors becomes impractical for data mining. Therefore, using much simpler and not so demanding descriptors for a first vast diversity screening of a very large dataset could facilitate the entire process. This task can be very well fulfilled by structural descriptors traditionally used in structure–property relationship and structure–activity relationship models (SPR/SAR models) [3–11]

More than 1,000 structural descriptors from five classes are used in SAR/SPR studies: constitutional, graph theoretical and topological indices, geometrical, electrostatic, and quantum descriptors. From these descriptors, the topological indices derived from molecular graphs were found to give good results in database screening. Topological indices have obvious advantages when compared with other descriptors: they are computed only from the information contained in the molecular graph, they have a unique value for a particular chemical compound, they offer a simple way of measuring molecular

branching, shape, size, and molecular similarity, and their calculation requires small computational resources. All the above characteristics indicate that topological indices are well suited for the virtual screening of databases of hundreds of thousands of compounds [2]. Structural descriptors derived from the molecular graph are widely used in modeling physical, chemical, or biological properties, in similarity and diversity assessment, database mining, and in the virtual screening of combinatorial libraries [12]. In a first approximation, the chemical structure of a molecule can be represented as a molecular graph. A topological index is a numerical descriptor of the molecular structure derived from the corresponding molecular graph. 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, whose vertices correspond to non-hydrogen atoms and whose edges correspond to covalent bonds [13]. In the graph representation of molecules, their geometrical features such as bond lengths or bond angles are not taken into account and the chemical bonding of atoms is regarded as being their most important characteristic. Numerous reviews and research papers describe the computation of structural descriptors from molecular graphs, and their use in drug design, toxicology, property prediction and design of chemical libraries [14–27].

Although a number of topological indices have been reported in the literature, only a handful of them have been employed successfully in SAR studies. These include *Hosoya's index* [28, 29] ,*Randic's molecular connectivity index*, χ [30], *the higher*-it order connectivity indices, ${}^{n}\chi$, for the paths of length *n* defined by Kier and Hall [15], *Balaban's index*, *J* [31–34], *Wiener's* index [35–37], *Zagreb group parameters*, M_1 and M_2 [38, 39], and the *eccentric connectivity index* [40–47].

Acquired immunodeficiency syndrome-(AIDS) is caused by the human immunodeficiency virus (HIV), a retrovirus of the lentivirus family. The variant of HIV that is the cause for almost all infections is known as HIV-1. The result of HIV infection is relentless destruction of the immune system. The immunologic profile exhibited in AIDS from early infection to terminal stage illness shows effects of the gradual devastation of the immune system [48].

HIV primarily infects cells with CD4 cell-surface receptor molecules, using them to gain entry. Many cell types share common epitopes with this protein, though CD4 lymphocytes play a crucial role. Cells of the mononuclear phagocyte system, principally blood monocytes and tissue macrophages, T lymphocytes, B lymphocytes, natural killer (NK) lymphocytes, dendritic cells (Langerhans cells of epithelia and follicular dendritic cells in lymph nodes), hematopoietic stem cells, endothelial cells, microglial cells in brain, and gastrointestinal epithelial cells are the primary targets of HIV infection.

Within the cell, the viral particle uncoats from the envelope to releases its RNA. Viral RNA serves as a template from which complementary DNA strands are transcribed. This reverse transcription is the distinguishing feature of retroviruses and is catalyzed by the HIV RNA-dependent DNA polymerase i.e. *reverse transcriptase*. After initial entry of HIV and establishment of infection, replication may at first occur within inflammatory cells at the site of infection or within peripheral blood mononuclear cells, but then the major site of replication quickly shifts to lymphoid tissues of the body, including those in lymph nodes, spleen, liver, and bone marrow. Besides lymph nodes, the gut-associated lymphoid tissue provides a substantial reservoir for HIV. The primary target of HIV is the immune system itself, which is gradually destroyed. Viral replication actively continues following initial HIV infection, and the rate of CD4 lymphocyte destruction is progressive.

Reverse transcriptase is necessary for the catalytic transformation of single stranded viral RNA into the double stranded linear DNA, which is integrated into host cell chromosomes. Reverse transcriptase inhibitors bind in a noncompetitive manner to a unique site on the enzyme, altering its ability to function and achieving a highly selective suppression of HIV-1 replication with little cytotoxicity [49]. The clinical benefit of antiretroviral therapy is determined by the magnitude and duration of plasma HIV RNA suppression [50, 51].

Retroviruses such as HIV-1 possess a unique replication cycle, so a variety of molecular targets are now available for therapeutic intervention [52]. One such target of considerable interest is the virally encoded RT, which mediates conversion of the viral RNA to proviral DNA.

Therapeutic intervention for the treatment of HIV-1, which eventually results in AIDS, is limited to AZT (Zidovudine) and a few other drugs approved for this disease. Unfortunately, these suffer from a number of limitations including limiting side effects and the revelation of the possible emergence of drug-resistant mutants of the virus [53]. Use of nucleoside analogues such as AZT [54, 55], which inhibit the operation of *reverse transcriptase*, is limited by significant toxicities [56–58]. Potent nonnucleoside, HIV-1 specific *reverse transcriptase* inhibitors have now been described in several chemically diverse series [59, 60].

The emergence of multidrug-resistance virus among patients treated with current anti-AIDS agents has created an urgent need for new antiretroviral agents. It is hoped that at least some agents directed at this target will control replication of drug-resistant HIV, have fewer side effects, and/or require less frequent dosing and fewer capsules or tablets than currently available agents. In view of the increasing incidence of resistance to current drug regimens and the frequency of adverse events, the development of novel, potent, safe, inexpensive antiviral agents

that are also effective against mutant HIV strains remains a high priority for medical research.

In the present investigations, the *eccentric connectivity index*— an adjacency-cum-distance based topological descriptor has been reformed to make it sensitive to both the presence and relative position of heteroatoms. The reformed *eccentric connectivity index*, termed the *eccentric connectivity topochemical index*, has been conceptualized and its discriminating power investigated with regard to the anti-HIV activity of 2, 3-diaryl-1, 3-thiazolidin-4-one derivatives. The degeneracy and discriminating power of the *eccentric connectivity topochemical index* was compared with that of the *eccentric connectivity index*—an adjacency-cum-distance based topological descriptor and *Wiener's index*—a distance-based topological descriptor for all possible structures up to five vertices containing one nitrogen atom.

Methodology

Calculations of topological descriptors

The *Wiener's* index, [35–37] a well-known distance-based topological index is defined as the sum of the distances between all the pairs of vertices in a hydrogen-suppressed molecular graph, that is

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

where P_i 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 *eccentric connectivity* index [40–47] 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, that is

$$\xi^c = \sum_{i=1}^{n} (E_i \times V_i) \tag{2}$$

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

The *eccentric connectivity index* has been modified in the present study so as to make it not only heteroatom sensitive but also to minimize degeneracy. Since the *eccentric connectivity index* has been used successfully recently for the development of numerous mathematical models for the prediction of biological activities of diverse nature, therefore, it was naturally expected that modification of the said descriptor would most probably result in vast improvement in its utility in modeling because of its sensitivity to both the presence and the relative position of heteroatoms and minimization of degeneracy. Accordingly, the *eccentric connectivity topo-chemical index* has been conceptualized to consider nature, number and position of heteroatoms.

The *eccentric connectivity topochemical index*, denoted by $\xi_c{}^c$, is defined as the summation of the product of chemical eccentricity and the chemical degree of each vertex in the hydrogen-suppressed molecular graph having n vertices, that is

$$\xi_c^c = \sum_{i=1}^{n} (E_{ic} \times V_{ic}) \tag{3}$$

where V_{ic} is the chemical degree of vertex i, E_{ic} is the chemical eccentricity of the vertex i and n is the number of the vertices in graph G.

The values of *Wiener's index* and the *eccentric connectivity index* can be calculated from the distance matrix (D) and the adjacency matrix (A). On the other hand, the value of proposed *eccentric connectivity topochemical index* could be easily calculated from the topochemical distance matrix (D_c) and the topochemical adjacency matrix (A_c). The chemical degree and chemical eccentricity of a vertex were obtained from the topochemical adjacency matrix (A_c) and the topochemical distance matrix (D_c), respectively, by substituting row elements corresponding to heteroatoms with the relative atomic weight with respect to carbon e.g. in the case of nitrogen 14/12 i.e. 1.167. The derivation of matrices for the calculation of *Wiener's index*, the *eccentric connectivity index* and the *eccentric connectivity topochemical index* for three isomers of butylamine is illustrated in Fig. 1. The sensitivity of the proposed *eccentric connectivity topochemical index* towards the relative position of a heteroatom is shown in Figs. 2 and 3 for structures having three to five vertices containing one nitrogen as heteroatom.

Model development

A data set comprising 31 derivatives [61] of 2, 3-diaryl-1, 3-thiazolidin-4-one was selected for the present study. The basic structure for these analogues is shown in Fig. 4 and various substituents listed in Table 1. The values of the *eccentric connectivity topochemical index* of all the derivatives in the data set were computed using an in-house computer program. The resultant data was analyzed and a suitable model developed after identification of the active range by the maximization of moving average with respect to active compounds [62]. Subsequently, each analogue was assigned a biological activity using this model, which was then compared with the reported anti-HIV activity.

Clercq et al. [63–66] reported the anti-HIV activity quantitatively as the concentration (EC_{50}) required to reduce the HIV-1-induced cytopathic effect by 50 percent in MT-4 cells. Derivatives exhibiting EC_{50} less than 0.1 μM were considered to be active for the purpose of the present study. The cytotoxic activity reported for 31 derivatives was defined as the concentration required to reduce MT-4 cell viability by 50 (CC_{50}). A selectivity index is defined as the quotient of CC_{50} and EC_{50}. Average EC_{50} values, from the reported data [61] for all the analogues as well as correctly predicted analogues present in each range in the proposed model were also calculated. Similarly, average selectivity index values for each range in the proposed model were also calculated.

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 predicted to that of the total number of compounds present in both the active and inactive ranges.

The aforementioned procedure was followed similarly for *Wiener's index* and the *eccentric connectivity index*. The results are summarized in Tables 1 and 2.

Fig. 1 Derivation of matrices for calculation of *eccentric connectivity topochemical index.*

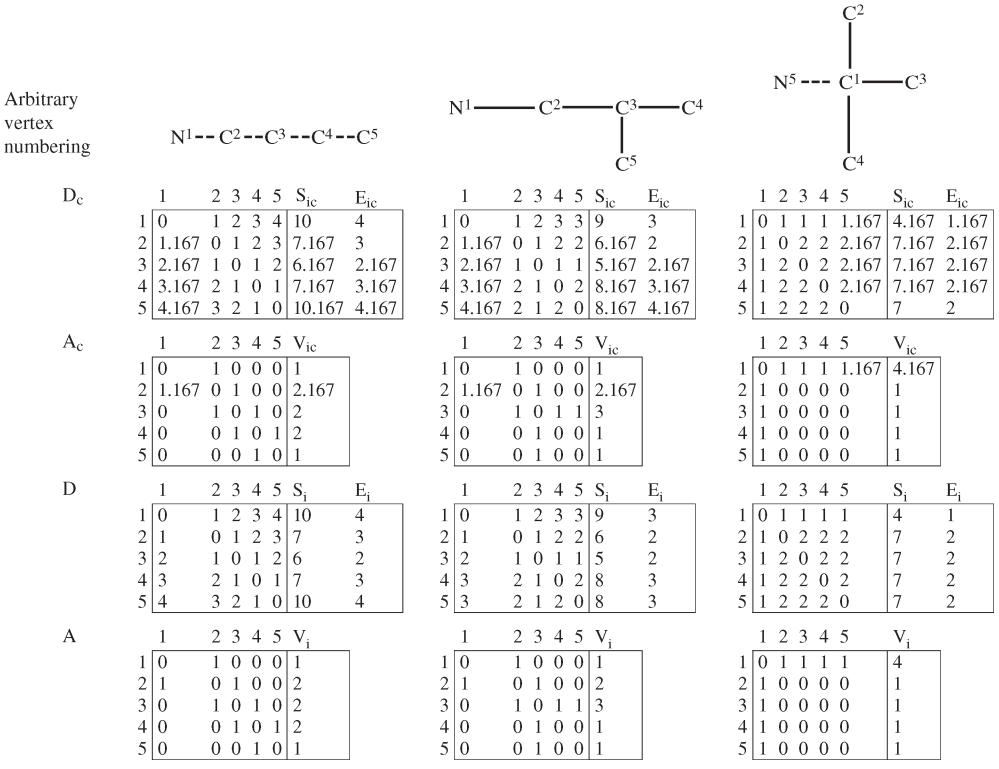
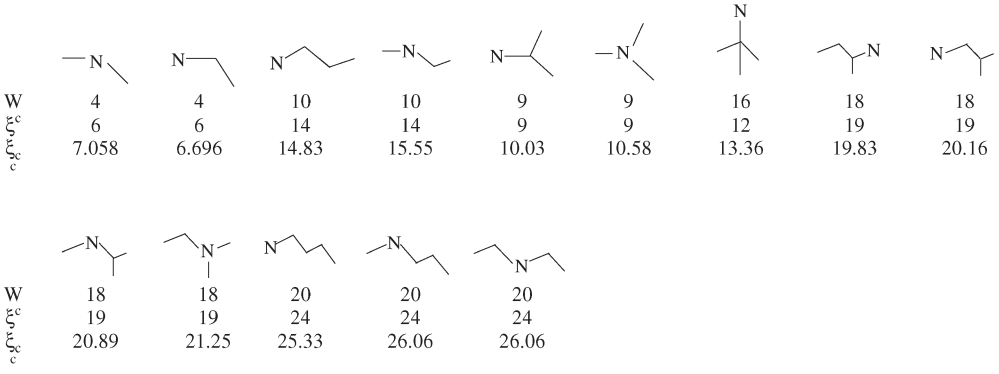


Fig. 2 W, ξ^c, $\xi_c{}^c$ values for structures having three to five vertices containing single nitrogen as heteroatom.



Result and discussion

Acquired immunodeficiency syndrome—first appeared in the west in the late seventies and needs no introduction. AIDS is a syndrome rather than a disease. It is a horrible way to die and should not be regarded lightly. To die of AIDS is to die ravaged, emaciated, generally in a degree of dementia and with a host of dreadful and disgusting symptoms. Human immunodeficiency virus 1 (HIV-1) infection is characterized by a progressive decline in both function and number of CD4+ T-lymphocytes secondary to ongoing viral replication. Without intervention, this ultimately leads to the development of the *acquired immunodeficiency syndrome* that places persons at risk for the acquisition of opportunistic infections and neoplasms [67]. The *acquired immunodeficiency syndrome* epidemic is one of the greatest challenges facing the medical community today. Infection with HIV is a dynamic process characterized by vigorous viral replication, CD4 lymphocyte depletion, and profound immunodeficiency. The error-prone nature of HIV *reverse transcriptase* promotes rapid evolution of genetic diversity and a remarkable propensity to develop resistance to antiretroviral agents [68].

Reverse transcriptase is a key enzyme, which plays an essential and multifunctional role in the replication of HIV [69] and thus represents an attractive target for the development of new drugs useful in AIDS therapy [70].

The discovery of drug leads is significantly accelerated by in silico screening of molecular libraries that starts from a collection of chemical compounds with a high structural diversity and selects molecules according to their similarity toward specific collections of active compounds. In this process, the molecular similarity/diversity and the drug-like character are characterized with structural descriptors, such as structure keys, fingerprints,

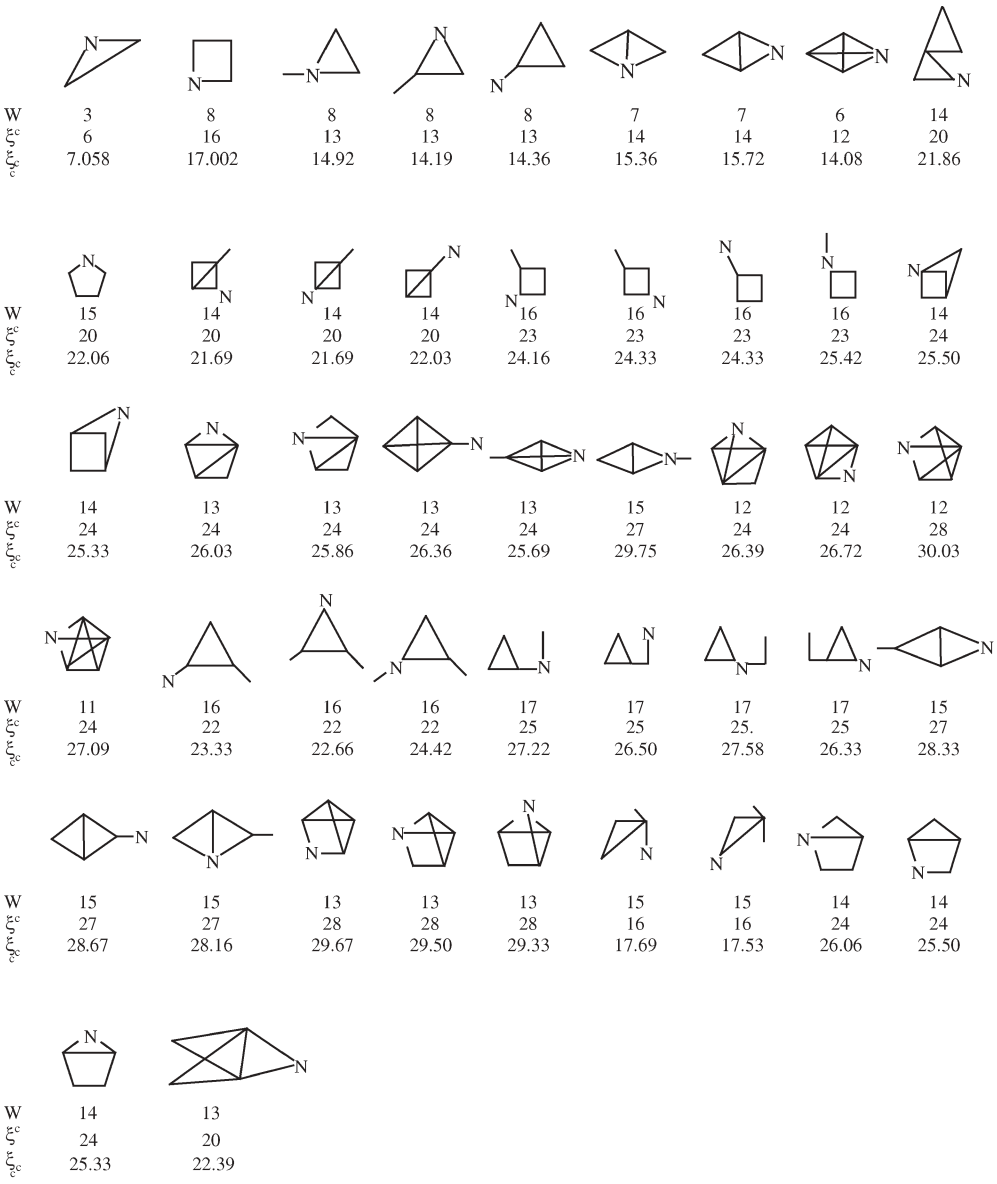


Fig. 3 W, ξ^c, ${\xi_c}^c$ values for cyclic structures having three to five vertices containing single nitrogen as heteroatom.

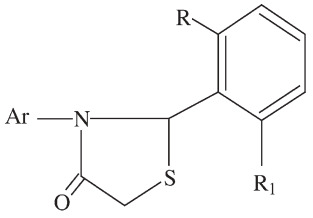


Fig. 4 Basic structure of 2,3-diaryl-1, 3-thiazoidin-4-ones

graph invariants and various topological indices computed from atomic connectivity or molecular matrices. The topological description of a molecule contains information on the atom–atom connectivity in the molecule, and encodes the size, shape, and branching features that determine the molecular properties. Mathematical descriptors of molecular structure, such as various topological indices [71], have been used in structure-activity/property studies, including multiple regression analysis, principal component analysis, pattern recognition, artificial neural networks, optimization of lead compound, search of combinatorial libraries and combinatorial optimization of the lead compound, and similarity–dissimilarity studies [72].

Though a large number of topological descriptors have been reported in the literature for structure-property–activity studies, only a handful of them have been used successfully in structure-activity/property studies. Most of these descriptors are incapable of taking heteroatoms in to considerations. As a consequence, there is a strong need to develop chemical topological descriptors with high discriminating power particularly with respect to structure-activity/property studies.

Table 1 Relationship of *Wiener's index*, *eccentric connectivity index* and *eccentric connectivity topochemical index* with anti-HIV activity

Comp. No.	Ar	R	R_1	W	ξ^c	$\xi_c{}^c$	Predicted activity			Reported activity
							W	ξ^c	$\xi_c{}^c$	
1	Phenyl	Cl	Cl	744	301	383.40	–	–	–	–
2	Phenyl	F	F	744	301	340.54	–	–	–	–
3	Pyridin-2-yl	Cl	Cl	744	301	389.83	–	–	–	–
4	Pyridin-2-yl	Cl	F	744	301	379.98	–	–	–	–
5	Pyridin-2-yl	F	F	744	301	346.49	–	–	–	–
6	Pyridin-3-yl	Cl	Cl	744	301	389.81	–	–	–	–
7	Pyridin-3-yl	F	F	744	301	346.46	–	–	–	–
8	Pyridin-4-yl	Cl	Cl	744	301	394.44	–	–	–	–
9	Pyridin-4-yl	F	F	744	301	350.64	–	–	–	–
10	5-Cl-pyridin-2-yl	Cl	Cl	866	344	524.16	–	–	–	–
11	5-Cl-pyridin-2-yl	Cl	F	866	344	510.24	–	–	–	–
12	5-Cl-pyridin-2-yl	F	F	866	344	479.59	–	–	–	–
13	5-Br-pyridin-2-yl	Cl	Cl	866	344	716.70	–	–	–	–
14	5-Br-pyridin-2-yl	Cl	F	866	344	697.68	–	–	–	–
15	5-Br-pyridin-2-yl	F	F	866	344	665.40	–	–	–	–
16	6-Br-pyridin-2-yl	Cl	Cl	852	318	676.03	+	+	–	–
17	6-Br-pyridin-2-yl	Cl	F	852	318	658.16	+	+	–	+
18	6-Br-pyridin-2-yl	F	F	852	318	627.11	+	+	–	+
19	3-Me-pyridin-2-yl	Cl	Cl	838	316	404.08	–	–	+	–
20	3-Me-pyridin-2-yl	F	F	838	316	361.83	–	–	–	–
21	4-Me-pyridin-2-yl	Cl	Cl	852	318	393.20	+	+	–	–
22	4-Me-pyridin-2-yl	Cl	F	852	318	399.23	+	+	+	+
23	4-Me-pyridin-2-yl	F	F	852	318	363.83	+	+	–	–
24	5-Me-pyridin-2-yl	Cl	Cl	866	344	438.83	–	–	–	–
25	5-Me-pyridin-2-yl	F	F	866	344	390.83	–	–	–	–
26	6-Me-pyridin-2-yl	Cl	Cl	852	318	414.05	+	+	+	+
27	6-Me-pyridin-2-yl	Cl	F	852	318	403.97	+	+	+	+
28	6-Me-pyridin-2-yl	F	F	852	318	368.33	+	+	–	+
29	4,6-di-Me-pyridin-2-yl	Cl	Cl	964	335	433.30	–	+	+	+
30	4,6-di-Me-pyridin-2-yl	Cl	F	964	335	423.22	–	+	+	+
31	4,6-di-Me-pyridin-2-yl	F	F	964	335	372.06	–	+	–	+

+ Active compound, – Inactive compound
NA not applicable

Degeneracy of descriptors

In the present work, the *eccentric connectivity index*, which has recently been employed widely and successfully for the development of numerous mathematical models [40–47] for the prediction of biological activities of diverse nature has been reformed to overcome its limitations caused by degeneracy and its insensitivity to heteroatoms. The reformed *eccentric connectivity index*, termed as the *eccentric connectivity topochemical index*, has been conceptualized. The value of *eccentric connectivity topochemical index* could be calculated from topochemical distance matrix (D_c) and topochemical adjacency matrix (A_c). The chemical degree and chemical eccentricity of a vertex were obtained from the topochemical adjacency matrix (A_c) and the topochemical distance matrix (D_c), respectively by substituting row elements corresponding to heteroatoms with relative atomic weight with respect to carbon atom. The derivation of matrices for calculating *Wiener's index*, the *eccentric connectivity index* and the *eccentric connectivity topochemical index* for three isomers of butylamine are shown in Fig. 1. The sensitivity of the proposed *eccentric connectivity topochemical index* towards the relative position of a heteroatom is shown in Figs. 2 and 3 for structures having three to five vertices containing one nitrogen as a heteroatom.

The results from Figs. 2, 3 and Table 3 clearly reveal the sensitivity of the *eccentric connectivity topochemical index* to both the presence and the relative position of a heteroatom as well as a steep decrease in degeneracy when compared to the *eccentric connectivity index* without adversely affecting discriminating power if one just compares the values for three, four and five vertices containing one nitrogen as heteroatom.

Anti-HIV activity

The relationship of the *eccentric connectivity topochemical index*, the *eccentric connectivity index* and *Wiener's index* to the anti-HIV activity of 2, 3-diaryl-1, 3-thiazolidin-4-one derivatives has been investigated. The selected data set comprises 31 analogues included both active and inactive compounds.

Retrofit analysis of the data in Tables 1 and 2 reveals the following information with regard to the model based on *Wiener's index*:

Table 2 The relationship between the anti-HIV activity and topological indices

Model index	Nature of range in proposed model	Index value	Number of analogues in the range	Number of analogues predicted correctly	Percent accuracy	Average selectivity index	Average EC_{50} (μM)
Wiener's index	Lower inactive	≤838	11	11	100	105.25 (105.25)	13.87 (13.87)
	Active	>838–<866	09	06	66	1580.44 (1750)	0.11 (0.06)
	Upper inactive	≥866	11	08	73	233.65 (17.15)	2.31 (3.15)
Eccentric connectivity index	Lower inactive	<318	11	11	100	105.25 (105.25)	13.87 (13.87)
	Active	318–335	12	09	75	1388.08 (1437)	0.10 (0.06)
	Upper inactive	> 335	08	08	100	17.15 (17.15)	3.15 (3.15)
Eccentric connectivity topochemical index	Lower inactive	<399.23	15	13	87	462.94 (359.46)	10.58 (12.16)
	Active	399.23–433.30	06	05	83	1518.66(1822.4)	0.05 (0.06)
	Upper inactive	>433.30	10	08	80	189.59 (43.23)	2.04 (2.54)

Values in the brackets are based upon correctly predicted compounds in the particular range

Table 3 Comparison of discriminating power and degeneracy* of *Wiener's Index* (**W**), *eccentric connectivity index* (ξ^c) and *eccentric connectivity topochemical index* ($\xi_c{}^c$) for all possible structures up to five vertices

	W	ξ^c	$\xi_c{}^c$
For three vertices			
Minimum value	3	6	6.696
Maximum value	4	6	7.058
Ratio	1:1.333	1:1	1:1.054
Degeneracy	1/3	3/3	1/3
For four vertices			
Minimum value	6	9	10.03
Maximum value	10	16	17.00
Ratio	1:1.167	1:1.778	1:1.69
Degeneracy	6/11	5/11	0/11
For five vertices			
Minimum value	11	12	13.36
Maximum value	20	28	30.03
Ratio	1:1.819	1:2.333	1:2.24
Degeneracy	38/47	36/47	8/47

Degeneracy number of compounds having same value/Total number of compounds with same number of vertices

Biological activity was assigned to a total of 31 analogues in both the active and inactive ranges, out of which the activity of 25 analogues was predicted correctly, resulting in ~81% accuracy with regard to anti-HIV activity.

- The active range had *Wiener's index* values of >838 to <866. 6 out of nine analogues in the active range exhibited anti-HIV activity resulting in 66% accuracy with regard to the active range of anti-HIV activity.
- The average selectivity index and average EC_{50} values of correctly predicted derivatives were observed to be 1,750 and 0.062 μM, respectively. The EC_{50} value in the active range is not only lowest but the average value of selectivity index is also the highest.
- Two inactive ranges, the lower and upper inactive ranges, ideally bracketed the active range. The lower inactive range had *Wiener's index* values of ≤838 and the upper inactive range had *Wiener's index* values of ≥866. All 11 analogues in the lower inactive range were correctly predicted resulting in 100% accuracy with regard to the lower inactive range of anti-HIV activity.

Retrofit analysis of data in Tables 1 and 2 reveals the following information with regard to the model based on the *eccentric connectivity index*

- Biological activity was assigned to a total of 31 analogues in both the active and inactive ranges, out of which the activity of 28 analogues was predicted correctly resulting in ~90% accuracy with regard to anti-HIV activity.
- The active range had *eccentric connectivity index* values of 318–335. Nine out of 12 analogues in the active range exhibited anti-HIV activity resulting in ~75% accuracy with regard to active range of anti-HIV activity.

– The average selectivity index and average EC_{50} values of correctly predicted derivatives were observed to be 1,437 and 0.062 μM, respectively. The EC_{50} value in the active range is not only lowest but the average value of selectivity index is also the highest.
– The lower and upper inactive ranges bracketed the active range. The lower inactive range had *eccentric connectivity index* values of less than 318 and the upper inactive range had *eccentric connectivity index* values of more than 335. Nineteen out of 19 analogues in both the inactive ranges were predicted correctly, resulting in 100% accuracy with regard to inactive range of anti-HIV activity.

Retrofit analysis of data in Tables 1 and 2 reveals the following information with regard to the model based on the *eccentric connectivity topochemical index*:

– Biological activity was assigned to a total of 31 analogues in both the active and inactive ranges, out of which activity of 26 analogues was predicted correctly, resulting in ~84% accuracy with regard to anti-HIV activity.
– The active range had *eccentric connectivity topochemical index* values of 399.23 to 433.30. Five out of six analogues in the active range exhibited anti-HIV activity, resulting in ~83% accuracy with regard to active range of anti-HIV activity.
– The average selectivity index and average EC_{50} values of correctly predicted derivatives were observed to be 1,822 and 0.065 μM, respectively. The EC_{50} value in the active range is not only lowest but the average value of selectivity index is also the highest.
– Two inactive ranges, the lower and upper inactive ranges, ideally bracketed the active range.

Although the EC_{50} value in the active range was lowest, surprisingly, the average value of selectivity index was also the highest in all the aforementioned models. Since the selectivity index is an indirect measure of therapeutic index, high potency amalgamated with low toxicity renders the active range in the proposed models highly significant for providing lead structures for the development of potent therapeutic agents of extremely low toxicity.

High discriminating power, low degeneracy and sensitivity towards both the presence and relative position of heteroatoms offers the *eccentric connectivity topochemical index* a vast potential for structure-property/activity studies, optimization of lead compounds, search of combinatorial libraries and combinatorial optimization of the lead compounds, and similarity–dissimilarity studies.

References

1. Wart WA (1998) Combinatorial chemistry. In: Schleyer PvR, Allinger NL, Clark T, Gasteiger J, Kollman PA, Schaefer III HF, Schreiner PR (eds) The Encyclopedia of combinatorial chemistry. Wiley, Chichester, pp 401–407
2. Ivanciuc O, Taraviras SL, Cabrol-Bass D (2000) J Chem Inf Comput Sci 40:126–134
3. Martin EJ, Blaney JM, Siani MA, Spellmeyer DC, Wong AK, Moos WH (1995) J Med Chem 38:1431–1436
4. Martin YC, Bures MG, Brown RD (1998) Pharm Pharmacol Commun 4:147–152
5. Grassy G, Calas B, Yasri A, Lahana R, Woo J, Tyer S (1998) Nat Biotechnol 16:748–752
6. Martin EJ, Critchlow RE (1999) J Comb Chem 1:32–45
7. Ghose AK, Vishwanadhan VN, Wendoloski JJ (1999) J Comb Chem 1:55–68
8. Bayada DM, Hamersma H, van Geerestein VJ (1999) J Chem Inf Comput Sci 39:1–10
9. Pearlman RS, Smith KM (1998) Perspect Drug Discov Des 9/10/11:330–353
10. Stanton DT (1999) J Chem Inf Comput Sci 39:11–20
11. Pearlman RS, Smith KM (1999) J Chem Inf Comput Sci 39:28–35
12. Ivanciuc O, Ivanciuc T, Cabrol-Boss D, Balaban AT (2002) Internet Electron J Mol Des 1:319–331
13. Ivanciuc O, Ivanciuc T, Balaban AT (1998) J Chem Inf Comput Sci 38:395–401
14. Kier LB, Hall LH (1976) Molecular connectivity in chemistry and drug research. Academic, New York
15. Kier LB, Hall LH (1986) Molecular connectivity in structure-Activity analysis. Research Studies Press, Letchworth
16. Trinajstic N (1992) Chemical graph theory. CRC Press, Boca Raton
17. Balaban AT (1997) From chemical topology to three-dimensional geometry. Plenum, New York
18. Diudea MV (2001) QSPR/QSAR studies by molecular descriptrors. Nova Science, Huntington
19. Ivanciuc O, Balaban AT (1998) Graph theory in chemistry. In: Schleyer PvR, Allinger NL, Clark T, Gasteiger J, Kollman PA, Schaefer III HF, Schreiner PR (eds) The Encyclopedia of computational chemistry. Wiley, Chichester, pp 1169–1190
20. Ivanciuc O, Balaban AT (1999) The graph description of chemical structure. In: Devillers J, Balaban AT (eds) Topological indices and related descriptors in QSAR and QSPR. Gordon and Breach Science Publishers, Amsterdam, pp 59–167
21. Ivanciuc O, Ivanciuc T, Balaban AT (1999) Vertex- and Edges-Weighted Molecular Graphs and Derived Structural Descriptors. In: Devillers J, Balaban AT (eds) Topological Indices and related Descriptors in QSAR and QSPR. Gordon and Breach Science Publishers, Amsterdam, pp 169–220
22. Ivanciuc O, Ivanciuc T (1999) Matrices and structural descriptors computed from molecular graph distances. In: Devillers J, Balaban AT (eds) Topological indices and related descriptors in QSAR and QSPR. Gordon and Breach Science Publishers, Amsterdam, pp 221–277
23. Ivanciuc O (2002) Internet Electron J Mol Des 1:1–9
24. Rios-Santamarina I, Garcia-Domench R, Cortijo J, Santamaria P, Morcillo J, Galvez J (2002) Internet Electron J Mol Des 1:70–79
25. Ivanciuc O (2000) J Chem Inf Comput Sci 40:1412–1422
26. Marino DGJ, Peruzzo PJ, Castro EA, Toropov AA (2002) Internet Electron J Mol Des 1:115–133
27. Toropov AA, Toropov AP (2002) Internet Electron J Mol Des 1:108–114
28. Hosoya H (1971) Bull Chem Soc Jpn 44:2332–2337
29. Hosoya H (1972) J Chem Doc 12:181–183
30. Randic M (1974) J Am Chem Soc 97:6609–6615
31. Balaban AT, Chiriac A, Motoc I, Simon Z (1980) Lect Notes Chem 15:22–27
32. Balaban AT (1982) Chem Phys Lett 89:399–404
33. Balaban AT (1985) J Chem Inf Comput Sci 25:334–343
34. Balaban AT, Quintas LB (1983) J Math Chem 14:163–233
35. Wiener H (1974) J Chem Phys 15:766–766
36. Wiener H (1947) J Am Chem Soc 69:2636–2638

37. Randic M, Guo X, Oxely T, Krishnapriyan H (1993) J Chem Inf Comput Sci 33:709–716
38. Gutman I, Randic M (1977) Chem Phys Lett 47:15–19
39. Gutman I, Ruscic B, Trinajstic N, Wilcox CF (1975) J Chem Phys 62:3399–3405
40. Sharma V, Goswami R, Madan AK (1997) J Chem Inf Comput Sci 37:273–282
41. Gupta S, Singh M, Madan AK (2002) J Math Anal Appl 266:259–268
42. Sardana S, Madan AK (2002) MATCH Commun. Math Comput Chem 45:36–53
43. Sardana S, Madan AK (2002) J Comput Aided Mol Des 16:1–6
44. Sardana S, Madan AK (2001) MATCH Commun. Math Comput Chem 43:85–98
45. Sardana S, Madan AK (2002) J Mol Model 8:258–265
46. Sardana S, Madan AK (2003) J Mol Str (THEOCHEM) 638:41–49
47. Kumar V, Madan AK (2004) MATCH Commun. Math Comput Chem 51:59–78
48. Wong S (1997) Agents affecting immune response. In: Wolff ME (ed) Burger's medicinal chemistry and drug discovery, vol 4, Therapeutic agents, 5th edn. Wiley, New York, p 605
49. Hajos G, Riedi S, Molnar J, Szabo D (2000) Drugs Fut 25:47–62
50. Marschner IC, Collier AC, Coombs RW, D'Aquila RT, DeGrutta V, Fischi MA, Hammer SM, Hughes MD, Johnsons VA, Katzenstein DA, Richman DD, Smeaton LM, Spector SA, Saag MS (1998) J Infect Dis 177:40–47
51. O'Brian WA, Hartigan PM, Martin D, Esinhart J, Hill A, Benoit S, Rubin M, Simberkoff MS, Hamilton JD (1996) New Engl J Med 40:196–202
52. Mitsuya H, Yarchoan R, Boredr S (1990) Science 249:1533–1544
53. Tomasselli AG, Howe WJ, Sawyer TK, Wlodawer A, Heinrikson RL (1991) Chimica Oggi 6–27
54. Moore RD, Hidalgo J, Sugland BW, Chaisson RE (1991) New Engl J Med 324:1412–1416
55. Lambert JS, Soidlil M, Reichman RC, Plank CS, Laverty M, Morse GD, Knoop C, Milaren C, Petinelli C, Valentine FT, Dolin R (1990) New Engl J Med 322:1333–1345
56. Richman DD, Fischl MA, Grico MH, Gottlieb MS, Volberding PA, Laskin OL, Leedom JM, Groopman JE, Mildvan D, Hirsch MS, Jackson GG, Durack DT, Nusinhoff-Lehrman S (1987) New Engl J Med 317:192–197
57. Larder BA, Darby G, Richman DD (1989) Science 243:1731–1734
58. Larder BA, Kemp SD (1989) Science 246:1155–1158
59. Pauwels R, Andries K, Desmyter J, Schols D, Kukla MJ, Breslin HJ, Raeymaeckers A, van Gelder JV, Heykants J, Schellenekens K, Janssen PAJ (1990) Nature 343:470–474
60. Merluzzi VJ, Hargrave KD, Labadia M, Grozinger E, Skoog M, Wu JC, Shih CK, Eckner K, Hattox X, Rosenthal AS, Faanes R, Eckner K, Koup RA (1990) Science 250:1411–1413
61. Barreca ML, Balzarani J, Chimirri A, de Clercq E, de Luca L, Holtje HD, Holtje M, Monforte AM, Monforte P, Pannecouque C, Rao A, Zappala M (2002) J Med Chem 45:5410–5413
62. Gupta S, Singh M, Madan AK (2001) J Comput Aided Mol Des 15:671–678
63. Tanaka H, Baba M, Haykawa H, Sakamaki T, Miyasaka T, Ubasawa M, Takashima H, Sekiya K, Nitta I, Shigeta S, Walker RT, Balzarani J, de Clercq E (1991) J Med Chem 34:349–357
64. Tanaka H, Baba M, Saito S, Miyasaka T, Ubasawa M, Takashima H, Sekiya K, Nitta I, Shigeta S, Takashima H, Walker RT, de Clercq E (1991) J Med Chem 34:1508–1511
65. Tanaka H, Baba M, Miyasaka T, Ubasawa M, Takashima H, Sekiya K, Nitta I, Shigeta S, Walker RT, de Clercq E (1992) J Med Chem 35:4713–4719
66. Tanaka H, Baba M, Miyasaka T, Ubasawa M, Takashima H, Sekiya K, Nitta I, Shigeta S, Walker RT, de Clercq E (1992) J Med Chem 35:337–345
67. Miedema F, Petit AJ, Terpstra FG, Schattenkerk JK, de Wolf F, Al BJ (1988) J Clin Invest 82:1908–1914
68. Raffanti SP, Haas DW (2001) Antiretroviral agents. In: Hardman JG, Limbird LE (eds) Goodman& Gilman's, The pharmacological basis of therapeutics, 10th edn. McGraw-Hill (Medical Publishing Division), pp 1349–1380
69. Jonckheere H, Anne J, de Clercq E (2000) Med Res Rev 20:129–154
70. de Clercq E (2000) Rev Med Virol 10:255–277
71. Randic M (1998) Topological indices. In: Schleyer PvR (ed) Encyclopedia of computational chemistry. Wiley, London, pp 3018–3032
72. Randic M (2000) Acta Chim Slov 47:143–151