ELSEVIER

Connective eccentricity index: A novel topological descriptor for predicting biological activity

Sunil Gupta,* Manjeet Singh,* and Anil K. Madan†

**Department of Pharmaceutical Sciences and Drug Research, Punjabi University, Patiala, India*
†Department of Pharmaceutical Sciences, M.D. University, Rohtak, India

A simple, adjacency-cum-path length based, topological descriptor termed the connective eccentricity index *has been conceptualized and its discriminating power investigated with regard to antihypertensive activity. A data set consisting of 81 derivatives of N-benzylimidazole was selected for the present investigation. These derivatives are potent, competitive, and nonpeptide angiotensin II receptor antagonists. The value of* connective eccentricity index *for each derivative was computed and active range identified. Subsequently, each derivative was assigned a biological activity that was compared with the reported antihypertensive activity. The results obtained using* connective eccentricity index *were better than the corresponding values obtained using Balaban's mean square distance index. The accuracy of prediction was found to be about 80% in the active range using* connective eccentricity index. *© 2000 by Elsevier Science Inc.*

Keywords: structure-activity relationships, topological descriptor, connective eccentricity index, angiotensin II receptor antagonists, antihypertensive activity

INTRODUCTION

Every year, new chemical entities are increasingly being synthesized worldwide and await their turn for societal use. However, not all compounds can be investigated experimentally at the rate at which they are synthesized; therefore, there is a need for reliable prediction of physical, chemical, and biological properties. Structure-activity relationship (SAR) models are the method of choice in such cases. The inherent problem of SAR is quantifying chemical structures. When applied to SAR, graph theory essentially involves translation of chemical structures into characteristic numerical descriptors.[1–5]

Corresponding author: A.K. Madan, Maharishi Dayanand University, Faculty of Pharmaceutical Sciences, Rohtak 124 000, India. Tel.: +91 11 461 8365; fax: +91 11 462 3649.

Graph theory has been of great help in constructing various molecular descriptors called topological indices. These indices can be derived from the topological distance of the molecule, such as the Wiener index,[6] or they may be based on molecular connectivity, i.e., indices of Randic,[7] distance-connectivity index reported by Balaban,[8] or information theoretic indices of Bonchev.[9] Molecular topology as represented by the connectivity of atoms can relate biological activity with the molecular graph or subgraph of the compounds. Because the three-dimensional structure of a compound depends on the connectivity of its constituent atoms, the numerical descriptors derived from information based on connectivity can reveal structural or substructural information about a molecule. Topological indices developed for predicting physicochemical properties and biological activities of chemical substances can be used for drug design.[10–16]

In the present study, a novel, adjacency-cum-path length based, topological descriptor termed the *connective eccentricity index* is presented. *Connective eccentricity index* (C^{ξ}) can be defined as the summation of the ratios of the degree of a vertex (Vi) and its eccentricity (Ei) for all vertices in the hydrogen suppressed molecular structure. It can be expressed by the following equation:

$$C^{\xi} = \sum_{i=1}^{n} (Vi/Ei).$$

The eccentricity *Ei* of a vertex *i* in a graph *G* is the path length from vertex *i* to the vertex *j* that is farthest from *i* (Ei = max d(ij); $j \in G$). Calculation of the *connective eccentricity index* values for three isomers of pentane is illustrated in Figure 1.

In order to explore the potential of *connective eccentricity index* in predicting biological activity, a relationship between antihypertensive activity and connective eccentricity index was investigated. The present study was carried out using nonpeptide N-benzylimidazole derivatives (Figure 2). They are reported to be potent, competitive, and nonpeptide angiotensin II (AII) receptor antagonists.[17] The predictability of *connective eccentricity index* with respect to antihypertensive activity sub-

Journal of Molecular Graphics and Modelling 18, 18–25, 2000

655 Avenue of the Americas, New York, NY 10010

1093-3263/00/$–see front matter
PII S1093-3263(00)00027-9

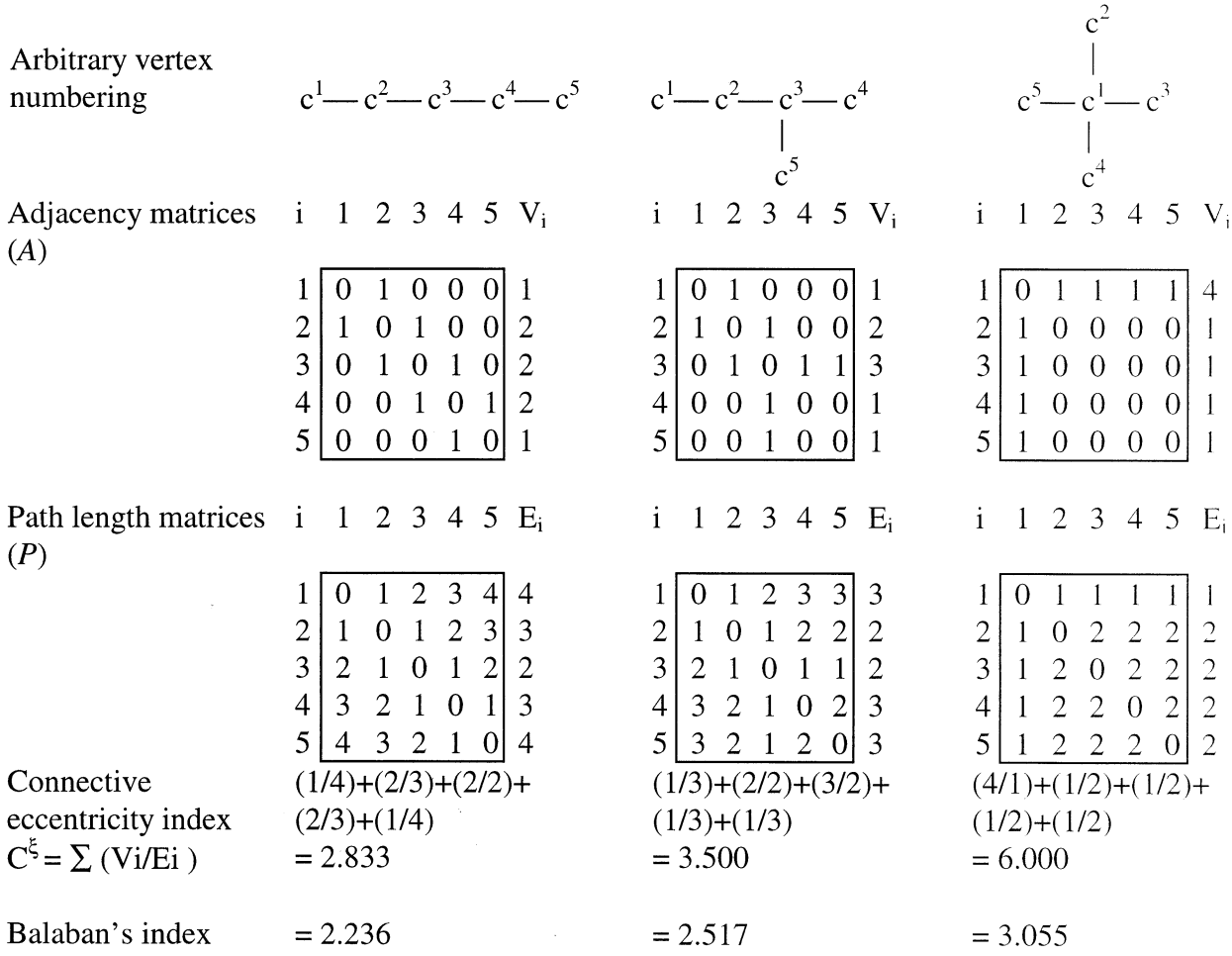

Figure 1. Calculation of connective eccentricity index *values for three isomers of pentane.*

sequently was compared with the predictability of Balaban's mean square distance index using the same data set.

Balaban[18] proposed a mean square distance topological index based on the distribution of distance. This index, termed Balaban's mean square distance index (D), can be expressed by the following equation:

$$D = (\Sigma g_i(i)^2/\Sigma i)^{1/2}$$

where g_i is number of occurrences of path lengths (i).

METHODOLOGY

A data set comprising 81 derivatives of N-benzylimidazole was selected for the present study. This data set consisted of both active and inactive derivatives. The values of *connective eccentricity index* of each derivative in the data set were computed using an in-house computer program. The resultant data were analyzed and active range identified.

Subsequently, each derivative was assigned a biological activity that was compared with the reported antihypertensive activities of N-benzylimidazole derivatives.[17] The antihypertensive activity was reported in terms of specific binding to AII receptors. IC_{50} was defined as inhibitory concentration of a derivative required for 50% displacement of the labeled AII (2 nM) from AII receptors. Derivatives exhibiting IC_{50} <6.0 μmol were considered to be active for the purposes of present study.

The percent degree of classification for each range was calculated from the ratio of the number of derivatives with correctly predicted activity to that of the number of derivatives present in the respective range. The overall degree of classification was obtained from the ratio of total number of derivatives with correctly predicted activity to that of total number of derivatives present in both the active and inactive ranges.

The values of Balaban's mean square distance index were computed for all the derivatives using hydrogen suppressed structures. The aforementioned procedure was similarly used to identify the active range and to calculate degree of classification. The results are summarized in Tables 1–3 and Figure 3.

RESULTS AND DISCUSSION

A novel adjacency-cum-path length based index termed the *connective eccentricity index* has been conceptualized in the present investigation. *Connective eccentricity index* takes into consideration the eccentricity as well as degree of all vertices in the graph.

In order to explore the utility of *connective eccentricity index* in structure-activity studies, a relationship between connective eccentricity index and antihypertensive activity was

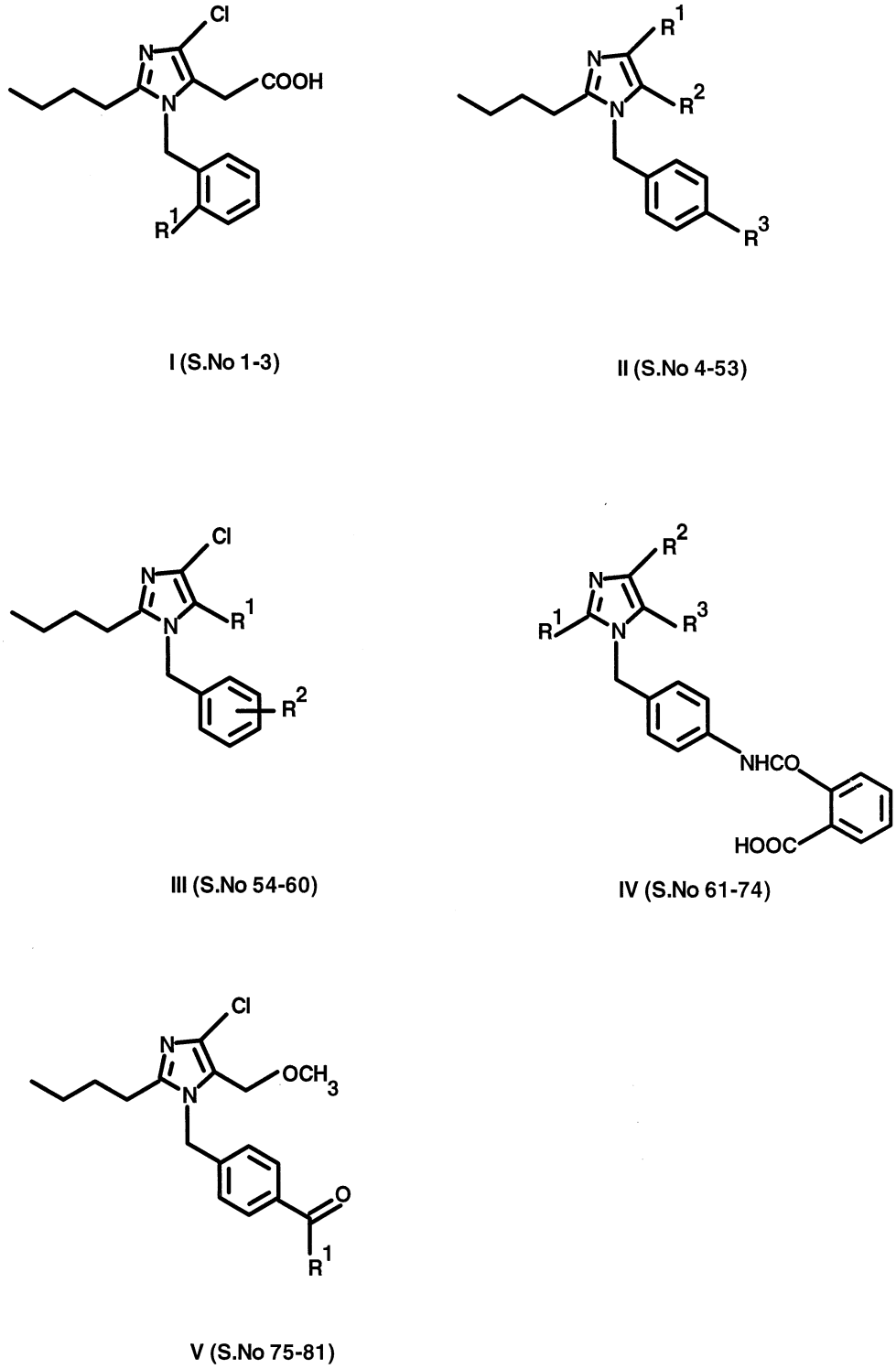

Figure 2. Parent structures of various N-benzylimidazoles.

investigated. The renin-angiotensin system is a major regulator of cardiovascular and renal function. In the treatment of hypertension, angiotensin-converting enzyme (ACE) inhibitors act by preventing the enzymatic cleavage of angiotensin I and consequently interfere with the formation of the powerful vasoconstrictor octapeptide AII. However, in addition to blocking the renin-angiotensin system, ACE inhibitors disrupt bradykinin, tachykinin, and prostaglandin biosynthesis. AII receptor antagonists overcome the side effects associated with ACE inhibitors as a result of their increased specificity and selectivity in blocking the effects of AII at the receptor level. AII receptor antagonists to date have been peptides, such as saralasin. They show antihypertensive action but have short plasma half-lives and are ineffective when administered orally. The

Table 1. Relationship of *connective eccentricity index* (C^{ξ}) and Balaban's mean square distance index (D) with antihypertensive activity of N-benzylimidazole derivatives

S. No.	BS	R^1	R^2	R^3	C^{ξ}	D	Antihypertensive activity: Assigned C^{ξ}	Assigned D	Reported
1	I	NO_2			6.745	11.897	−	−	−
2	I	Cl			6.323	10.386	−	−	−
3	I	H			6.087	9.771	+	−	−
4	II	Cl	CH_2COOH	COOH	5.671	10.875	±	−	+
5	II	Cl	CH_2COOCH_3	NH_2	5.928	10.632	−	−	
6	II	CH_2COOH	Cl	COOH	5.618	11.167	−	−	−
7	II	Cl	CH_2OH	COOH	5.289	9.540	−	−	+
8	II	CH_2OH	Cl	COOH	5.269	9.643	−	−	−
9	II	Cl	CH_2OCOCH_3	COOH	5.829	11.819	+	−	+
10	II	Cl	CH_2COOH	CH_2COOH	5.442	11.165	−	−	−
11	II	Cl	CH_2COOCH_3	NO_2	5.846	11.658	+	−	−
12	II	Cl	CH_2COOCH_3	$NHCH_2Ph$	5.386	13.489	−	±	−
13	II	Cl	CH_2OH	CHO	5.115	8.707	−	−	−
14	II	Cl	CH_2OH	CH=NOH	4.933	9.015	−	−	−
15	II	Cl	CH_2OH	OCH_3	5.115	8.707	−	−	−
16	II	Cl	CH_2COOCH_3	diphenylmaleimidyl	6.941	19.513	−	±	−
17	II	Cl	CH_2COOCH_3	$NHCO(CH_2)_2COOH$	5.260	13.531	−	±	−
18	II	Cl	CH_2COOCH_3	$NHCO(CH_2)_3COOH$	5.099	13.878	−	±	−
19	II	Cl	CH_2COOCH_3	NHCOCH=CHCOOH(cis)	5.260	13.531	−	±	−
20	II	Cl	CH_2OCH_3	$NHCOCH_2CH(Ph)COOH$	5.386	15.259	−	±	+
21	II	Cl	CH_2OCH_3	$NHCOCH(Ph)CH_2COOH$	5.661	15.563	±	±	−
22	II	Cl	CH_2COOCH_3	2-carboxybenzamido	5.943	16.394	+	+	+
23	II	CH_2COOCH_3	Cl	2-carboxybenzamido	5.656	15.820	±	+	+
24	II	Cl	CH_2OCH_3	2-carboxybenzamido	5.676	14.930	±	±	+
25	II	Cl	CH_2COOCH_3	4-phthalimido	6.330	15.557	−	±	+
26	II	Cl	CH_2COOCH_3	3-carboxy-2-napthylamido	6.308	18.210	−	+	+
27	II	Cl	CH_2COOCH_3	2-carboxytetrafluorobenzamido	6.186	18.187	+	+	+
28	II	Cl	CH_2COOCH_3	2-carboxy-4,5-dimethylbenzamido	5.790	16.291	+	+	+
29	II	Cl	CH_2COOCH_3	2-carboxy-4,5-dinitrobenzamido	5.787	17.117	±	+	−
30	II	Cl	CH_2COOCH_3	2-carboxy-4,5-nitrobenzamido	6.040	17.747	+	+	+
31	II	Cl	CH_2COOCH_3	2-carboxy-6-methoxybenzamido	6.210	17.792	+	+	+
32	II	Cl	CH_2COOCH_3	2-carboxy-6-methylbenzamido	6.081	17.044	+	+	+
33	II	Cl	CH_2COOCH_3	2-carboxy-3,6-dichlorobenzamido	6.210	17.778	+	+	+
34	II	Cl	CH_2OCH_3	2-carboxy-3,6-dichlorobenzamido	5.943	16.235	+	+	+

(*Continued*)

Table 1. (*Continued*)

S. No.	BS	R^1	R^2	R^3	C^{ξ}	D	Antihypertensive activity: Assigned C^{ξ}	Assigned D	Reported
35	II	Cl	CH_2OCH_3	2-carboxy-3,6-dichloro-N-methyl benzamido	6.117	16.593	+	+	+
36	II	Cl	CH_2COOCH_3	2-(2′-carboxyphenyl)benzamido	6.258	19.239	+	+	+
37	II	Cl	CH_2COOCH_3	$NHSO_2CF_3$	5.893	14.478	+	±	+
38	II	Cl	CH_2COOCH_3	$NHCOCF_3$	5.733	13.813	±	±	−
39	II	Cl	(1-H-tetrazole-5-yl)methyl	(1-H-tetrazole-5-yl)	6.244	13.340	+	±	+
40	II	Cl	CH_2COOCH_3	2-(trifluoromethanesulfonamido) benzamido	6.015	18.687	+	+	+
41	II	Cl	CH_2COOCH_3	N-(phenylmethyl)-N-[2-(trifluoro methanesulfonamido)benzoyl]amino	7.082	21.341	−	±	+
42	II	Cl	CH_2COOCH_3	2-(trifluoromethanesulfonamido) benzamido	5.903	17.215	+	+	+
43	II	Cl	CH_2COOCH_3	5-chloro-2-(trifluoromethyl sulfonamido) benzamido	6.144	19.275	+	+	+
44	II	Cl	CH_2COOCH_3	5-bromo-2-(trifluoromethyl sulfonamido)benzamido	6.144	19.275	+	+	−
45	II	Cl	CH_2COOCH_3	5-iodo-2-(trifluoromethyl sulfonamido)benzamido	6.144	19.275	+	+	+
46	II	Cl	CH_2COOCH_3	5-methyl-2-(trifluoromethyl sulfonamido)benzamido	6.144	19.275	+	+	+
47	II	Cl	CH_2COOCH_3	4-nitrobenzamido	5.390	15.305	−	±	−
48	II	Cl	CH_2COOCH_3	4-chlorbenzamido	5.393	14.177	−	±	−
49	II	Cl	CH_2COOCH_3	3-methyl-2-(trifluoromethane sulfonamido)benzamido	6.144	19.245	+	+	+
50	II	Cl	CH_2COOCH_3	trans-2-(trifluoromethane sulfonamido)cyclohexane carboxamiod	6.015	18.687	+	+	+
51	II	Cl	CH_2COOCH_3	cis-2-(trifluoromethanesulfonamido) cyclohexane carboxamido	6.015	18.687	+	+	+
52	II	Cl	CH_2COOCH_3	2-hydroxy-3,5-dinitrobenzamido	6.154	18.608	+	+	−
53	II	Cl	CH_2COOCH_3	2-amidobenzenesulphonic acid	5.269	14.803	−	±	+
54	III	CH_2COOH	3-COOH		6.139	11.322	+	−	−

(*Continued*)

Table 1. (*Continued*)

S. No.	BS	R^1	R^2	R^3	C^{ξ}	D	Antihypertensive activity: Assigned C^{ξ}	Assigned D	Reported
55	III	CH_2COOH	2-COOH		6.745	11.89	−	−	−
56	III	(1-H-tetrazole-5-yl) methyl	3-(1-H-tetrazole-5-yl)		6.755	13.638	−	±	+
57	III	CH_2COOCH_3	4-[3-(trifluoromethane sulfonamido)benzamido]		6.068	18.386	+	+	+
58	III	CH_2COOCH_3	4-[4-(trifluoromethane sulfonamido)benzamido		5.457	18.273	−	+	−
59	III	CH_2COOCH_3	2-[2-carboxybenzamido]		6.854	16.435	−	+	−
60	III	CH_2COOCH_3	2-[2-(trifluormethane sulfonamido)benzamido]		6.844	18.580	−	+	−
61	IV	H	H	H	5.276	11.251	−	−	−
62	IV	CH_3	H	H	5.436	12.075	−	−	−
63	IV	CH_3CH_2	H	H	5.329	12.112	−	−	−
64	IV	$CH_3(CH_2)_2$	H	H	5.184	12.240	−	±	+
65	IV	$CH_3(CH_2)_3$	H	H	5.081	12.439	−	±	+
66	IV	$CH_3(CH_2)_4$	H	H	4.951	12.694	−	±	+
67	IV	$CH_3(CH_2)_5$	H	H	4.854	12.993	−	±	+
68	IV	$CH_3(CH_2)_6$	H	H	4.748	13.327	−	±	+
69	IV	$PhCH_2CH_2$	H	H	5.218	14.424	−	±	−
70	IV	4-CH_3OPhCH_2	H	H	5.136	14.410	−	±	+
71	IV	cyclohexyl-CH_2	H	H	5.339	14.078	−	±	+
72	IV	$(CH_3)_2CH$	H	H	5.477	12.926	−	±	−
73	IV	$Ph(CH_2)_3$	H	H	5.548	16.109	−	+	−
74	IV	H	H	CH_2OH	5.329	12.112	−	−	−
75	V	2-carboxyphenyl amino			5.690	14.871	±	±	−
76	V	3-methyl-2-carboxy phenylamino			5.819	15.591	+	+	+
77	V	2-(1-H-tetrazole-5-yl) phenylamino			5.792	15.87	+	+	+
78	V	L-Phe			5.428	14.990	−	±	−
79	V	D-Phe			5.428	14.990	−	±	−
80	V	L-Pro			5.727	13.682	±	±	−
81	V	D-pro			5.727	13.682	±	±	−

BS = basic structure.

Table 2. Relationship between antihypertensive activity and connective eccentricity index (C^{ξ})

Range	C^{ξ} Index Value	Classified derivatives		Accuracy (%)	Average IC_{50} (μmol)
		Total	Correct		
Lower inactive	<5.65	32	22	68.75	32.744 (46.909)
Transitional	5.65–5.79	9	NA	NA	19.789 (NA)
Active	5.79–6.32	31	25	80.65	18.259 (1.842)
Upper inactive	≥6.32	9	6	66.67	34.733 (50.500)

Values in parentheses are based on correctly classified compounds in the particular range.

present study was carried out using nonpeptide derivatives. All 81 nonpeptide N-benzylimidazole derivatives whose IC_{50} values were reported by Dunica et al.[17] constituted the data set for the present study. These derivatives, which are nonpeptide, potent, and competitive AII receptor antagonists, are reported to constitute a novel class of antihypertensives.

Using *connective eccentricity index*, the biologically active derivatives were found to be present in the narrow range of index values, thus facilitating rapid identification of the active range. The results revealed the following information:

- A total of 72 derivatives were classified into active and inactive ranges, of which 53 were correctly classified with respect to antihypertensive activity.
- The presence of a transitional range indicated a gradual change in antihypertensive activity from inactive to active range (Figure 3). A total of nine derivatives were present in the transitional range.
- The active range for antihypertensive activity had connective eccentricity index value of 5.79–6.32. As much as 81% of these derivatives exhibited antihypertensive activity. The average IC_{50} of correctly classified derivatives was found to be 1.842 μmol.

Similarly, the study using Balaban's mean square distance index revealed the following information:

- A total of 50 derivatives were classified into active and inactive ranges, of which 40 were correctly classified, resulting in about 80% accuracy with regard to antihypertensive activity. However, the upper inactive range was not observed.
- Lower and upper transitional ranges were observed, indicating a gradual change in antihypertensive activity from the inactive to active range. A total of 31 derivatives were present in the transitional range. Derivatives present in the transitional ranges were not considered while determining the overall accuracy of prediction.
- The active range for antihypertensive activity had Balaban's mean square distance index value of 15.58–19.50. About 77% of the derivatives in the active range exhibited antihypertensive activity. The average IC_{50} of correctly classified derivatives was 1.646 μmol.

Comparison of the results obtained by *connective eccentricity index* and Balaban's mean square distance index disclosed the following information:

- The discriminating power of connective eccentricity index is better than that of Balaban's mean square distance index, as evidenced by the case of three isomers of pentane. Figure 1 shows that the relative change in the value of the connective eccentricity index with the change in molecular structure is almost double the corresponding change in the value of Balaban's mean square distance index.
- The gradual movement from the inactive to active range resulted in a steep increase in the potency of the derivatives for both connective eccentricity index and Balaban's mean square distance index. The average potency in the active range obtained by connective eccentricity index (1.864 μmol) is comparable to the corresponding value obtained by Balaban's mean square distance index (1.646 μmol).
- In a data set comprising 81 derivatives of N-benzylimidazole, 88% of derivatives were classified into active and inactive ranges using connective eccentricity index compared to 62% using Balaban's mean square distance index. Connective eccentricity index is more sensitive to changes in molecular structure compared to Balaban's index because connective

Table 3. Relationship between antihypertensive activity and Balaban's mean square distance index (D)

Range	D Index Value	Classified derivatives		Accuracy (%)	Average IC_{50} (μmol)
		Total	Correct		
Inactive	<12.20	20	17	85.00	61.005 (71.265)
Lower transtional	12.20–15.58	29	NA	NA	13.552 (NA)
Active	15.58–19.50	30	23	76.667	15.959 (1.646)
Upper transitional	≥19.50	2	NA	NA	6.350 (NA)

Values in parentheses are based on correctly classified compounds in the particular range.

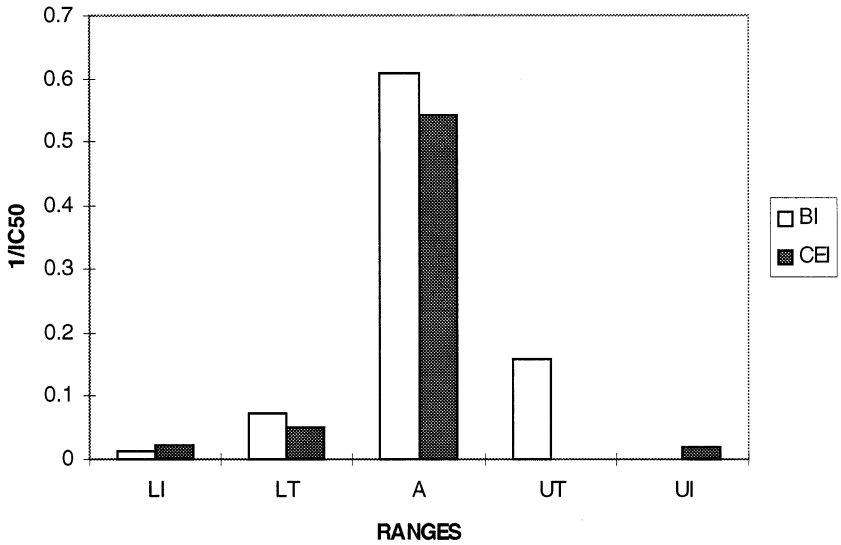

Figure 3. Average $1/IC_{50}$ values of various ranges (LI = lower inactive, LT = lower transitional, A = active, UT = upper transitional, UI = upper inactive). BI = Balaban's mean square distance index, CEI = connective eccentricity index.

eccentricity index considers the shape as well as the size of the molecule.

In the present study, *connective eccentricity index* proved its discriminating power with respect to antihypertensive activity in a data set that consisted of a relatively large number of derivatives. Therefore, *connective eccentricity index*, along with other mathematical tools, can provide leads about therapeutic agents used to treat hypertension. Moreover, the high discriminating power of *connective eccentricity index* offers a vast potential in other structure-property and structure-activity studies.

REFERENCES

1 Trinajstic, N. In *Chemical Graph Theory, Volumes I and II*. CRC Press, Boca Raton, Florida, 1983

2 Balaban, A.T., Motoc, J., Bonchev, D., and Mekennyan, O. Topological indices for structure activity correlation. *Top. Curr. Chem.* 1983, **114**, 21–55

3 Katritzky, A.R., and Gordeeva, E.V. Traditional topological indices versus electronic, geometric and combined molecular descriptors in QSAR/QSPR research. *J. Chem. Inf. Comput. Sci.* 1993, **33**, 835–857

4 Basak, S.C., Bertlsen, S., and Grunwold, G.D. Applications of graph theoretical parameters in quantifying molecular similarity and structure-activity relationships. *J. Chem. Inf. Comput. Sci.*, 1994, **34**, 270–276

5 Estrada, E., and Ramirez, A. Edge adjacency relationships and molecular topographic descriptors: Definition and QSAR applications. *J. Chem. Inf. Comput Sci.* 1996, **36**, 837–843

6 Wiener, H. Structural determination of paraffin boiling points. *J. Am. Chem. Soc.* 1947, **69**, 17–20

7 Randic, M. On characterisation of molecular branching. *J. Am. Chem. Soc.* 1975, **97**, 6609–6615

8 Balaban, A.T. Highly discriminating distance based numerical descriptor. *Chem. Phys. Lett.* 1982, **89**, 399–404

9 Bonchev, D. In *Information-Theoretic Indices for Characterisation of Chemical Structures*. Research Studies Press, Chichester, England, 1983, pp. 1–249

10 Kier, L.B., and Hall, L.H. In: *Molecular Connectivity in Chemistry and Drug Research*. Academic Press, New York, 1976, pp. 1–257

11 Sablic, A., and Trinajstic, N. Quantitative structure activity relationships: The role of topological indices. *Acta Pharm. Jugosl.* 1981, **31**, 189–214

12 Rouvery, D.H. The limits of applicability of topological indices. *J. Mol. Struct. (THEOCHEM)* 1989, **185**, 187–201

13 Galvez, J., Garcia-Domenec, R., deJulian-Ortiz, J.V., and Soler R. Topological approach to drug design. *J. Chem. Inf. Comput. Sci.* 1995, **35**, 272–284

14 Garcia-Domenec, R., deGregario-Alapont, C., deJulian-Ortiz, J.V.. Galvez, J., and Popa, L. Molecular connectivity to find β-blockers with low toxicity. *Bioorg. Med. Chem. Lett.* 1997, **7**, 567–572

15 Sharma, V., Goswami, R., and Madan, A.K. Eccentric connectivity index: A novel highly discriminating topological descriptor for structure property and structure activity studies. *J. Chem. Inf. Comput. Sci.* 1997, **37**, 273–282

16 Gupta, S., Singh, M., and Madan, A.K. Superpendentic index: A novel topological descriptor for predicting biological activity. *J. Chem. Inf. Comput. Sci.* 1999, **39**, 272–277

17 Dunica, J.V., Chiu, A.T., Carini, D.J., Gregory, G.B., Johnson, A.L., Wells, G.J., Wong, P.C., Calabrese, J.C., and Timmermans, P.B.M.W.M. The discovery of potent nonpeptide angiotensin II receptor antagonists: A new class of potent antihypertensives. *J. Med. Chem.* 1990, **33**, 1312–1329

18 Balaban, A.T. Topological indexes based on topological distances in molecular graphs. *Pure Appl. Chem.* 1983, **55**, 199–206