J Mol Model (2007) 13:137–145
DOI 10.1007/s00894-006-0134-y

ORIGINAL PAPER

Topochemical models for the prediction of voltage-gated sodium channel binding activity of hydantoins and related non-hydantoins

Meenal Gupta · Anil Kumar Madan

Received: 27 February 2006 / Accepted: 30 June 2006 / Published online: 7 September 2006

Abstract The relationship of *Wiener's topochemical index*—a distance based topochemical index, *molecular connectivity topochemical index*—an adjacency based topochemical index and *eccentric connectivity topochemical index*—an adjacency-cum-distance based topochemical index with sodium channel binding activity has been investigated. A dataset comprising 50 hydantoins and related non-hydantoins was selected. The dataset was divided equally into training and test sets. The values of the three topochemical indices for all the compounds present in both the training and test sets were computed using an in-house computer program. The resulting data was analyzed and suitable models were developed after identification of the active ranges in the training set. Subsequently, a biological activity was assigned to each compound involved in the training set using these models, which was then compared with the reported sodium channel binding activity. An accuracy of prediction of the order of >99% was observed using the proposed models. Cross-validation of these models using the test set revealed an exceptionally high accuracy of ~95%.

Keywords Topochemical indices · Wiener's topochemical index · Molecular connectivity topochemical index · Eccentric connectivity topochemical index · Hydantoins · Voltage-gated sodium channel · Sodium channel binding activity

M. Gupta · A. K. Madan (✉)
Faculty of Pharmaceutical Sciences, M. D. University,
124001 Rohtak, India
e-mail: madan_ak@yahoo.com

Introduction

An upsurge of interest in pharmaceutical drug design is the prediction of physicochemical, toxicological and pharmacological properties of chemicals directly from their structure [1]. This has become increasingly dependent on predictive quantitative structure-activity/property relationships (QSARs/QSPRs) [2]. Developing SARs/SPRs for drug compounds using computational or theoretical methods relies on an appropriate representation of the molecular structure [3]. The use of graph-theoretic techniques for the characterization of structures and for exploration of SARs/SPRs has received considerable attention [4, 5].

Molecular structures can be represented by planar graphs, $G=[V, E]$, where the non-empty set V represents the set of atoms and the set E represents covalent bonds [6]. Such a graph represents the "topology of a molecule" in the sense that it depicts the pattern of connectedness of atoms in the molecule, being, at the same time, independent of such metric aspects of molecular structure as equilibrium distance between nuclei, valence angles etc. [7, 8]. In order to study relationships between chemical structure and property, the bonding topology of a molecule is converted into an expression, which may be a matrix, a polynomial, a sequence of numbers or a numerical index [9]. Such numerical indices uniquely determined from a molecular graph are called topological (graph-theoretic) indices (TIs) [10–12]. Graph invariants are graph-theoretical properties that remain invariant for isomorphic graphs [13, 14]. TIs are derived from different classes of weighted graphs, representing various levels of chemical structural information. They are numerical quantifiers of molecular topology and encode information regarding the size, shape, branching pattern, cyclicity and symmetry of molecular graphs [1].

TIs that encode information strictly about the adjacency and distance of atoms in molecular structures are considered as topostructural (TS) indices [1, 14]. Unlike TS indices, topochemical (TC) indices take into consideration not only the presence but also the relative position of heteroatom(s), thereby reducing the degeneracy of TS indices to a large extent [15]. A number of TIs have received great attention due to their applications in QSAR/QSPR studies and drug research [16–19]. These include TS indices such as Wiener's index, [20–22] Hosoya's index, [23] Balaban's J indices, [24–26] Randic's molecular connectivity index, [27, 28] Zagreb indices M_1 and M_2, [29–31] eccentric connectivity index, [32–36] information theoretic indices [37, 38] and neighborhood complexity indices [13, 39]. Topochemical indices that have been reported and used for SAR studies include valence connectivity indices, [19, 40, 41] molecular connectivity topochemical index, [42] eccentric adjacency topochemical index, [43] eccentric connectivity topochemical index, [44] Wiener's topochemical index [45] and superadjacency topochemical index [46].

Classical work by Hodgkin and Huxley [47] defined the three key features of sodium channels-voltage-dependent activation, rapid inactivation and selective ion conductance. Building on this foundation, more recent structure-function studies using molecular, biochemical and electrophysiological techniques have provided a good understanding of the molecular basis of sodium-channel function [48]. Voltage-gated sodium channels (Na_VChs) play an essential role in the initiation and propagation of action potentials in neurons and other electrically excitable cells such as myocytes and endocrine cells [48] in concert with other voltage-dependent ion channels, which in turn triggers other physiological events leading to muscular contraction and neuronal firing. Na_VChs have long been recognized as important targets for local anesthetics, anticonvulsants and anti-arrhythmic agents [49]. All these drugs share a common preference for the inactive state of Na_VChs and a common site of interaction with α subunit, located in correspondence to the inner pore [50]. In addition to receptor sites for drugs, Na_VChs also contain at least six distinct neurotoxin-binding sites [51, 52].

The essential nature of Na_VChs is emphasized by the existence of inherited disorders (sodium "channelopathies"), which were among the first recognized ion-channel diseases and still continue to attract widespread clinical and specific interest [49]. Nearly 20 disorders affecting skeletal muscle contraction, cardiac rhythm or neuronal function and ranging in severity from mild to life-threatening or incapacitating conditions have been traced to mutations in genes encoding human Na_VChs [49]. Studies over the past decade have demonstrated that a sustained sodium-ion influx through Na_VChs can trigger reverse sodium-calcium exchange which imports damaging levels of calcium into axons [53], thereby activating injurious calcium-mediated processes [54]. Sodium-channel blockers have been shown to have a protective effect in neuro-inflammatory disorders, preventing axonal degeneration [55] and to be useful clinically in the treatment of various sodium channelopathies [49]. The sodium channelopathies provide outstanding illustrations of the delicate balances that maintain normal operation of critical physiological events such as muscle contraction and conduction of electrical signals [49]. Further, the multiplicity of sodium channels and the dynamic nature of their expression make them important targets in the search for novel therapeutics for epilepsy, cardiac arrhythmia and persistent pain syndromes [56].

In the present study, three topochemical indices—*Wiener's topochemical index* (a distance based topochemical index), *molecular connectivity topochemical index* (an adjacency based topochemical index) and *eccentric connectivity topochemical index* (an adjacency-cum-distance based topochemical index) have been used for the development of models for prediction of sodium channel binding activity of hydantoins and related non-hydantoins.

Methodology: calculations of topochemical indices

Wiener's topochemical index (W_c)

Wiener's topochemical index is the modification of the oldest and widely used distance based topostructural index—Wiener's index [20–22]. Various modifications of Wiener's index have been reported in literature. These include hyper-Wiener's index, [57] revised Wiener index, [58] new hyper-Wiener index [59] and Wiener's topochemical index [45]. Wiener's topochemical index is defined as half the sum of chemical distances between all the pairs of vertices in hydrogen suppressed molecular graph, i.e.

$$W_c = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} Pij^c$$

where, Pij^c is the chemical length of path that contains the least number of edges between vertices, i and j in graph G and n is the maximum possible number of i and j [45].

Molecular connectivity topochemical index (χ^A)

The molecular connectivity topochemical index is a modified form of one of the most widely used adjacency-based topostructural indices—the first order molecular connectivity index [27, 60]. The molecular connectivity topochemical index was originally reported in the literature as the atomic molecular connectivity index [42] and subsequently renamed the molecular connectivity topo-

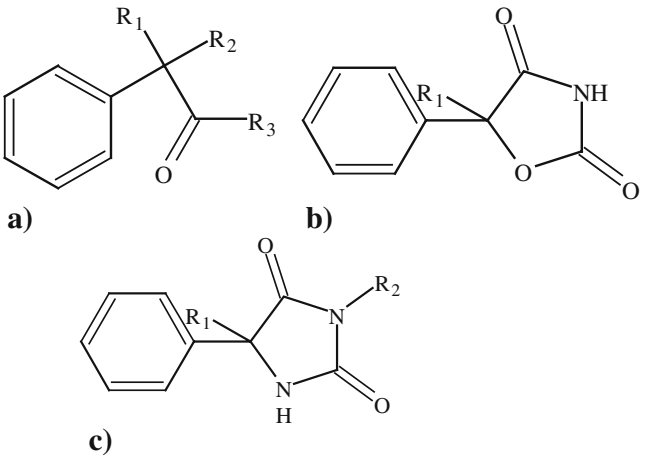

Fig. 1 Basic structures of various compounds in the dataset

chemical index for the sake of simplicity [61]. The molecular connectivity topochemical index, denoted by χ^A, is defined as the summation of the modified bond values of adjacent vertices for all edges in the hydrogen-suppressed molecular graph and is expressed as

$$\chi^A = \sum_{i=1}^{n} \left(V_i^c V_j^c \right)^{-1/2}$$

where, n is the number of vertices, V_i^c and V_j^c are the modified degrees of adjacent vertices, i and j forming the edge ij in a graph G [42].

Table 1 Structures of various compounds in the training set

Cpd. no.	Substituents		
	R_1	R_2	R_3
A2	–OH	$–C_2H_5$	$–NH_2$
A4	–OH	$–C_5H_{11}$	$–NH_2$
A6	–OH	$–C_9H_{19}$	$–NH_2$
A8	–OH	$–C_3H_7$	–OH
A10	–OH	$–C_7H_{15}$	–OH
B12	$–C_3H_7$	–	–
B14	$–C_7H_{15}$	–	–
C16	$–C_2H_5$	–H	–
C18	$–C_4H_9$	–H	–
C20	$–C_7H_{15}$	–H	–
C22	$–C_2H_5$	$–SO_2CH_3$	–
C24	$–C_5H_{11}$	$–SO_2CH_3$	–
C26	$–C_9H_{19}$	$–SO_2CH_3$	–
A28	$–NH_3^+$	$–C_5H_{11}$	$–O^-$
A30	$–NH_3^+$	$–C_9H_{19}$	$–O^-$
A32	–OH	$–C_2H_5$	$–OC_2H_5$
C34	$–C_5H_{11}$	–benzyl	–
C36	$–C_2H_5$	–4-(diphenylmethyl)-1-piperazinyl methyl	–
C38	$–C_7H_{15}$	–4-(diphenylmethyl)-1-piperazinyl methyl	–
C40	$–C_2H_5$	–4-(diphenylmethyl)-1-piperazinyl ethyl	–
A42	–OH	$–C_2H_5$	$–NH(CH_2)_2Ph$
A44	–OH	$–C_2H_5$	$–NH(CH_2)_3Ph$
A46	–OH	$–C_2H_5$	$–NH(CH_2)_3Ph_2$
A48	–OH	$–CH_3$	–*N*-diphenylmethyl piperazinyl
A50	$–(CH_3)_3Ph$	$–C_2H_5$	$–NH–(CH_3)_3–Ph$

Table 2 Structures of various compounds in the test set

Cpd. no.	Substituents		
	R_1	R_2	R_3
A1	–OH	$–CH_3$	$–NH_2$
A3	–OH	$–C_3H_7$	$–NH_2$
A5	–OH	$–C_7H_{15}$	$–NH_2$
A7	–OH	$–C_2H_5$	–OH
A9	–OH	$–C_5H_{11}$	–OH
A11	–OH	$–C_9H_{19}$	–OH
B13	$–C_5H_{11}$	–	–
B15	$–C_9H_{19}$	–	–
C17	$–C_3H_7$	–H	–
C19	$–C_5H_{11}$	–H	–
C21	$–C_9H_{19}$	–H	–
C23	$–C_3H_7$	$–SO_2CH_3$	–
C25	$–C_7H_{15}$	$–SO_2CH_3$	–
A27	$–NH_3^+$	$–C_2H_5$	$–O^-$
A29	$–NH_3^+$	$–C_7H_{15}$	$–O^-$
C31	$–C_2H_5$	$–SO_2–Ph–CH_3$	–
A33	–OH	$–C_7H_{15}$	$–OC_2H_5$
C35	$–C_9H_{19}$	–benzyl	–
C37	$–C_5H_{11}$	–4-(diphenylmethyl) - 1-piperazinyl methyl	–
C39	$–C_6H_5$	–4-(diphenylmethyl) - 1-piperazinyl methyl	–
C41	$–C_6H_5$	–4-(diphenylmethyl) - 1-piperazinyl ethyl	–
A43	–OH	$–C_5H_{11}$	$–NH(CH_2)_2Ph$
A45	–OH	$–C_5H_{11}$	$–NH(CH_2)_3Ph$
A47	–OH	$–C_5H_{11}$	$–NH(CH_2)_3Ph_2$
A49	–OH	$–C_5H_{11}$	–*N*-diphenyl methylpiperazinyl

Eccentric connectivity topochemical index (ξ_c^c)

The eccentric connectivity topochemical index is a modification of an adjacency-cum-distance based eccentric connectivity index [32–36]. It is defined as the summation of the product of chemical eccentricity and chemical degree of each vertex in a hydrogen suppressed molecular graph having n vertices, i.e. where, V_{ic} is the chemical degree of vertex, E_{ic}

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

is the chemical eccentricity of the vertex i and n is the number of the vertices in graph G [44].

Model development

A dataset comprising the 50 compounds reported by Zha et al. [52] was selected for the present investigation. The dataset contained structurally diverse compounds including hydantoins and related non-hydantoins such as hydroxy amides, oxazolidinediones, hydroxy acids, hydroxy esters and amino acids. The basic structure for these compounds is shown in Fig. 1 and various substituents are listed in Tables 1 and 2.

The dataset was equally divided into two sets. Compounds with even serial numbers were designated as the training set and those with odd serial numbers were separated as the test set. The values of the *Wiener's topochemical index,* W_c, were computed for all the compounds in both training and test sets using an in-house computer program. The resulting data pertaining to the training set was analyzed and a suitable model was developed after identification of the active range by maximization of the moving average with respect to the active compounds (<35%=inactive, 35–65%=transitional, >65%=active) [62]. Subsequently, each compound in the training set was assigned a biological activity using this model, which was then compared with the reported [52] sodium-channel binding activity. The activity was reported in terms of binding affinity, expressed as IC_{50}, which represents the micromolar concentration of compound required to displace 50% of specifically bound [^{3}H]-batrachotoxinin A 20-α-benzoate. The compounds possessing IC_{50} values of <40 μM were considered to be active and the compounds possessing an IC_{50} values of >40 μM were considered to be inactive for the purpose of the present study [52, 63–65]. The percent accuracy of prediction of active and inactive ranges in the proposed model was calculated. The overall accuracy of prediction was calculated from the ratio of the total number of compounds with correctly predicted activity to that of total number of compounds in both the active and inactive ranges. This model was subsequently cross-validated using the test set.

Table 3 Sodium channel binding activity of various compounds in the training set

Cpd. no.	W_c	χ^A	ξ_c^c	Sodium channel binding activity			
				Assigned			Reported
				W_c	χ^A	ξ_c^c	
A2	236.998	5.947	129.161	−	−	−	−
A4	452.248	7.447	219.665	−	−	−	−
A6	967.914	9.447	410.997	±	+	±	+
A8	295.493	6.417	151.662	−	−	−	−
A10	675.492	8.417	310.659	±	−	±	−
B12	415.073	7.383	202.097	−	−	−	−
B14	862.069	9.383	372.437	±	±	±	+
C16	341.674	6.934	184.449	−	−	−	−
C18	495.342	7.934	218.949	−	−	−	−
C20	855.844	9.434	368.951	±	±	±	±
C22	744.039	7.722	367.059	±	−	±	−
C24	1,113.047	9.222	450.977	±	±	±	−
C26	1,917.725	11.222	720.951	+	+	+	+
A28	452.248	7.443	219.831	−	−	−	−
A30	967.914	9.443	411.163	±	±	±	−
A32	370.985	6.881	190.207	−	−	−	−
C34	1,492.549	11.898	496.942	+	+	+	+
C36	4,104.468	16.651	1,042.688	+	+	+	+
C38	6,012.06	19.151	1,346.147	+	+	+	+
C40	4,601.65	17.141	1,141.692	+	+	+	+
A42	1,046.362	9.976	424.514	±	+	±	+
A44	1,237.782	10.476	483.681	+	+	+	+
A46	2,203.305	13.459	643.019	+	+	+	+
A48	2,621.969	14.253	757.661	+	+	+	+
A50	3,178.995	14.986	874.141	+	+	+	+

+: Active compound, −: inactive compound, ±: compound in the transitional range where activity could not be specifically assigned.

Table 4 Predictability of the proposed models for sodium channel binding activity using training set

Model index	Nature of range in proposed model	Index value	No. of compounds in the range		Percent accuracy of prediction	Average IC_{50}[a](μM)	Overall accuracy of prediction
			Total	Correctly predicted			
W_c	Inactive	≤495.342	8	8	>99	>328.62 (>328.62)	
	Transitional	>495.342 to 1,113.047	8	NA[b]	NA	>182.18 (NA)	>99
	Active	>1,113.047	9	9	>99	8.7 (8.7)	
χ^A	Inactive	≤8.417	10	10	>99	>362.9 (>362.9)	
	Transitional	>8.417 to 9.443	4	NA	NA	105.25 (NA)	>99
	Active	>9.443	11	11	>99	10.43 (10.43)	
ξ_c^c	Inactive	≤219.831	8	8	>99	>328.62 (>328.62)	
	Transitional	>219.831 to 450.977	8	NA	NA	>182.18 (NA)	>99
	Active	>450.977	9	9	>99	8.7 (8.7)	

[a] Values in the bracket indicate average IC_{50} values of correctly predicted compounds of the particular range.
[b] Not applicable.
W_c Wiener's topochemical index, χ^A molecular connectivity topochemical index, ξ_c^c eccentric connectivity topochemical index.

Table 5 Cross-validation of proposed models for sodium channel binding activity using test set

Cpd. no.	W_c	χ^A	ξ_c^c	Sodium channel binding activity			
				Assigned			Reported
				W_c	χ^A	ξ_c^c	
A1	192.583	5.38	118.161	−	−	−	−
A3	294.415	6.447	150.832	−	−	−	−
A5	674.081	8.447	309.331	±	±	±	+
A7	237.994	5.917	129.991	−	−	−	−
A9	453.493	7.417	220.661	−	−	±	−
A11	969.491	9.417	412.657	±	±	±	−
B13	602.571	8.383	269.105	±	−	±	−
B15	1,201.567	10.383	487.769	+	+	+	+
C17	410.508	7.434	197.449	−	−	−	−
C19	597.176	8.434	266.283	±	±	±	+
C21	1,194.512	10.434	483.619	+	+	+	+
C23	846.708	8.222	384.393	±	−	±	−
C25	1,467.386	10.222	570.947	+	+	+	+
A27	236.998	5.943	129.327	−	−	−	−
A29	674.081	8.443	309.497	±	±	±	−
C31	1,688.1	10.703	647.749	+	+	+	−
A33	896.813	9.381	358.762	±	±	±	+
C35	2,448.56	13.898	755.995	+	+	+	+
C37	5,136.022	18.151	1,182.529	+	+	+	+
C39	5,348.541	18.697	1,194.698	+	+	+	+
C41	5,950.391	19.187	1,303.702	+	+	+	+
A43	1,499.619	11.476	513.683	+	+	+	+
A45	1,736.541	11.976	578.85	+	+	+	+
A47	2,912.069	14.959	752.188	+	+	+	+
A49	3,662.678	16.319	905.168	+	+	+	+

+: Active compound, −: inactive compound, ±: compound in the transitional range where activity could not be specifically assigned.

Table 6 Cross-validation of proposed topochemical models using test set

Model index	Nature of range in proposed model	Index value	Number of compounds in the range		Percent accuracy of prediction	Average IC_{50}[a] (μM)	Overall accuracy of prediction
			Total	Correctly predicted			
W_c	Inactive	≤495.342	6	6	>99	>443.67 (>443.67)	
	Transitional	>495.342 to 1,113.047	7	NA[b]	NA	154.29 (NA)	94.45
	Active	>1,113.047	12	11	91.67	9.06 (5.247)	
χ^A	Inactive	≤8.417	8	8	>99	>409.25 (>409.25)	
	Transitional	>8.417 to 9.443	5	NA	NA	93.6 (NA)	95
	Active	>9.443	12	11	91.67	9.06 (5.247)	
ξ_c^c	Inactive	≤219.831	5	5	>99	>432.4 (>432.4)	
	Transitional	>219.831 to 450.977	8	NA	NA	>197.5 (NA)	94.11
	Active	>450.977	12	11	91.67	9.06 (5.247)	

[a] Values in the bracket indicate average IC_{50} values of correctly predicted compounds of the particular range.
[b] Not applicable.
W_c Wiener's topochemical index, χ^A molecular connectivity topochemical index, ξ_c^c eccentric connectivity topochemical index.

The aforementioned procedure was also adopted for molecular connectivity topochemical index, χ^A and eccentric connectivity topochemical index, ξ_c^c. Results are summarized in Tables 3, 4, 5 and 6.

Results and discussion

The critical step in drug discovery remains the identification and optimization of lead compounds in a time and cost-effective manner, which depends on the ability to explore and quantify the relationships between molecular structure and functional properties—particularly the biological activity [66]. Among the numerous computational techniques employed, mathematical-topological methods occupy an eminent place in the field of prediction of properties and activities of chemical compounds and even materials [67]. These methods, known as QSPR/QSAR, are normally, but not always, based upon graph-theoretical descriptors, where molecules are seen as graphs. The basis for these methods is that a correlation exists between various properties of drug molecules and topological indices [68].

The voltage-gated sodium channels are of fundamental importance in the control of sodium fluxes in neurons and other electrically excitable cells, triggering various physiological events. This is clearly demonstrated by rare genetic disorders of sodium-channel activity affecting skeletal muscle contraction, cardiac rhythm or neuronal function and ranging in severity from mild to life threatening. The multiplicity of sodium channels and the dynamic nature of

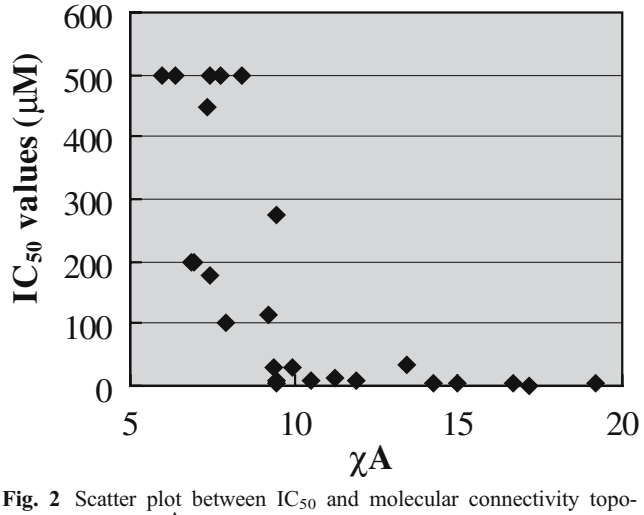

Fig. 2 Scatter plot between IC_{50} and molecular connectivity topochemical index (χ^A) values of various compounds in the training set

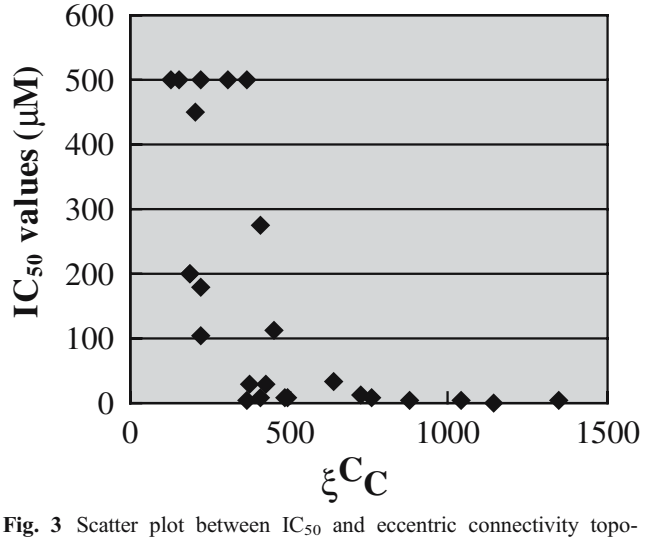

Fig. 3 Scatter plot between IC_{50} and eccentric connectivity topochemical index ξ_c^c values of various compounds in the training set

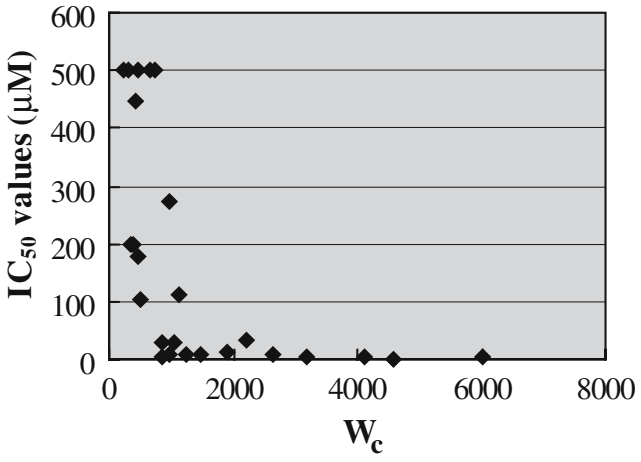

Fig. 4 Scatter plot between IC_{50} and Wiener's topochemical index (W_c) values of various compounds in the training set

their expression make them important targets in the search for novel therapeutics for epilepsy, cardiac arrhythmia and persistent pain syndromes [56].

In the present investigation, three topochemical indices—*Wiener's topochemical index*, *molecular connectivity topochemical index* and *eccentric connectivity topochemical index* were selected for development of models for prediction of sodium channel binding activity, on the basis that these indices provide molecular structural information on different concepts. Wiener's topochemical index is based upon inter-atomic distances, the molecular connectivity topochemical index is based upon connectivity of atoms (adjacency) within a molecule and eccentric connectivity topochemical index is based upon the combination of both adjacency as well as interatomic distances.

A dataset comprising of compounds having diverse nature including hydantoins and related non-hydantoins was selected. The compounds in the dataset possess varying degree of biological activity but only those with IC_{50} values of <40 μM were considered to be active for the purpose of present study. The rational behind the selection of this cutoff limit was that in the sodium channel binding assay, diphenylhydantoin (phenytoin or DPH) was used as a standard reference, and only assays that gave an IC_{50} value for DPH of 40±4 (10%) μM were considered valid [52]. Retrofit analysis of the data in Tables 3, 4, 5 and 6 and Figs. 2, 3 and 4 reveals the following information:

Model based upon Wiener's topochemical index:

- Active and inactive ranges were ideally separated in the proposed model by a transitional range, which clearly indicates gradual change in biological activity. Index values of various activity ranges in the proposed model are given in Table 4.
- The average IC_{50} value of correctly predicted active compounds was found to be 8.7 and 5.247 μM for the training and test sets, respectively, thus indicating high potency of active range of the proposed model.
- The overall ratio of average IC_{50} values of active and inactive ranges was found to be 1:37.77 and 1:48.97 in the training and test sets, respectively. However, this ratio in the case of correctly predicted compounds was found to be 1:37.77 and 1:84.55 in the training and test sets, respectively.

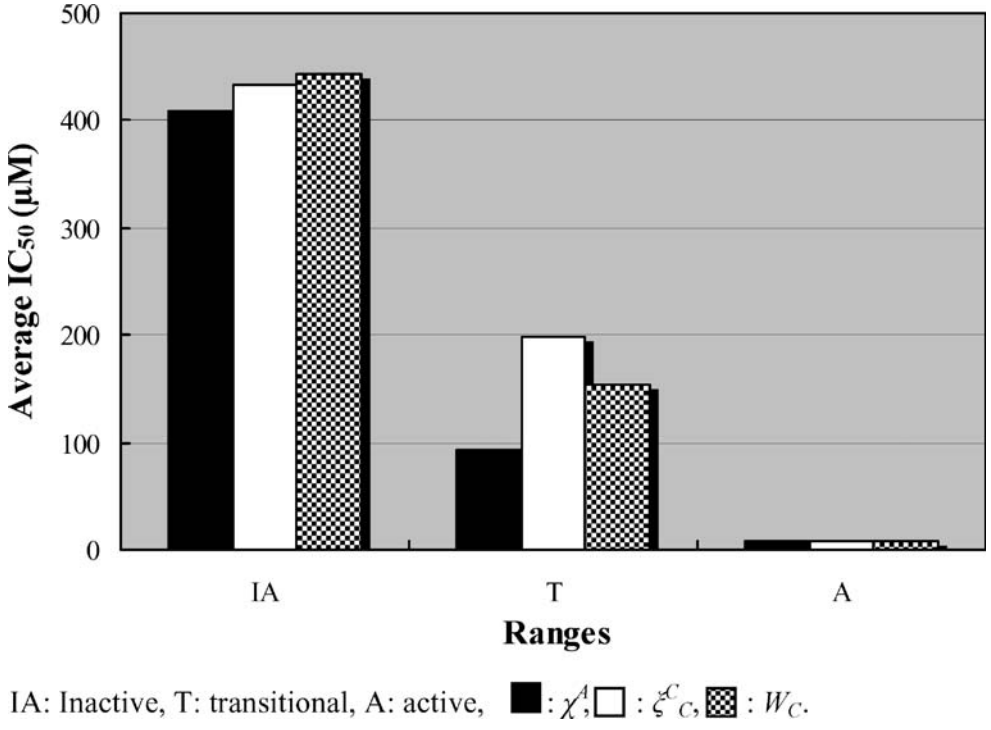

Fig. 5 Average activity values of test compounds and characterization of various activity ranges using Wiener's topochemical index (W_c), molecular connectivity topochemical index (χ^A) and eccentric connectivity topochemical index ξ_c^c

– The overall predictability of the model using the training set was >99%. Cross validation of the proposed model using the test set revealed an exceptionally high accuracy of 94.45%.

Model based upon Molecular connectivity topochemical index:

– Active and inactive ranges were ideally separated in the proposed model by a transitional range, which clearly indicated gradual change in biological activity. Index values of various activity ranges in the proposed model are given in Table 4.
– The average IC_{50} value of correctly predicted active compounds was found to be 10.43 and 5.247 μM for the training and test sets, respectively, thus indicating high potency of the active range in the proposed model based upon molecular connectivity index.
– The overall ratio of IC_{50} values of active and inactive ranges was found to be 1:34.79 and 1:45.17 in the training and test sets, respectively. However, this ratio in the case of correctly predicted compounds was 1:34.79 and 1:77.99 in the training and test sets, respectively.
– The overall predictability of model using the training set was >99%. Cross validation of the proposed model using the test set surprisingly revealed an accuracy of 95%.

Model based upon Eccentric connectivity topochemical index:

– Active and inactive ranges were ideally separated in the proposed model by a transitional range, which clearly indicated gradual change in biological activity. Index values of various activity ranges in the proposed model are shown in Table 4.
– The average IC_{50} value of correctly predicted active compounds in the training set was 8.7 μM. It was only 5.247 μM in the test set, which indicated high potency of active range of the proposed model based upon the eccentric connectivity topochemical index.
– The overall ratio of IC_{50} values of active and inactive ranges was found to be 1:37.77 and 1:47.72 in the training and test sets, respectively. However, this ratio in case of correctly predicted compounds was found to be 1:37.77 and 1:82.40 in the training and test sets, respectively.
– The overall predictability of model using the training set was >99%. Cross validation of the proposed model using the test set revealed an accuracy of 94.11%.

As is evident from Fig. 5 and Table 6, all three models show a large difference in average activity values of compounds falling into the inactive and active ranges (~80 times), indicating the capability of the present models to classify highly inactive and highly potent compounds. Careful examination of the structures of the compounds present in the active ranges of these models reveal that, in general, compounds containing diphenylmethyl-1-piperazinylalkyl substituent possess high sodium-channel binding activity. An *N*-aryl alkyl substituent in the structure also seems to be important for activity.

The methodology used in the present study aims at the development of suitable models for providing lead molecules through exploitation of the active ranges in the proposed models based on topochemical indices. The proposed models are unique and differ widely from conventional QSAR system of modeling. Both systems of modeling have their advantages and limitations. In the instant modeling, the system adopted has distinct advantage of identification of narrow active ranges, which may be erroneously skipped during routine regression analysis in conventional QSAR. Since the ultimate goal of modeling is to provide lead structures, active ranges of the proposed models can play a vital role in providing lead structures [69]. The exceptionally high predictability of the proposed models offers a vast potential for providing lead structures for development of potent sodium-channel binding hydantoins and related non-hydantoins.

References

1. Basak SC, Gute BD, Grunwald GD (1999) Topological indices and related descriptors in QSAR and QSPR. In: Devillers J, Balaban AT (eds) Gordon and Breach Science, The Netherlands, pp 675–696
2. Basak SC, Denise M (2001) J Chem Inf Comput Sci 41:692–701
3. Estrada E, Patlewicz G, Uriarte E (2003) Indian J Chem 42A:1315–1329
4. Basak SC, Grunwald GD (1994) Math Model Sci Comput 4: 464–469
5. Balasubramanian K, Basak SC (1998) J Chem Inf Comput Sci 38:367–373
6. Trinajstic N (1983) Chemical graph theory, vols 1 and 2. CRC, Boca Raton, Florida
7. Dmitriev IS (1981) Molecules without chemical bonds. Mir publishers, Moscow, p 8
8. Basak SC (1987) Med Sci Res 15:605–609
9. Balaban AT, Chiriac A, Motoc I, Simon Z (1980) Steric fit in quantitative structure-activity relations. Springer, Berlin Heidelberg New York, pp 22–51
10. Trinajstic N (1983) Chemical graph theory, vol 2. CRC, Boca Raton, Florida, pp 105–140
11. Gutman I, Polansky O (1986) Mathematical concepts in organic chemistry. Springer, Berlin Heidelberg New York
12. Trinajstic N (1992) Chemical graph theory, 2nd edn., vols 1 and 2. CRC, Boca Raton, Florida, pp 31–34
13. Roy AB, Basak SC, Harriss DK, Magnuson VR (1983) In: Avula XJR, Kalman RE, Liapis AI, Rodin EY (eds) Proceedings of the fourth international conference on mathematical modeling in science and technology. Pergamon, New York, pp 745–750

14. Basak SC, Gute BD, Grunwald GD (1998) In: Carbo-Dorca R, Mezey PG (eds) Advances in molecular similarity vol 2. JAI, Stanford, CT, pp 171–185
15. Bajaj S, Sambi SS, Madan AK (2005) Acta Chim Slov 52:292–296
16. Rouvray DH (1976) In: Balaban AT (ed) Chemical applications of graph theory. Academic, London, pp 180–181
17. Hansen PJ, Jurs PC (1988) J Chem Educ 65:574–580
18. Katritzky AR, Gordeeva EVJ (1993) J Chem Inf Comput Sci 33:835–857
19. Basak SC, Balaban AT, Grunwald GD, Gute BD (2000) J Chem Inf Comput Sci 40:891–898
20. Wiener HJ (1947) Chem Phys 15:766
21. Wiener HJ (1947) J Am Chem Soc 69:17–20
22. Wiener HJ (1947) J Am Chem Soc 69:2636–2638
23. Hosoya H (1971) Bull Chem Soc Jpn 44:2332–2337
24. Balaban AT (1982) Chem Phys Lett 80:399–404
25. Balaban AT (1983) Pure Appl Chem 55:199–206
26. Balaban AT (1986) Math Chem (MATCH) 21:115–122
27. Randic M (1975) J Am Chem Soc 97:6609–6615
28. Kier LB, Murray WJ, Randic M, Hall LH (1976) J Pharm Sci 65:1226–1230
29. Gutman I, Trinajstic N (1972) Chem Phys Lett 17:535–538
30. Nikolic S, Tolic IM, Trinajstic N, Baucic I (2000) Croat Chem Acta 73:909–921
31. Nikolic S, Kovacevic G, Milicevic A, Trinajstic N (2003) Croat Chem Acta 76:113–124
32. Sharma V, Goswami R, Madan AK (1997) J Chem Inf Comput Sci 37:273–282
33. Sardana S, Madan AK (2001) MATCH 43:85–98
34. Sardana S, Madan AK (2002) J Mol Model 8:258–265
35. Sardana S, Madan AK (2002) MATCH 45:36–53
36. Gupta S, Singh M, Madan AK (2002) J Math Anal Appl 266: 259–268
37. Bonchev D, Trinajsic N (1977) J Chem Phys 67:4517–4533
38. Raychoudhary C, Ray SK, Ghosh JJ, Roy AB, Basak SC (1984) J Comput Chem 5:581–588
39. Basak SC, Roy AB, Ghosh JJ (1980) In: Avula XJR, Bellman R, Luke YL, Rigler AK (eds) Proceedings of the second international conference on mathematical modeling. University of Missouri, Rolla, Missouri, p 851
40. Kier LB, Hall LH (1976) Molecular connectivity in chemistry and drug research. Academic, New York
41. Kier LB, Hall LH (1986) Molecular connectivity in structure-activity analysis. Research Studies, Letchworth, Hertfordshire, England, p 262
42. Goel A, Madan AK (1995) J Chem Inf Comput Sci 35:510–514
43. Gupta S, Singh M, Madan AK (2003) Indian J Chem 42A: 1414–1425
44. Kumar V, Sardana S, Madan AK (2004) J Mol Model 10:399–407
45. Bajaj S, Sambi SS, Madan AK (2004) J Mol Struct, Theochem 684:197–203
46. Bajaj S, Sambi SS, Madan AK (2004) QSAR & Comb Sci 23:506–514
47. Hodgkin AL, Huxley AF (1952) J Physiol 117:500–544
48. Yu FH, Catterall WA (2003) Genome Biol 4:207
49. George AL Jr (2005) J Clin Invest 115:1990–1999
50. Roselli F, Livrea P, Jirillo E (2006) Recent patents on CNS drug discovery 1:83–91
51. Cestele S, Catterall WA (2000) Biochimie 82:883–892
52. Zha C, Brown GB, Brouillette WJJ (2004) Med Chem 47: 6519–6528
53. Waxman SG (2005) Brain 128:5–6
54. Stys PK, Waxman SG, Ransom BRJ (1992) Neurosci 12:430–439
55. Stys PK, Ransom BR, Waxman SG (1992) J Neurophysiol 67:236–240
56. Waxman SG, Cummins TR, Dib-Hajj SD, Black JA (2000) J Rehabil Res Dev 37:517–528
57. Randic M (1993) Chem Phys Lett 211:478–483
58. Randic M (2002) Acta Chim Slov 49:483–496
59. Gutman I (2004) Croat Chem Acta 77:61–64
60. Gupta S, Singh M, Madan AK (2001) J Mol Struct, Theochem 571:147–152
61. Dureja H, Madan AK (2005) J Mol Model 11:525–531
62. Gupta S, Singh M, Madan AK (2001) J Comput-aided Mol Des 15:671–678
63. Willow M, Catterall WA (1982) Mol Pharmacol 22:627–635
64. Brown ML, Brown GB, Brouillette WJ (1997) J Med Chem 40:602–607
65. Brown ML, Zha C, Dyke CCV, Brown GB, Brouillette WJ (1999) J Med Chem 42:1537–1545
66. Lather V, Madan AK (2005) J Mol Graph Model 23:339–345
67. Pogliani L (2003) Indian J Chem 42A:1347–1353
68. Bajaj S, Sambi SS, Madan AK (2005) Drug Dev Ind Pharm 31:1041–1051
69. Dureja H, Madan AK (2006) J Mol Graph Model 25:373–379