ACADEMIC PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

J. Math. Anal. Appl. 275 (2002) 386–401

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS

www.academicpress.com

Eccentric distance sum: A novel graph invariant for predicting biological and physical properties

S. Gupta,[a] M. Singh,[a] and A.K. Madan [b,*]

[a] *Department of Pharmaceutical Sciences and Drug Research, Punjabi University, Patiala 147002, India*

[b] *Faculty of Pharmaceutical Sciences, M.D. University, Rohtak 124001, India*

Received 28 December 2000

Submitted by H.P. Benson

Abstract

Eccentric distance sum—a novel graph invariant with vast potential in structure activity/property relationships has been conceptualized in the present study. This graph invariant displayed high discriminating power with respect to both biological activity and physical properties. The structure activity relationship of *eccentric distance sum* was investigated with regard to anti-HIV activity of dihydroseselins. The values *of eccentric distance sum* of each analogue in the data set were computed and active range identified. Subsequently, biological activity was assigned to each analogue in the data set, which was then compared with the reported anti-HIV activity of dihydroseselin analogues. Surprisingly the accuracy of prediction was found to be more than 88% with regard to anti-HIV activity. On the other hand, investigations pertaining to quantitative structure property relationship of the novel graph invariant with regard to various physical properties of diverse nature, for data sets consisting of primary amines, secondary amines and alcohols revealed correlation percentages ranging from 93% to 99%. The over all results with regard to structure–activity and quantitative structure–property studies using eccentric distance sum were better than the corresponding values obtained using Wiener's index.

* Corresponding author.
E-mail address: madan_ak@yahoo.com (A.K. Madan).

0022-247X/02/$ – see front matter
PII: S0022-247X(02)00373-6

1. Introduction

Graph invariant is a graph theoretic property that is preserved by isomorphism. The chemical information derived through graph invariants has been found useful in chemical documentation, isomer discrimination, structure property correlations, etc. [1]. Structure activity/property relationships (SAR/SPRs) are models that relate structural aspects of a molecule to its physicochemical properties. The inherent problem in the development of a suitable correlation between chemical structures and physical properties can be attributed to the non-quantitative nature of chemical structures. Graph theory can be employed through translation of chemical structures into characteristic polynomial, matrix, sequence or numerical graph invariants [2–7]. Since structure of an analogue depends on the connectivity of its constituent atoms, the numerical graph invariants derived from information based on connectivity can reveal structural or substructural information of a molecule. Molecular topology as represented by the connectivity of the atoms can relate physical properties and biological activity [8] with the analogues.

Acquired immunodeficiency syndrome (AIDS), the immunosuppressive disease caused by human immunodeficiency virus (HIV), by virtue of high mortality rate and incurability is a serious threat to public health. The causative agents, termed as HIV-1 and HIV-2, are retroviruses. Various compounds have been reported by De Clercq [9] to inhibit the replication of HIV-1 *in vitro*. First generation anti-HIV compound such as 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. 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 acts as inhibitor 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. The search of more selective and effective agents against HIV now focuses on derivatives with novel structures or active through new mechanisms of action.

In present study the relationship of *eccentric distance sum*—a novel graph invariant and Wiener's topological index with anti-HIV activity of dihydroseselins has been investigated to facilitate the development of potent and safe anti-HIV agents. The relationship of *eccentric distance sum* and Wiener's topological index with physical properties in data sets of diverse nature was also investigated.

1.1. Calculation of graph invariant

Eccentric distance sum, denoted by ξ^{DS}, can be defined as the summation of product of eccentricity and distance sum of each vertex in the hydrogen

suppressed molecular graph having n vertices,

$$\xi^{\mathrm{DS}}(G) = \sum_{i=1}^{n} (E_i * S_i),$$

where S_i is distance sum of vertex i, E_i is eccentricity of vertex i and n is the number of vertices in graph G. *Eccentric distance sum* takes into consideration the eccentricity and distance sum of all vertices in the graph. The eccentricity E_i of a vertex i in a graph G is the path length from vertex i to the vertex j that is farthest from i ($E_i = \max d(ij),\ j \in G$).

The Wiener's topological index [10] of a hydrogen suppressed molecular graph is defined as the sum of the distances between all pairs of vertices,

$$W(G) = \frac{1}{2}\left(\sum_{i=1}^{n} d(ij)\right),$$

where $d(ij)$ is length of the path that contains the least number of edges between vertex i and vertex j in graph G; n is the maximum possible number of i and j.

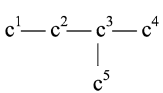
Eccentric distance sum and Wiener's index can be easily calculated from the distance matrix. Calculation of *eccentric distance sum* and *Wiener's index* of hydrogen suppressed molecular graphs for three isomers of pentane has been exemplified in Fig. 1.

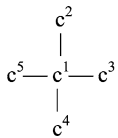
1.2. Model development/analysis

A data set of 48 analogues used for investigating structure–activity relationship consisted of both active and inactive dihydroseselins. The values of *eccentric distance sum* and Wiener's index of each analogue in the data set were computed using an in-house computer program and active ranges identified based on the maximization of moving average with respect to active analogues. Subsequently, each analogue was assigned a biological activity which was then compared with the reported [11] anti-HIV activity. The biological activity was reported [11] in terms of effective concentration (EC_{50}, μM) for anti-HIV activity in acutely infected H9 lymphocytes. The analogues exhibiting $EC_{50} < 25$ μM were considered to be potentially active for the purpose of present study (Table 1). The percent degree of prediction for each range was calculated from the ratio of the number of analogues with correctly predicted activity to that of total number of analogues present in the respective range. The overall degree of prediction was obtained from the ratio of total number of analogues with correctly predicted activity to that of total number of analogues present in both active and inactive ranges (Tables 2 and 3).

Various data sets selected for the structure-property correlation comprised of cavity surface areas (csa) of 51 and boiling points (bp) of 62 straight chain, branched chain and cyclic alcohols [12,13]; boiling points of 21 primary

Arbitrary vertex numbering





Path length matrices (P)

i	1	2	3	4	5	Si	E_i
1	0	1	2	3	4	10	4
2	1	0	1	2	3	7	3
3	2	1	0	1	2	6	2
4	3	2	1	0	1	7	3
5	4	3	2	1	0	10	4

i	1	2	3	4	5	Si	E_i
1	0	1	2	3	3	9	3
2	1	0	1	2	2	6	2
3	2	1	0	1	1	5	2
4	3	2	1	0	2	8	3
5	3	2	1	2	0	8	3

i	1	2	3	4	5	Si	E_i
1	0	1	1	1	1	4	1
2	1	0	2	2	2	7	2
3	1	2	0	2	2	7	2
4	1	2	2	0	2	7	2
5	1	2	2	2	0	7	2

Eccentric distance sum

$\xi^{DS}(G) = \sum_{i=1}^{n} (E_i * S_i)$

(10*4)+(7*3)+(6*2)+ (7*3)+(10*4) = 134

(9*3)+(6*2)+(5*2)+(8*3)+(8*3) = 97

(4*1)+(7*2)+(7*2)+ (7*2)+(7*2) = 60

Wiener's index

$W(G) = \frac{1}{2} (\sum_{i=1}^{n} d(ij))$

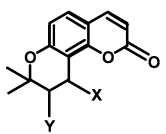
½ (10+7+6+7+10) =20

½ (9+6+5+8+8) = 18

½ (4+7+7+7+7) = 16

Fig. 1. Calculation of *eccentric distance sum* and Wiener's index values for three isomers of pentane.

Table 1
Relationship of anti-HIV activity of dihydroseselin analogues with *eccentric distance sum* and Wiener's index



Y = OR^1 (S.NO. 02-48)
X = OR^2 (S.NO. 02-43)
X = N_3 (S.NO. 44-45)
X = NHR^2 (S.NO. 46-48)

S.No.	C/OA	R^1	R^2	W	ξ^{DS}	Anti-HIV activity		
						Predicted		Reported
						W	ξ^{DS}	
1		–	–	1814	6172	±	–	+
2	Cis	$COCH_3$	$COCH_2CH(CH_3)_2$	1791	29940	±	±	+
3		$COCH_3$	$COC(CH_3)=CHCH_3$	474	29375	–	±	+
4	(±) cis	H	H	614	8059	–	–	NS
5	(+) cis pure	H	H	614	8059	–	–	NS
6	(±) trans	H	3-chlorobenzoyl	1865	33843	±	±	+
7	(±) trans	H	H	614	8059	–	–	NS
8	Cis	$COCH_3$	$COCH_3$	1298	19476	–	–	–
9	Cis	$COCH_2CH(CH_3)_2$	$COCH_2CH(CH_3)_2$	2444	44834	+	+	+
10		$COCH(C_2H_5)CH_3$	$COCH(C_2H_5)CH_3$	2392	43437	±	±	+
11		$COCH=C(CH_3)_2$	$COCH=C(CH_3)_2$	2444	44834	+	+	+
12	Racemic	$COCHC(CH_3)_3$	$COCHC(CH_3)_3$	2894	53864	+	+	NS
13		4-tert-butylbenzoyl	4-tert-butylbenzoyl	6326	170585	±	±	+
14	Pure	3-menthyloxycarbonyl	3-menthyloxycarbonyl	6988	173785	±	±	NS
15	Pure	3-menthyloxycarbonyl	3-menthyloxycarbonyl	6988	173785	±	±	+
16	(+) cis pure	Camphanoyl	Camphanoyl	6186	133714	±	±	+
17	(−) cis pure	Camphanoyl	Camphanoyl	6186	133714	±	±	–
18		H	$COCH(C_2H_5)CH_3$	1338	21706	–	–	NS
19		$COCH(C_2H_5)CH_3$	H	1380	24492	–	–	NS

Table 1 (*Continued*)

S.No.	C/OA	R^1	R^2	W	ξ^{DS}	Anti-HIV activity		
						Predicted		Reported
						W	ξ^{DS}	
20		$COCH_3$	$COCH(C_2H_5)CH_3$	1791	29375	±	±	NS
21	Pure	$COCH_2C(CH_3)_3$	$COCH_2C(CH_3)_3$	2894	53864	+	+	+
22		$COCH_3$	3-chlorobenzoyl	2402	43987	±	±	NS
23		$COCH(CH_3)_2$	3-chlorobenzoyl	3131	61119	+	+	+
24		CH_3	3-chlorobenzoyl	2024	36741	±	±	−
25		$CH_2C_6H_5$	3-chlorobenzoyl	3440	72485	+	+	+
26	R	Tetrahydropyran-2-yl	3-chlorobenzoyl	3098	60192	+	+	NS
27	S	Tetrahydropyran-2-yl	3-chlorobenzoyl	3098	60192	+	+	+
28		$CH_2C_6H_5$	H	1599	30855	−	±	−
29		CH_3	H	708	9490	−	−	NS
30		Tetrahydropyran-2-yl	H	1376	24410	−	−	NS
31		$CH_2C_6H_5$	$COCH_3$	2066	39950	±	±	−
32		CH_3	$COCH_3$	1028	14108	−	−	NS
33		Tetrahydropyran-2-yl	$COCH_2CH(CH_3)_2$	2414	44049	±	±	NS
34		Tetrahydropyran-2-yl	$COCH_3$	1808	32050	±	±	NS
35		H	$COCH_2CH(CH_3)_2$	1358	22192	−	−	NS
36	Trans	$COCH_3$	$COCH_2CH(CH_3)_2$	1818	29940	±	±	NS
37	Trans	$COCH_3$	$COCH_3$	1298	19476	−	−	−
38	(−) trans pure	Camphanoyl	Camphanoyl	6186	133714	±	±	+
39	(+) trans pure	Camphanoyl	Camphanoyl	6186	133714	±	±	−
40		$COOC_6H_5$	$COOC_6H_5$	4124	97608	±	±	NS
41		$COCH_2CH_3$	$COCH_2CH_3$	3550	75789	±	±	NS
42	Trans	$COCH_2CH(CH_3)_2$	$COCH_2CH(CH_3)_2$	2444	44834	+	+	+
43		CH_3	CH_3	800	10754	−	−	NS
44	Cis/trans	H	–	808	10876	−	−	−
45	Cis/trans	$COCH_3$	–	1169	17476	−	−	−
46		H	$COCH_3$	917	12473	−	−	NS
47	Trans	H	$COCH_2CH(CH_3)_2$	1358	22192	−	−	NS
48	Trans	$COCH_3$	$COCH_2CH(CH_3)_2$	1814	29940	±	±	NS

C/OA = configuration/optical activity, NS = no suppression; +, active compound, −, inactive compound, and ±, compound in the transitional range.

Table 2
The relationship between anti-HIV activity and *eccentric distance sum*

Range	Value	Predicted analogues		Accuracy	NS	Anti-HIV activity
		Total	Correct	(%)		(μM)
Lower inactive	< 25000	17	16	94.12	12	69.458
Lower transitional	25000–44750	13	NA	NA	6	48.627
Active	44750–75000	9	7	77.78	2	10.728
Upper transitional	$\geqslant 75000$	9	NA	NA	3	35.599

NA = not applicable, NS = no suppression.

Table 3
The relationship between anti-HIV activity and *Wiener index*

Range	Value	Predicted analogues		Accuracy	NS	Anti-HIV activity
		Total	Correct	(%)		(μM)
Lower inactive	< 1600	18	17	94.44	12	65.050
Lower transitional	1600–2425	12	NA	NA	6	49.566
Active	2425–3500	9	7	77.78	2	10.728
Upper transitional	$\geqslant 3500$	9	NA	NA	3	35.599

NA = not applicable, NS = no suppression.

and 13 secondary amines; molar refraction (mr) values of 48 heterogeneous compounds consisting of ethers, amines and alcohols [14]. The values of *eccentric distance sum* of each analogue in the data sets were computed using an in-house computer program. The resultant data was subjected to regression analysis and mathematical models, based on eccentric distance sum, were developed for various physical properties of primary amines, secondary amines and alcohols. The results are summarized in Tables 4–7.

2. Results and discussion

Although, in contemporary medicinal chemistry, quantum chemical and physicochemical parameters have been extensively utilized in the prediction of biological activity of molecules, the inherent problem in structure activity relationship is that qualitative changes in chemical structure can not be directly related to quantum change in biological activity. The translation of qualitative chemical structures into numerical graph invariants overcomes this problem and facilitates development of relationship between quantified chemical structures (using graph invariants) and quantitative biological activity. Since a minor alteration of chemical structure results in major change in the characteristic numerical value of graph invariant, the graph invariants offer a vast potential in structure activity/property relationships.

Table 4
Relationship of *eccentric distance sum* with cavity surface areas of straight chain, branched chain and cyclic alcohols

S.No.	Compound	ξ^{DS}	csa		S.No.	Compound	ξ^{DS}	csa	
			Reported	Predicted				Reported	Predicted
1	1-butanol	134	272.1	271.5	27	3-methyl-3-hexanol	556	337.7	340.1
2	2-methyl-1-propanol	97	263.8	257.9	28	3-ethyl-3-pentanol	434	324.4	327.0
3	2-butanol	97	264.1	257.9	29	2,3-dimethyl-2-pentanol	439	323.8	327.6
4	1-pentanol	292	303.9	307.1	30	2,3-dimethyl-3-pentanol	418	321.8	325.1
5	3-methyl-1-butanol	222	291.4	294.1	31	2,4-dimethyl-2-pentanol	478	328.6	332.1
6	2-methyl-1-butanol	205	289.4	290.4	32	2,4-dimethyl-3-pentanol	455	331.7	329.5
7	2-pentanol	222	295.9	294.1	33	2,2-dimethyl-3-pentanol	439	326.1	327.6
8	3-pentanol	205	293.5	290.4	34	3-heptanol	752	357.1	356.6
9	3-methyl-2-butanol	160	284.3	279.2	35	4-heptanol	725	357.1	354.8
10	2-methyl-2-butanol	154	282.5	277.5	36	1-octanol	1560	399.4	400.6
11	1-hexanol	552	335.7	339.8	37	2,2,3-trimethyl-3-pentanol	564	335.2	340.9
12	2-hexanol	447	327.7	328.6	38	2-octanol	1342	391.0	391.1
13	3-hexanol	416	325.3	324.9	39	2-ethyl-1-hexanol	1058	371.3	376.7
14	3-methyl-3-pentanol	288	305.8	306.5	40	1-nonanol	2410	431.2	429.2
15	2-methyl-2-pentanol	326	314.3	312.5	41	2-nonanol	2104	423.2	420.0
16	2-methyl-3-pentanol	314	314.3	310.7	42	3-nonanol	2005	420.8	416.9
17	3-methyl-2-pentanol	314	311.3	310.7	43	4-nonanol	1940	420.8	414.6
18	2,3-dimethyl-2-butanol	235	301.2	296.9	44	5-nonanol	1901	420.8	413.3
19	3,3-dimethyl-1-butanol	326	307.5	312.5	45	2,6-dimethyl-4-heptanol	1410	394.0	392.6
20	3,3-dimethyl-2-butanol	235	296.7	296.7	46	3,5-dimethyl-4-heptanol	2093	379.3	385.2
21	4-methyl-1-pentanol	447	323.0	328.6	47	2,2-diethyl-1-pentanol	622	372.5	376.4
22	4-methyl-2-pentanol	342	314.9	314.9	48	7-methyl-1-octanol	2104	418.7	420.0
23	2-ethyl-1-butanol	330	308.6	313.1	49	3,5,5-trimethyl-1-hexanol	1327	376.6	390.4
24	Cyclohexanol	280	290.5	305.1	50	1-decanol	3550	463.0	456.3
25	1-heptanol	964	367.5	371.1	51	1-dodecanol	7014	527.0	508.3
26	2-methyl-2-hexanol	622	346.1	346.2					

csa = cavity surface area.

Table 5
Relationship of *eccentric distance sum* with boiling points of primary amines, secondary amines and alcohols

S.No.	Compound	ξ^{DS}	bp	
			Reported	Predicted
	Alcohols			
1	Ethanol	14	78.5	73.3
2	2-propanol	33	82.4	87.1
3	1-propanol	52	97.4	95.5
4	1-butanol	134	117.7	115.6
5	2-methyl-1-propanol	97	107.9	108.3
6	2-butanol	97	99.5	108.3
7	1-pentanol	292	137.8	135.3
8	3-methyl-1-butanol	222	131.2	128.0
9	2-methyl-1-butanol	205	128.7	125.9
10	2-pentanol	222	119.0	128.0
11	3-pentanol	205	115.3	125.9
12	3-methyl-2-butanol	160	111.5	119.8
13	2-methyl-2-butanol	154	102.0	118.9
14	1-hexanol	552	157.0	153.8
15	2-hexanol	447	139.9	147.4
16	3-hexanol	416	135.4	145.3
17	3-methyl-3-pentanol	288	122.4	134.9
18	2-methyl-2-pentanol	326	121.4	138.3
19	2-methyl-3-pentanol	314	126.5	137.3
20	3 -methyl-2-pentanol	314	134.2	137.3
21	2,3-dimethyl-2-butanol	235	118.6	129.6
22	3,3-dimethyl-1-butanol	326	143.0	138.3
23	3,3-dimethyl-2-butanol	235	120.0	129.5
24	4-methyl-1-pentanol	447	151.8	147.4
25	4-methyl-2-pentanol	342	131.7	139.6
26	2-ethyl-1-butanol	330	146.5	138.6
27	Cyclohexanol	280	161.0	134.1
28	1-heptanol	964	176.3	172.1
29	2-methyl-2-hexanol	622	142.5	157.5
30	3-methyl-3-hexanol	556	142.4	154.0
31	3-ethyl-3-pentanol	434	142.5	146.5
32	2,3-dimethyl-2-pentanol	439	139.7	146.8
33	2,3-dimethyl-3-pentanol	418	139.0	145.4
34	2,4-dimethyl-2-pentanol	478	133.0	149.4
35	2,4-dimethyl-3-pentanol	455	138.8	147.9
36	2,2-dimethyl-3-pentanol	439	136.0	146.8
37	3-heptanol	752	156.8	163.7
38	4-heptanol	725	155.0	162.5
39	1-octanol	1560	195.2	189.6
40	2,2,3-trimethyl-3-pentanol	564	152.5	154.5
41	2-octanol	1342	179.8	183.9
42	2-ethyl-1-hexanol	1058	184.6	175.3
43	1-nonanol	2410	213.1	207.0
44	2-nonanol	2104	198.5	201.4

Table 5 (*Continued*)

S.No.	Compound	ξ^{DS}	bp	
			Reported	Predicted
45	3-nonanol	2005	194.7	199.5
46	4-nonanol	1940	193.0	198.1
47	5-nonanol	1901	195.1	197.3
48	2,6-dimethyl-4-heptanol	1410	178.0	184.8
49	3,5-dimethyl-4-heptanol	2093	187.0	180.4
50	2,2-diethyl-1-pentanol	622	192.0	175.2
51	7-methyl-1-octanol	2104	206.0	201.4
52	3,5,5-trimethyl-1-hexanol	1327	193.0	183.5
53	1-dodecanol	7014	230.2	356.7
54	Cyclopentanol	133	140.8	115.4
55	Cycloheptanol	417	185.0	145.3
56	1-ethylcyclohexanol	674	166.0	160.1
57	2-ethylcyclohexanol	689	181.0	160.8
58	1 -methylcyclohexanol	405	155.0	144.5
59	2-methylcyclohexanol	428	165.0	146.1
60	3-methylcyclohexanol	434	174.5	146.5
61	4-methylcyclohexanol	480	173.5	149.5
62	1,3,5-trimethylcyclohexanol	817	181.0	166.4
	Primary amines			
63	1-propylamine	52	49	65.6 (62.1)
64	2-propylamine	33	33	63.3 (60.0)
65	2-methyl-2-propylamine	60	46	66.5 (63.0)
66	2-butylamine	97	63	70.9 (67.1)
67	2-methylpropylamine	97	69	70.9 (67.1)
68	1-butylamine	134	77	75.1 (71.0)
69	2-methyl-2-butylamine	154	78	77.4 (73.1)
70	2-pentylamine	222	92	85.1 (80.3)
71	3-methylbutylamine	222	96	85.1 (80.3)
72	2-methylbutylamine	205	96	83.2 (78.5)
73	1-pentylamine	292	104	92.7 (87.4)
74	4-methylpentylamine	447	125	109.0 (102.4)
75	1-hexylamine	552	130	119.6 (112.1)
76	3-methylpentylamine	416	114	105.9 (99.5)
77	4-heptylamine	725	139	135.9 (127.1)
78	2-heptylamine	801	142	142.7 (133.3)
79	1-heptylamine	964	155	156.5 (145.8)
80	1-octylamine	1560	180	197.9 (182.4)
81	1-nonylamine	2410	201	232.4 (210.1)
82	2-undecylamine	4565	237	190.4 (150.9)
83	3-pentylamine	205	91	83.2 (78.5)
	Secondary amines			
84	N-(methyl) ethylamine	52	36	40.6 (62.1)
85	N-methyl-1-methylethylamine	97	50	52.1 (67.1)
86	Diethylamine	134	56	59.3 (71.0)
87	N-methyl-1-methylpropylamine	205	78.5	70.3 (78.5)
88	N-(ethyl) propylamine	292	80.5	81.0 (87.4)

Table 5 (*Continued*)

S.No.	Compound	ξ^{DS}	bp	
			Reported	Predicted
89	Bis(1-methylethyl)amine	342	84	86.3 (92.3)
90	N-(methyl)butylamine	292	90.5	81.0 (87.4)
91	N-methyl-1-methylbutylamine	416	105	93.4 (99.5)
92	Dipropylamine	552	109.5	104.6 (112.1)
93	Bis(2-methylpropyl)amine	1122	139	139.0 (156.9)
94	Dibutylamine	1560	159	158.6 (182.4)
95	Bis(3-methylbutyl)amine	2776	187.5	199.9 (213.2)
96	Dipentylamine	3550	205	220.6 (201.9)

bp = boiling point.

A novel distance based graph invariant termed as *eccentric distance sum* was conceptualized in the present investigation. *Eccentric distance sum* is highly discriminating and can be easily calculated from the distance matrix. The consideration of both distance sum and eccentricity of the vertices results in significant changes in the graph invariant value with a minor change in the branching of a molecule. As evidenced by Fig. 1, the *eccentric distance sum* value changes by more than twice (from 134 to 60) following branching of five membered linear carbon structure whereas Wiener's index value changes by only 1.25 times (from 20 to 16) for the identical changes. Thus, the proposed novel index is about 1.8 times more sensitive to the changes in the molecular structure when compared with Wiener's topological index.

Structure activity relationship of *eccentric distance sum* was investigated with regard to anti-HIV activity for the treatment of AIDS in a data set containing analogues of dihydroseselin. It is reported that plant constituents, suksdorfin (S.No. 2) isolated from *Lomatium suksdorfii* and related analogues of dihydroseselin act as anti-HIV agents without inhibiting HIV reverse transcriptase and, therefore, have potential for use in combination with reverse transcriptase inhibitors for AIDS therapy. This combination can be clinically useful to reduce the toxicities of HIV reverse transcriptase inhibitors and development of drug resistant virus.

The study on structure activity relationship of dihydroseselins with eccentric distance sum revealed the following information:

- A total of 26 analogues were classified into active and inactive ranges. Out of these analogues 88% were classified correctly with respect to anti-HIV activity.
- The bracketing of active range by transitional ranges indicated a gradual diminution in anti-HIV activity towards inactive range as evident from the analogues showing no suppression. A total of 22 analogues were present in the transitional range.

Table 6
Relationship of *eccentric distance sum* with molar refractivity of primary amines, secondary amines and alcohols

S.No.	Compound	ξ^{DS}	mr		S.No.	Compound	ξ^{DS}	mr	
			Reported	Predicted				Reported	Predicted
1	2-propanol	33	17.705	18.506	26	1-propylamine	52	19.400	20.455
2	2-methyl-1-propanol	97	22.103	23.463	27	1-butylamine	134	24.079	25.193
3	3-methyl-1-butanol	222	26.904	28.153	28	3 -methylbutylamine	222	28.672	28.153
4	2-methyl-1-butanol	205	26.697	27.664	29	1-pentylamine	292	28.727	29.904
5	2-pentanol	222	26.680	28.153	30	1-hexylamine	552	33.290	34.403
6	3-pentanol	205	26.639	27.664	31	2-heptylamine	801	38.037	37.341
7	2-methyl-2-butanol	154	26.721	25.976	32	1-heptylamine	964	38.003	38.895
8	1-hexanol	552	31.428	34.403	33	1-nonylamine	2410	47.277	47.586
9	3-methyl-3-pentanol	288	31.182	29.813	34	3-pentylamine	205	28.617	27.664
10	2-methyl-2-pentanol	326	31.210	30.638	35	Butyldimethylamine	447	33.816	32.842
11	2-methyl-3-pentanol	314	31.138	30.386	36	Methyl-2-methylpropylamine	222	33.852	28.153
12	4-methyl-1-pentanol	447	31.489	32.842	37	Dimethylpentylamine	801	38.281	37.341
13	4-methyl-2-pentanol	342	31.355	30.963	38	Triethylamine	330	33.793	30.720
14	2-ethyl-1-butanol	330	31.180	30.720	39	Trimethylamine	33	19.594	18.506
15	1-heptanol	964	36.093	38.895	40	Tripropylamine	1386	47.783	42.131
16	3-ethyl-3-pentanol	434	35.821	32.629	41	Butyl methyl ether	292	27.020	29.904
17	1-octanol	1560	40.637	43.242	42	Dibutyl ether	1560	40.987	43.242
18	2-ethyl-1-hexanol	1058	40.625	39.700	43	Dipropyl ether	552	32.226	34.403
19	2-methyl-1-pentanol	416	31.164	32.327	44	Ethyl 1-methylethyl ether	222	27.678	28.153
20	2,2-dimethyl-1-butanol	288	31.266	29.813	45	Ethyl pentyl ether	964	36.363	38.895
21	2-methyl-1-hexanol	752	35.930	36.826	46	1-methylpropyl ethyl ether	416	31.560	32.327
22	4-ethyl-4-heptanol	1210	44.919	40.890	47	Butyl 1-methylethyl ether	801	36.027	37.341
23	6-methyl-1-heptanol	1342	40.736	41.833	48	1-methylpropyl methyl ether	205	31.337	27.664
24	3-methyl-3-heptanol	972	40.446	38.966					
25	4-methyl-4-heptanol	914	40.439	38.442					

mr = molar refractivity.

Table 7
Mathematical models for prediction of various physical properties using *eccentric distance sum* and Wiener's index [15]

Property	N	Equation	Correlation (%)	Average error (%)
Cavity surface area of alcohols	51	$\mathrm{csa} = 124.900(\xi^{DS})\text{^}0.158$	99.63	0.052
		$\ln(\mathrm{csa}) = 5.229 + 0.144\ln(W)$	97.00	4.960
Boiling point of alcohols	62	$\mathrm{bp} = 43.006(\xi^{DS})\text{^}0.202$	92.97	0.378
		$\ln(\mathrm{bp}) = 4.279 + 0.181\ln(W)$	92.00	10.070
Boiling points of primary amines	21	$\mathrm{bp} = -2\mathrm{E}{-}05(\xi^{DS})\text{^}2 + 0.120(\xi^{DS}) + 59.406$	95.94	11.846
		$\ln(\mathrm{bp}) = 1.275 + \ln(W) - 0.006$	93.00	12.030
Boiling points of secondary amines	13	$\mathrm{bp} = 8.317(\xi^{DS})\text{^}0.401$	98.90	0.694
		$\ln(\mathrm{bp}) = 1.081 + \ln(W) - 0.006$	95.00	10.470
Boiling points of combined primary and secondary amines	34	$\mathrm{bp} = -2\mathrm{E}{-}05(\xi^{DS})\text{^}2 + 0.112(\xi^{DS}) + 56.375$	95.38	1.555
		$\ln(\mathrm{bp}) = 1.210 + \ln(W) - 0.006$	92.00	12.960
Molar refraction of heterogeneous compounds	48	$\mathrm{mr} = 8.5723(\xi^{DS})\text{^}0.220$	96.00	0.176
		$\ln(\mathrm{mr}) = 0.826 + 0.690\ln(W)$	96.00	20.660

- The active range for anti-HIV activity had *eccentric distance sum* value of 44750–75000. About 78% of analogues in the active range exhibited anti-HIV activity. The average EC_{50} of correctly predicted analogues in the active range was found to be 10.728 μM.
- Transitional ranges should ideally bracket the active range that should be subsequently bracketed by inactive ranges. However, upper inactive range was not observed in this case. In the lower inactive range 12 out of 16 correctly classified analogues did not show any suppression, while remaining were not potentially active.

One of the limitations of the graph invariant is that it can not identify optical and configurational isomers (S.No. 12 and 21, 26 and 27 in the active range). However, it is noteworthy that despite these limitations all analogues (S.No. 4, 5, 7) were correctly identified as inactive in the instant case. Similarly other analogues (S.No. 2 and 36, 14 and 15, 16 and 17, 38 and 39) were correctly identified in the transitional range.

Similarly, the study on structure activity relationship of dihydroseselins with Wiener's index revealed the following information:

- A total of 27 analogues were classified into active and inactive ranges. Out of these about 88% were classified correctly with respect to anti-HIV activity.
- The active range for anti-HIV activity had Wiener value of 2445–3500. The average EC_{50} of correctly predicted analogues was found to be 10.728 μM.
- Transitional ranges should ideally bracket the active range that should be subsequently bracketed by inactive ranges. The bracketing by lower transitional range (1600–2425) and upper transitional range (more than 3500) indicated a gradual diminution in anti-HIV activity towards inactive range. In the lower inactive range average EC_{50} of analogues was found to be 65.050 μM while upper inactive range was not observed in this case.

The quantitative structure property relationship of the *eccentric distance sum* was investigated with regard to various physical properties, for data sets consisting of primary amines, secondary amines and alcohols. Values of eccentric distance sum of all the compounds in various data sets were computed and the resultant data subjected to regression analysis. The mathematical models along with statistical analysis for various data sets and physical properties involved are compiled in Table 7. Excellent correlations were obtained using *eccentric distance sum* in all six data sets employed in present investigations. Correlation percentages ranging from 93% to more than 99% were obtained in data sets using *eccentric distance sum.* The average errors were also on the lower side (from 0.17% to 11.84%) indicating higher correlation abilities of the novel graph invariant. Correlation percentages ranging from 92% to 97% and average errors (from 4.96% to 20.66%) obtained in same data sets using *Wiener's topological*

index has been reported previously [15]. Comparatively, *eccentric distance sum* exhibited much better correlation and lesser average errors than the Wiener's topological index. The excellent prediction of the physical properties by *eccentric distance sum* can be attributed to probable contribution of distance sum in addition to eccentricity. The physical properties are significantly responsible for the biological activity of a chemical compound.

Results using *eccentric distance sum* were highly encouraging due to it's high discriminating power and excellent predictability both with regard to biological and physical properties. Though both *eccentric distance sum* and *Wiener's topological index* showed almost same predictability of anti-HIV activity of dihydroseselins but *eccentric distance sum* exhibited far superior discriminating power and correlating ability with regard to physical properties. *Eccentric distance sum* offers a vast potential for structure activity/property relationships. *Eccentric distance sum* can provide valuable leads for the development of safe and potent therapeutic agents of diverse nature.

References

[1] S.C. Basak, G.J. Niemi, G.D. Veith, Predicting properties of molecules using graph invariants, J. Math. Chem. 7 (1991) 243–272.
[2] D. Bonchev, O. Mekenyan, A.T. Balaban, Algorithms for coding chemical compounds, in: N. Trinajstic (Ed.), Mathematical and Computational Concepts in Chemistry, Ellis Horwood, Chichester, 1986, pp. 34–47.
[3] A.T. Balaban, I. Motoc, D. Bonchev, O. Mekennyan, Graph invariant indices for structure property correlation, Top. Curr. Chem. 114 (1983) 21–55.
[4] L.B. Kier, L.H. Hall, in: Molecular Connectivity in Structure Property Analysis, Research Studies Press, Letchworth, 1986, pp. 1–257.
[5] A.R. Katritzky, E.V. Gordeeva, Traditional graph invariant indices versus electronic, geometric and combined molecular graph invariants in QSAR/QSPR research, J. Chem. Inf. Comput. Sci. 33 (1993) 835–857.
[6] S.C. Basak, S. Bertlsen, G.D. Grunwold, Applications of graph theoretical parameters in quantifying molecular similarity and structure–property relationships, J. Chem. Inf. Comput. Sci. 34 (1994) 270–276.
[7] E. Estrada, A. Ramirez, Edge adjacency relationships and molecular topographic graph invariants: Definition and QSAR applications, J. Chem. Inf. Comput. Sci. 36 (1996) 837–843.
[8] 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. (2000), in press.
[9] E. De Clercq, Chemotherapeutic approaches to the treatment of acquired immune deficiency syndrome (AIDS), J. Med. Chem. 29 (1986) 1561–1569.
[10] H. Wiener, Structural determination of paraffin boiling points, J. Am. Chem. Soc. 69 (1947) 17–20.
[11] L. Huang, Y. Kashiwada, L.M. Cosentino, S. Fan, C.H. Chen, A.T. McPhail, T. Fujioka, K. Mihashi, K.H. Lee, Anti-AIDS agents. 15. Synthesis and anti-HIV activity of dihydroseselins and related analogs, J. Med. Chem. 37 (1994) 3947–3955.
[12] L.B. Kier, L.H. Hall, W.J. Murray, M. Randic, Molecular connectivity I: Relationship to nonspecific local anesthesia, J. Pharm. Sci. 64 (1975) 1971–1974.

[13] L.H. Hall, L.B. Kier, W.J. Murray, Molecular connectivity II: Relationship to water solubility and boiling point, J. Pharm. Sci. 64 (1975) 1974–1977.
[14] L.B. Kier, L.H. Hall, Molecular connectivity VII. Specific treatment of heteroatoms, J. Pharm. Sci. 65 (1976) 1806–1809.
[15] 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.