Journal of Computer-Aided Molecular Design, **15:** 671–678, 2001.
KLUWER/ESCOM

Predicting anti-HIV activity: computational approach using a novel topological descriptor

S. Gupta[1], M. Singh[1] & A.K. Madan[2,*]
[1]*Department of Pharmaceutical Sciences and Drug Research, Punjabi University, Patiala 147 002, India;* [2]*Faculty of Pharmaceutical Sciences, M.D. University, Rohtak 124001, India*

Received 25 June 2000; accepted 29 May 2001

Key words: additive adjacency matrices, anti-HIV activity, eccentric adjacency index, first order molecular connectivity index, 1-[(2-hydroxyethoxy)methyl]-6-(phenylthio)thymine, HEPT

Summary

The discriminating power of a novel topological descriptor termed as *eccentric adjacency index* in the estimation of anti-HIV activity, for a data set of 107 1-[(2-hydroxyethoxy)methyl]-6-(phenylthio)thymine (HEPT) derivatives was investigated in the present study. The value of *eccentric adjacency index* of each derivative was computed and active range was identified using moving average analysis. Subsequently, each derivative was assigned a biological activity which was then compared with the reported anti-HIV activity. The accuracy of prediction was found to be more than ninety percent in the active range using *eccentric adjacency index.* The proposed index offers a vast potential for structure-activity/property studies.

Introduction

Acquired immuno deficiency syndrome (AIDS) is an immunosuppressive disease caused by the depletion of helper T-lymphocytes. The causative agent, termed as HIV-1, is a retrovirus. A similar retrovirus, HIV type 2 also causes AIDS. Various compounds have been reported by De Clercq [1] to inhibit the replication of HIV-1 *in vitro*. A thymidine derivative 3′-azido-3′-deoxy-thymidine (AZT) is clinically effective in the treatment of AIDS but associated with side effects like bone marrow suppression besides emergence of AZT-resistant HIV variants. A purine dideoxy nucleoside, 2′, 3′-dideoxyinosine (DDI) is used as an alternate drug for the patients who do not tolerate AZT, but it also has unfavorable side effects. The reverse transcriptase of HIV-1 is an essential enzyme required to catalyze the conversion of viral RNA into proviral DNA and therefore is a target for antiviral therapy against AIDS. AZT and DDI act as inhibitors of viral reverse transcriptase after phosphorylation by cellular kinases. These phosphates may also interact nonspecifically with host cellular DNA polymerases and account for toxic side effects. In the search of more selective and effective agents against HIV a large number of nucleoside derivatives have been evaluated for their antiviral activities [2-6]. 1-[(2-hydroxyethoxy) methyl]-6-(phenylthio) thymine (HEPT) derivatives have been reported by Miyasaka et al. [7] to be selective inhibitor of reverse transcriptase of human immuno deficiency virus type 1 (HIV-1). They have no inhibitory activity against other retroviruses related to HIV like SIV_{MAC} (simian immuno deficiency virus), SRV (simian AIDS-related virus), MSV (murine Moloney sarcoma virus), and HIV-2. Therefore, these models are unsuitable for testing anti-HIV activity.

Molecular topology overcomes the inherent problem in structure activity relationship (SAR) to quantify chemical structures by translation of chemical structures into characteristic numerical descriptors [8–13]. When vertices are interpreted as atoms and edge as bonds the resulting graphs represent connectivities within a molecule. The topological descriptors are derived from information based on this connec-

*To whom correspondence should be addressed: E-mail: madan_ak@yahoo.com

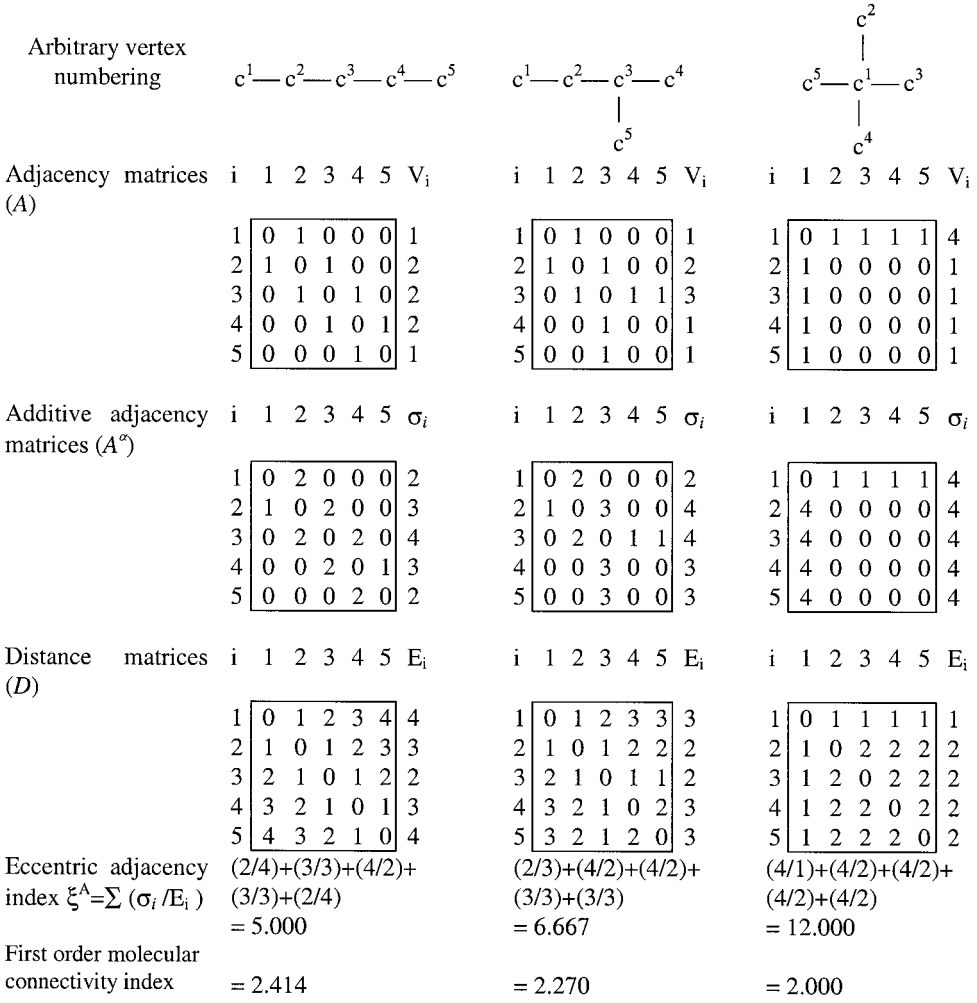

Figure 1. Calculation of *eccentric adjacency index* for three isomers of pentane.

tivity of a molecule.Since 3-D structure of a compound depends on the connectivity of its constituent atoms, the numerical topological descriptors derived from information based on connectivity can reveal structural or sub-structural information of a molecule. Although numerous numerical topological descriptors of diverse nature have been reported in literature but only a handful of them have been widely employed in structure-activity/property studies. These include Randic's molecular connectivity index [14–15], Hosoya's index [16–17], and Balaban's index [18–21]. Molecular topology as represented by the connectivity of the atoms can relate biological activity with the molecular graph or sub-graph of the compounds. Topological descriptors developed for predicting physicochemical properties and biological activities, of chemical substances; can be used for drug design [22–27].

Relationship of anti-HIV activity of HEPT derivatives with the first order molecular connectivity index and a novel topological descriptor termed as *eccentric adjacency index* has been investigated in the present study.

Materials and methods

Calculation of topological descriptor

In the present study a novel, adjacency-cum-distance based, topological descriptor termed as *eccentric adjacency index* has been proposed. *Eccentric adjacency index,* denoted by ξ^A, can be defined as the summation

of ratios, of sum of the degrees of adjacent vertices and eccentricity of the concerned vertex, for all vertices in the hydrogen suppressed molecular structure

$$\xi^A = \sum_{i=1}^{n} (\sigma_i / E_i),$$

where, σ_i is sum of degrees of vertices adjacent to vertex i, E_i is eccentricity of vertex i and n is the number of vertices in graph G. The eccentricity *Ei* of a vertex i in a graph G is the distance from vertex i to the vertex j that is farthest from i ($Ei = \max d(ij)$; $j \in G$). Distance d(ij) between vertices i, j in graph G is the length of the path that contain least number of edges between vertices i and j. The value of *eccentric adjacency index* can be calculated from the distance matrix (D) and a modified adjacency matrix termed as *additive adjacency matrix* (A^{α}). When, non-zero column elements in adjacency matrix represent the degree of corresponding vertex in the molecular graph, the matrix may be defined as *additive adjacency matrix*. Summations of the elements in each row of A^{α} yield the value of σ_i of the corresponding vertex. Calculation of the *eccentric adjacency index* values for three isomers of pentane is illustrated in Figure 1.

Randic [14], in 1975, proposed an index for characterization of branching. This index termed as molecular connectivity index and more specifically as first order molecular connectivity index by Kier and Hall [15], denoted by $^1\chi$, is defined as the summation of bond values (inverse square root of product of the degrees of adjacent vertices), for all the edges in the hydrogen suppressed molecular structure

$$^1\chi = \sum_{i=1, j=i+1}^{n-1,n} (Vi \times Vj)^{-1/2},$$

Where n is the number of vertices, *Vi* and *Vj* are the degrees of adjacent vertices i and j in graph G.

Model development/analysis

The data set comprising of 107 derivatives of 1-[(2-hydroxyethoxy)methyl]-6-(phenylthio) thymines (HEPT) comprised of both active and inactive derivatives (Figure 2). The values of *eccentric adjacency index* of all the derivative in the data set were computed using an in-house computer program. The resultant data was analyzed and active range was identified based on the maximization of moving average with respect to active compounds (<35% = inactive, 35–65% = transitional, ≥ 65% = active). Subsequently, each derivative was assigned a biological activity,

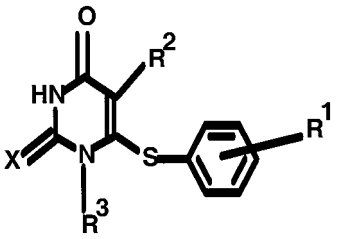

Figure 2. Basic structure of 6-(phenylthio) thymine.

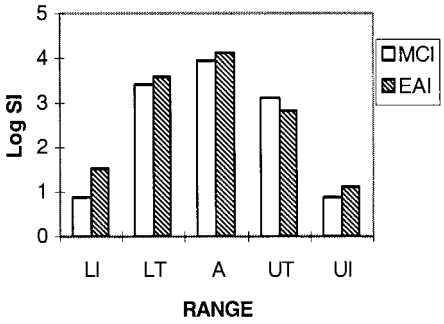

Figure 3. Average selectivity index (SI) values of various ranges (LI = lower inactive, LT = lower transitional, A = active, UT = upper transitional, UI = upper inactive). MCI = first order molecular connectivity index, EAI = eccentric adjacency index.

which was then compared with the anti-HIV activities of HEPT derivatives reported by Tanaka et al. [28–31]. The values of first order molecular connectivity index were also computed for all the derivatives using hydrogen-suppressed structures. Aforementioned procedure was similarly used to identify the active range.

The anti-HIV activity was quantitatively reportedby Tanaka et al. [28–31] as effective concentration required to achieve 50 percent protection (EC_{50}) of MT-4 cells against the cytopathic effect of HIV-1. Derivatives exhibiting EC_{50} less than 0.5 μMol were considered to be active. The cytotoxic activity reported for 101 derivatives was defined as cytotoxic concentration required to reduce the viability of mock infected MT-4 cells by 50% (CC_{50}). Selectivity index (SI) is defined as the quotient of CC_{50} and EC_{50}.

Results

The percent degree of prediction for each range was calculated from the ratio of number of derivatives with correctly predicted activity to that of the number of derivatives present in the range. The overall degree of prediction was obtained from the ratio of total number of derivatives with correctly predicted activity to that

Table 1. Relationship of *eccentric adjacency index* and first order molecular connectivity index with anti-HIV activity of 6-(phenylthio) thymine derivatives

S. No.	R^1	R^2	R^3	X	ξ^A	$^1\chi$	Anti-HIV activity			Selectivity index
							Assigned		Reported	
							ξ^A	$^1\chi$		
1	2-Me	Me	$CH_2OCH_2CH_2OH$	O	13.766	9.541	−	±	−	3.5
2	2-NO_2	Me	$CH_2OCH_2CH_2OH$	O	14.733	10.452	±	±	−	1.8
3	2-OMe	Me	$CH_2OCH_2CH_2OH$	O	14.159	10.079	−	±	−	13.2
4	3-Me	Me	$CH_2OCH_2CH_2OH$	O	13.704	9.542	−	±	−	161.5
5	3-Et	Me	$CH_2OCH_2CH_2OH$	O	13.408	10.062	−	±	–	67.0
6	3-t-Bu	Me	$CH_2OCH_2CH_2OH$	O	14.623	10.735	±	±	−	6.3
7	3-CF_3	Me	$CH_2OCH_2CH_2OH$	O	14.623	10.735	±	±	−	4.4
8	3-F	Me	$CH_2OCH_2CH_2OH$	O	13.704	9.524	−	±	−	85.5
9	3-Cl	Me	$CH_2OCH_2CH_2OH$	O	13.704	9.524	−	±	−	16.2
10	3-Br	Me	$CH_2OCH_2CH_2OH$	O	13.704	9.524	−	±	−	24.7
11	3-I	Me	$CH_2OCH_2CH_2OH$	O	13.704	9.524	−	±	−	10.6
12	3-NO_2	Me	$CH_2OCH_2CH_2OH$	O	13.392	10.435	−	±	−	5.0
13	3-OH	Me	$CH_2OCH_2CH_2OH$	O	13.704	9.542	−	±	−	5.4
14	3-OMe	Me	$CH_2OCH_2CH_2OH$	O	13.408	10.062	−	±	−	11.4
15	3,5-Me_2	Me	$CH_2OCH_2CH_2OH$	O	14.279	9.918	±	±	+	934.6
16	3,5-Cl_2	Me	$CH_2OCH_2CH_2OH$	O	14.279	9.918	±	±	−	100.0
17	3,5-Me_2	Me	$CH_2OCH_2CH_2OH$	S	14.279	9.918	±	±	+	781.8
18	3-COOMe	Me	$CH_2OCH_2CH_2OH$	O	13.274	10.973	−	+	−	28.0
19	3-COMe	Me	$CH_2OCH_2CH_2OH$	O	13.932	10.435	−	±	−	31.2
20	3-CN	Me	$CH_2OCH_2CH_2OH$	O	13.403	10.062	−	±	−	23.4
21	H	$CH_2CH{=}CH_2$	$CH_2OCH_2CH_2OH$	O	14.109	10.168	−	±	−	73.2
22	H	Et	$CH_2OCH_2CH_2OH$	S	13.673	9.668	−	±	+	1345.5
23	H	Pr	$CH_2OCH_2CH_2OH$	S	14.109	10.168	−	±	−	23.0
24	H	i-Pr	$CH_2OCH_2CH_2OH$	S	14.386	10.041	±	±	+	6793.0
25	3,5-Me_2	Et	$CH_2OCH_2CH_2OH$	S	14.712	10.502	±	±	+	35512.8
26	3,5-Me_2	i-Pr	$CH_2OCH_2CH_2OH$	S	15.535	10.829	+	+	+	10375.4
27	3,5-Cl_2	Et	$CH_2OCH_2CH_2OH$	S	14.823	10.456	±	±	+	1500.3
28	H	Et	$CH_2OCH_2CH_2OH$	O	13.673	9.668	−	±	+	3327.1
29	H	Pr	$CH_2OCH_2CH_2OH$	O	14.109	10.168	−	±	−	71.8
30	H	i-Pr	$CH_2OCH_2CH_2OH$	O	14.386	10.041	±	±	+	3661.1
31	3,5-Me_2	Et	$CH_2OCH_2CH_2OH$	O	14.823	10.456	±	±	+	11566.1
32	3,5-Me_2	i-Pr	$CH_2OCH_2CH_2OH$	O	15.535	10.829	+	+	+	47407.4
33	3,5-Cl_2	Et	$CH_2OCH_2CH_2OH$	O	14.823	10.456	±	±	+	3610.5
34	4-Me	Me	$CH_2OCH_2CH_2OH$	O	13.061	9.524	−	±	−	1.1
35	H	Me	$CH_2OCH_2CH_2OH$	S	12.475	9.130	−	−	−	105.7
36	H	Me	$CH_2OCH_2CH_2OH$	O	12.475	9.130	−	−	−	125.5
37	H	I	$CH_2OCH_2CH_2OH$	O	12.475	9.130	−	−	−	5.6
38	H	$CH{=}CH_2$	$CH_2OCH_2CH_2OH$	O	13.173	9.668	−	±	−	69.1
39	H	CH=CHPh	$CH_2OCH_2CH_2OH$	O	14.983	12.186	±	−	−	15.8
40	H	CH_2Ph	$CH_2OCH_2CH_2OH$	O	15.845	11.686	±	±	−	1.0
41	H	$CH{=}CPh_2$	$CH_2OCH_2CH_2OH$	O	18.125	14.669	−	−	−	25.0
42	H	Me	$CH_2OCH_2CH_2OMe$	O	12.609	9.630	−	±	−	34.4
43	H	Me	$CH_2OCH_2CH_2OAc$	O	12.302	10.486	−	±	−	NA
44	H	Me	$CH_2OCH_2CH_2OCOPh$	O	11.893	12.559	−	−	−	NA
45	H	Me	CH_2OEt	O	13.887	8.630	−	−	+	700.0
46	H	Me	$CH_2OCH_2CH_2Cl$	O	12.475	9.130	−	−	−	130.7

Table 1 continued.

S. No.	R^1	R^2	R^3	X	ξ^A	$^1\chi$	Anti-HIV activity			Selectivity index
							Assigned		Reported	
							ξ^A	$^1\chi$		
47	H	Me	$CH_2OCH_2CH_2N_3$	O	11.820	10.130	−	±	−	32.1
48	H	Me	$CH_2OCH_2CH_2F$	O	12.475	9.130	−	−	−	190.0
49	H	Me	CH_2OPr	O	12.475	9.130	−	−	−	40.8
50	H	Me	CH_2OCH_2Ph	O	13.247	10.648	−	±	+	1079.5
51	H	Et	CH_2OEt	O	14.500	9.168	±	±	+	8449.4
52	H	Et	CH_2OEt	S	14.500	9.168	±	±	+	3079.5
53	3,5-Me_2	Et	CH_2OEt	O	15.772	9.956	+	±	+	18518.5
54	3,5-Me_2	Et	CH_2OEt	S	15.772	9.956	+	±	+	22727.3
55	H	Et	CH_2OCH_2Ph	O	13.640	11.186	−	+	+	5762.7
56	3,5-Me_2	Et	CH_2OCH_2Ph	O	14.604	11.974	±	±	+	6250.0
57	H	Et	CH_2OCH_2Ph	S	13.640	11.186	−	+	+	12820.5
58	3,5-Me_2	Et	CH_2OCH_2Ph	S	14.604	11.974	±	±	+	2898.6
59	H	i-Pr	CH_2OEt	O	15.309	9.941	+	±	+	8833.3
60	H	i-Pr	CH_2OCH_2Ph	O	14.214	11.559	±	+	+	7407.4
61	H	i-Pr	CH_2OEt	S	15.309	9.541	+	±	+	7142.9
62	H	i-Pr	CH_2OCH_2Ph	S	14.214	11.559	±	+	+	2941.2
63	H	Me	CH_2OMe	O	14.199	8.130	−	−	−	116.2
64	H	Me	CH_2OBu	O	12.609	9.630	−	±	−	17.7
65	H	Me	Et	O	14.574	7.630	±	−	−	42.7
66	H	Me	Bu	O	13.887	8.630	−	−	−	74.2
67	3,5-Cl_2	Et	CH_2OEt	S	15.772	9.556	+	±	+	3493.1
68	H	Et	CH_2O-i-Pr	S	15.136	9.524	+	±	+	454.5
69	H	Et	CH_2O-c-Hex	S	14.538	10.686	±	±	−	139.4
70	H	Et	$CH_2OCH_2CH_2OH$	S	13.640	11.186	−	+	+	285.7
71	H	Et	$CH_2OCH_2C_6H_4O$(4-Me)	S	13.172	11.580	−	+	+	256.4
72	H	Et	$CH_2OCH_2C_6H_4O$(4-Cl)	S	13.172	11.580	−	+	+	1663.5
73	H	Et	$CH_2OCH_2CH_2Ph$	S	12.900	11.686	−	±	+	219.3
74	3,5-Cl_2	Et	CH_2OEt	O	15.772	9.956	+	±	+	6081.1
75	H	Et	CH_2O-i-Pr	O	15.136	9.524	+	±	+	420.6
76	H	Et	CH_2O-c-Hex	O	14.613	10.686	±	±	−	25.0
77	H	Et	CH_2OCH_2-c-Hex	O	13.640	11.186	−	+	+	37.8
78	H	Et	$CH_2OCH_2CH_2Ph$	O	12.900	11.686	−	±	+	395.8
79	H	c-Pr	CH_2OEt	S	16.095	9.203	±	±	+	481.7
80	H	c-Pr	CH_2OEt	O	16.095	9.203	±	±	+	2240.0
81	H	Me	$CH_2OCH_2CH_2OC_5H_{11}$-n	O	10.287	11.842	−	±	−	1.0
82	2-Cl	Me	$CH_2OCH_2CH_2OH$	O	13.766	9.541	−	±	−	1.0
83	3-CH_2OH	Me	$CH_2OCH_2CH_2OH$	O	13.408	10.062	−	±	−	1.0
84	4-F	Me	$CH_2OCH_2CH_2OH$	O	13.016	9.524	−	±	−	1.0
85	4-Cl	Me	$CH_2OCH_2CH_2OH$	O	13.016	9.524	−	±	−	1.0
86	3-NO_2	Me	$CH_2OCH_2CH_2OH$	O	12.840	10.435	−	±	−	1.0
87	3-CN	Me	$CH_2OCH_2CH_2OH$	O	12.358	10.062	−	±	−	1.0
88	3-OH	Me	$CH_2OCH_2CH_2OH$	O	13.016	9.524	−	±	−	1.0
89	3-OMe	Me	$CH_2OCH_2CH_2OH$	O	12.358	10.062	−	±	−	1.0
90	3-COMe	Me	$CH_2OCH_2CH_2OH$	O	12.840	11.435	−	+	−	1.0
91	3-COOH	Me	$CH_2OCH_2CH_2OH$	O	13.932	10.435	−	±	−	1.0
92	3-$CONH_2$	Me	$CH_2OCH_2CH_2OH$	O	13.932	10.435	−	±	−	1.0
93	H	COOMe	$CH_2OCH_2CH_2OH$	O	14.822	10.579	±	±	−	1.0
94	H	CONHPh	$CH_2OCH_2CH_2OH$	O	14.509	12.597	±	−	−	1.0
95	H	SPh	$CH_2OCH_2CH_2OH$	O	15.845	11.686	±	±	−	1.0

Table 1 continued.

S. No.	R^1	R^2	R^3	X	ξ^A	$^1\chi$	Anti-HIV activity			Selectivity index
							Assigned		Reported	
							ξ^A	$^1\chi$		
96	H	C≡CH	$CH_2OCH_2CH_2OH$	O	13.673	9.668	−	±	−	1.0
97	H	C≡CPh	$CH_2OCH_2CH_2OH$	O	14.983	12.186	±	−	−	1.0
98	3-NH_2	Me	$CH_2OCH_2CH_2OH$	O	13.704	9.524	−	±	−	NA
99	H	$COCHMe_2$	$CH_2OCH_2CH_2OH$	O	15.458	10.952	+	+	−	1.0
100	H	COPh	$CH_2OCH_2CH_2OH$	O	16.558	12.113	−	−	−	1.0
101	H	C≡CMe	$CH_2OCH_2CH_2OH$	O	14.109	10.168	−	±	−	1.0
102	H	F	$CH_2OCH_2CH_2OH$	O	13.130	9.130	−	−	−	NA
103	H	Cl	$CH_2OCH_2CH_2OH$	O	13.130	9.130	−	−	−	NA
104	H	Br	$CH_2OCH_2CH_2OH$	O	13.130	9.130	−	−	−	NA
105	H	Me	$CH_2OCH_2CH_2OCH_2Ph$	O	11.411	12.148	−	−	−	2.3
106	H	Me	H	O	13.076	6.665	−	−	−	1.0
107	H	Me	Me	O	14.014	7.092	−	−	−	0.6

Table 2. The relationship between anti-HIV activity of 6-(phenylthio) thymine derivatives and first order molecular connectivity index

Range	Value	Predicted derivatives		Accuracy (%)	Average EC_{50}	Average selectivity index
		Total	Correct			
Lower inactive	<9.15	15	14	93.33	44.92 (48.10)	127.75 (7.54)
Lower transitional	9.15–10.75	65	NA	NA	55.28 (NA)	2430.24 (NA)
Active	10.75–11.60	13	10	76.92	10.06 (0.092)	6845.23 (8895.80)
Upper transitional	11.60–12.00	7	NA	NA	14.17 (NA)	1395.24 (NA)
Upper inactive	≥12.00	7	7	100.00	9.83 (9.83)	7.68 (7.68)

Values in the bracket are based upon correctly predicted compounds in the particular range.

of total number of derivatives present in both the active and inactive ranges. The results are summarized in Tables 1–3. Using *eccentric adjacency index* the study revealed the following information:

- Biological activity was assigned to a total of 79 derivatives in both the active and inactive ranges, out of which activity of 65 derivatives was correctly predicted resulting in about 82 percent accuracy, with respect to anti-HIV activity.
- Two transitional ranges bracketed the active range, indicating a gradual change in anti-HIV activity from inactive to active range (Figure 3). A total of 28 derivatives were present in the transitional range.
- The active range had eccentric adjacency index value of 15.00–15.80. As many as 91 percent derivatives in the active range exhibited anti-HIV activity. The average EC_{50} of correctly predicted derivatives was observed to be 0.062 μMol.

The accuracy of prediction for active and inactive ranges is compiled in Tables 2 and 3.

Similarly, the study using first order molecular connectivity index revealed following information

- Biological activity was assigned to a total of 35 derivatives in both the active and inactive ranges, out of which anti-HIV activity of 31 derivatives was correctly predicted.
- Lower and upper transitional ranges were observed, indicating a gradual change in anti-HIV activity from inactive to active range. A total of 72 derivatives were present in the transitional range.
- The active range had first order molecular connectivity index value of 10.75–11.60. About 77% of the derivatives exhibited anti-HIV activity. The average EC_{50} of correctly predicted derivatives was observed to be 0.092 μMol.

It is noteworthy that eccentric adjacency index possess eccentricity features that differentiates it from first

Table 3. The relationship between anti-HIV activity of 6-(phenylthio) thymine derivatives and eccentric adjacency index

Range	Value	Predicted derivatives		Accuracy (%)	Average EC_{50}	Average selectivity index
		Total	Correct			
Lower inactive	<14.20	66	53	80.30	62.94 (78.35)	540.14 (34.38)
Lower transitional	14.20–15.00	24	NA	NA	10.03 (NA)	3867.75 (NA)
Active	15.00–15.80	11	10	90.91	1.15 (0.062)	11405.01 (12545.41)
Upper transitional	15.80–16.50	4	NA	NA	11.05 (NA)	680.92 (NA)
Upper inactive	≥16.50	2	2	100.00	6.92 (6.92)	13.00 (13.00)

Values in the bracket are based upon correctly predicted compounds in the particular range.

order molecular connectivity index. This is evidenced by the fact that only three derivatives (26, 32, and 99) are predicted commonly in the active range.

Discussion

Though a large number of topological descriptors have been reported in literature for structure property studies but only a handful of them have been successfully used in structure activity relationship studies. As a consequence there is a strong need to develop topological descriptors with high discriminating power particularly with respect to structure activity studies.

A novel adjacency-cum-distance based index termed as *eccentric adjacency index* conceptualized in the present investigation can be simply calculated from a modified adjacency matrix termed as *additive adjacency matrix*. *Eccentric adjacency index* is not based on any physicochemical properties of the molecules but on the simple fact that a necessary condition for drug-receptor interaction is proper fit like lock and key. Hence the index in the present form is independent of the bonding environment. *Eccentric adjacency index* facilitates direct structural interpretations because it takes into consideration the eccentricity in addition to degree of adjacent vertices in the graph.

In order to explore the utility of *eccentric adjacency index* in structure-activity studies, a relationship between eccentric adjacency index and anti-HIV activity of 1-[(2-hydroxyethoxy)methyl 6-(phenylthio) thymines was investigated. These derivatives are selective inhibitors of reverse transcriptase of HIV-1. Both active and inactive derivatives were included in the set by random selection. *Eccentric adjacency index* values were computed for all the derivatives in the data set and the resulting data was analyzed. The active derivatives were found to be present in the narrow range of index values, thus, facilitating rapid identification of the active range. However, active range bracketed by transitional ranges basically comprises of a general model while the active range constitutes a specific model. Afterall, specific model is always a subset of the general model and consequently account for less number of compounds.

The absence of any direct correlation between *eccentric adjacency index* and first order molecular connectivity index indicate that the two indices are distinctive and consider different structural components. Besides this, the comparison of results obtained by *eccentric adjacency index* and first order molecular connectivity index disclose the following information:

- The high sensitivity of eccentric adjacency index with changes in the molecular structure as compared to first order molecular connectivity index is evident in case of three isomers of pentane (Figure 1). As the eccentric adjacency index increases, the first order molecular connectivity index decreases non-linearly. Also relative change in the value of eccentric adjacency index with the change in the molecular structure is almost double the corresponding relative change in the value of first order molecular connectivity index.
- In a data set comprising of 107 derivatives of HEPT 74 percent derivatives were classified into active and inactive ranges using eccentric adjacency index as compared to only 33 percent using first order molecular connectivity index.
- The gradual movement from inactive to active range results in steep increase in the value of selectivity index in case of both eccentric adjacency index and first order molecular connectivity index. But the average selectivity index value (11405) in the active range obtained by eccentric adjacency index is higher than the corresponding value (6845) obtained by first order molecular connectivity index. *Since selectivity index is an indirect indicator of therapeutic index, therefore, anti-HIV*

activity amalgamated with low toxicity highlights the potential of this range.

In order to explain the predicted activity it is important to offer some insight into the significance of these indices:

- Eccentric adjacency index encodes information about the sensitivity of R^2 substituents with regard to presence of cyclic fragments, degree of branching or connectivity and shows that a maximum activity can be encountered when an optimum value is reached. The substitution of methyl group at the C-5 position of phenylthio thymines by either ethyl or isopropyl increase the anti-HIV activity while the presence of aromatic rings result in inactive derivatives.
- First order molecular connectivity index on the other hand conveys information about the shape characteristics of whole molecule. Increasing the number of R^1 substituents result in greater anti-HIV activity while unsubstituted derivatives are inactive. The combined effects of disubstitution and branching increases activity. The presence of aromatic ring in R^2 substituents results in inactive derivatives while phenyl ring in R^3 substituents enhances anti-HIV activity.
- Inspite of structural interpretations the prediction using eccentric adjacency index was superior to that using first order molecular connectivity index in the active range. However, these two indices can be used in combination, as same compounds are not categorized in a particular range. Infact the biological activity of a derivative predicted as transitional by eccentric adjacency index can be confirmed by using first order molecular connectivity index.

In the present study *eccentric adjacency index* has proved the discriminating power with regard to atleast one property – anti-HIV activity – in a data set consisting of relatively large number of derivatives. *Eccentric adjacency index* along with other mathematical tools can provide valuable leads for development of potent therapeutic agents for treatment of HIV infection. Moreover, the high discriminating power of *eccentric adjacency index* offers a vast potential for other structure-activity/property studies.

References

1. De Clercq, E., J. Med. Chem., 29 (1986) 1561–1569.
2. Lin, T. -S., Chen, M.S., McLaren, C., Gao, Y.S., Ghazzouli, I. and Prusoff, W.H., J. Med. Chem., 30 (1987) 440-44-4.
3. Kim, C.H., Marquez, V.E., Broder, S., Mitsuya, H. and Driscoll, J.S., J. Med. Chem., 30 (1987) 862-866.
4. Herdewijn, P., Balzarani, J., De Clercq, E., Pauwels, R., Baba, M., Broder, S. and Vanderhaeghe, H., J. Med. Chem., 30 (1987) 1270–1278.
5. Lin, T.S., Guo, J.Y., Schinazi, R.F., Chu, C.K., Xiang, J.N. and Prusoff, W.H., J. Med. Chem., 31 (1988) 336–40.
6. Herdewijn, P., Balzarani, J., Baba, M., Pauwels, R., Van Aerschot, A., Janssen, G. and De Clercq, E., J. Med. Chem., 31 (1988) 2040-2048.
7. Miyasaka, T., Tanaka, H., Baba, M., Hayakawa, H., Walker R.T., Balzarani, J. and De Clercq, E., J. Med. Chem., 32 (1989) 2507-2509.
8. Trinajstic, N., Chemical Graph Theory, CRC Press, Boca Raton, FL, 1983, Vols. I and II.
9. Balaban, A.T., Motoc, J., Bonchev, D. and Mekennyan, O., Top. Curr. Chem., 114 (1983) 21–55.
10. Kier, L.B. and Hall, L.H., Molecular Connectivity in Structure Activity Analysis, Research Studies Press, Letchworth, UK, 1986.
11. Katritzky, A.R. and Gordeeva, E.V., J. Chem. Inf. Comput. Sci., 33 (1993) 835–857.
12. Basak, S.C., Bertlsen, S. and Grunwold, G.D. J. Chem. Inf. Comput. Sci., 34 (1994) 270–276.
13. Estrada, E. and Ramirez, A., J. Chem. Inf. Comput Sci., 36 (1996) 837–843.
14. Randic, M., J. Am. Chem. Soc., 97 (1975) 6609–6615.
15. Kier, L.B., Hall, L.H., Murray W.J. and Randic, M. J. Pharm. Sci., 64 (1975) 1971–1974.
16. Hosoya, H., Bull. Chem. Soc. Jpn., 44 (1971) 2332–2337.
17. Hosoya, H., J. Chem. Doc., 12 (1972) 181-183.
18. Balaban, A.T., Chiriac, A., Motoc I. and Simon, Z., Lect. Notes Chem., 15 (1980) 22–27.
19. Balaban, A.T., Chem. Phys. Lett., 89 (1982) 399-404.
20. Balaban, A.T. and Fillip, P., MATCH Commun. Math. Chem., 16 (1984) 163–190.
21. Balaban, A.T., J. Chem. Inf. Comput. Sci., 25 (1985) 334–343.
22. Sabljic, A. and Trinajstic, N., Acta Pharm. Jugsol., 31 (1981) 189–214.
23. Rouvray, D.H., J. Mol. Struct. (THEOCHEM), 185 (1989) 187–201.
24. Galvez, J., Garcia-Domenec, R., deJulian-Ortiz, J.V. and Soler, R., J. Chem. Inf. Comput. Sci., 35 (1995) 272–284.
25. Sharma, V., Goswami, R. and Madan, A.K., J. Chem. Inf. Comput. Sci., 37 (1997) 273-82.
26. Garcia-Domenec, R., deGregario-Alapont, C., deJulian-Ortiz, J.V., Galvez, J. and Popa, L., Bioorg. Med. Chem. Lett., 7 (1997) 567–572.
27. Gupta, S., Singh, M. and Madan, A.K., J. Mol. Struct. (THEOCHEM), in press.
28. Tanaka, H., Baba, M., Hayakawa, H., Sakamaki, T., Miyasaka, T., Ubasawa, M., Takashima, H., Sekiya, K., Nitta, I., Shigeta, S., Walker, R.T., Balzarani, J. and De Clercq, E., J. Med. Chem., 34 (1991) 349–357.
29. Tanaka, H., Baba, M., Saito, S., Miyasaka, T., Takashima, H., Sekiya, K., Ubasawa, M., Nitta, I., Walker, R.T., Takashima, H. and De Clerecq, E., J. Med. Chem., 34 (1991) 1508–1511.
30. Tanaka, H., Takashima, H., Ubasawa, M., Sekiya, K., Nitta, I., Baba, M., Shigeta, S., Walker, R.T., De Clercq, E. and Miyasaka, T., J. Med. Chem., 35 (1992) 337–345.
31. Tanaka, H., Takashima, H., Ubasawa, M., Sekiya, K., Nitta, I., Baba, M., Shigeta, S., Walker, R.T. and De Clercq, E. and Miyasaka, T. J. Med. Chem., 35 (1992) 4713–4719.