J Mol Model (2003) 9:259–270
DOI 10.1007/s00894-003-0135-z

ORIGINAL PAPER

Kunal Roy · Achintya Saha

Comparative QSPR studies with molecular connectivity, molecular negentropy and TAU indices

Part I: Molecular thermochemical properties of diverse functional acyclic compounds

Received: 6 March 2003 / Accepted: 7 April 2003 / Published online: 20 June 2003

Abstract Molecular thermochemical properties (heats of formation and atomization) of diverse functional acyclic compounds have been correlated with TAU indices and the relations have been compared to those involving molecular negentropy and first order valence molecular connectivity indices to unravel the diagnostic feature of the TAU scheme and to explore the relative suitability of the scheme in describing physicochemical parameters. For both the properties it was found that TAU relations could satisfactorily explain the variances of the thermochemical parameters and the relations were comparable to those involving molecular negentropy and molecular connectivity. Moreover, specific contributions of functionality, branchedness, shape and size factors to the thermochemical properties could be found from the relations involving TAU parameters.

Keywords QSPR · Molecular connectivity · Molecular negentropy · TAU · Heat of formation · Heat of atomization

Introduction

One of the recent trends in mathematical chemistry is the characterization of molecular structure using graph theoretic approaches [1, 2, 3, 4]. A plethora of topological indices have been defined in the last two decades and these have been extensively used in formulating structure–property/activity/toxicity relationships (QSPR/QSAR/QSTRs) of organic chemicals [5, 6, 7]. A topological index calculated from a molecular graph quantitatively describes the structural information of molecules taking into account parameters like molecular size, shape, adjacency pattern, symmetry, heteroatom variation, cyclicity etc. These indices have been found to be well correlated with numerous physicochemical, pharmacological and toxicological properties and have also been used to calculate or predict such properties from developed mathematical models [5, 6, 7]. QSPR/QSAR studies reveal in a quantitative manner how changes in composition or structure lead to changes in properties and functions [8] and such studies are increasingly being used in chemistry, biochemistry, pharmacology and environmental research.

The present communication attempts to correlate molecular thermochemical properties (heats of formation and atomization) with TAU indices and to compare those with relations involving molecular negentropy (I) and first order valence molecular connectivity ($^1\chi^v$) indices to explore the diagnostic features of the TAU scheme. The TAU scheme is unique in that it unravels specific contributions of functionality, branchedness, shape and size factors to the physicochemical property or biological activity, while other indices mainly give a global contribution of the molecule. Thus, a comparison among the relations involving these indices may explore the relative suitability of the schemes in describing physicochemical parameters. Recently, comparative QSARs of general anesthetic activity and tadpole narcosis with TAU, molecular negentropy and molecular connectivity have been reported [9, 10]. In the present paper, we have used the same data set as used by Kier and Hall [8] for modeling molecular thermochemical properties with the connectivity index.

K. Roy (✉)
Drug Theoretics and Cheminformatics Laboratory,
Division of Medicinal and Pharmaceutical Chemistry,
Department of Pharmaceutical Technology,
Jadavpur University,
700 032, Calcutta, India
e-mail: kunalroy_in@yahoo.com
Tel.: 91-33-2414 6676

A. Saha
Department of Chemical Technology,
University of Calcutta,
92 A P C Road, 700 009, Calcutta, India

Materials and methods

The physicochemical parameters were taken from the literature [8]. First order valence molecular connectivity [8, 11, 12] and molecular negentropy [13, 14, 15] values were calculated according to the original references. TAU indices were introduced by Pal et al. in the late eighties and early nineties [16, 17, 18, 19]. These are *T*opochemically *A*rrived *U*nique indices developed in a VEM (valence electron, mobile) environment. These include T (composite topochemical index), T_R (skeletal index), F (functionality index) and B (simple branchedness index).

In the TAU scheme, a vertex in the molecular graph is considered to be composed of a core and a valence electronic environment. The valence electronic environment is partitioned into two components, localized (identified as the valence electron localized count θ') and mobile (identified as the VEM count θ).

The first order topochemical composite index (T) is defined as

$$T = \sum_{i<j} E_{ij} = \sum_{i<j} \left(V_i V_j\right)^{0.5} \quad (1)$$

where E_{ij}=VEM edge weight of the edge between ith and jth vertices

$$V_i = \text{VEMvertexweightofthe}i\text{thvertex} = \lambda_i/\theta_i \quad (2)$$

$$\lambda_i = \text{Corecountofthe}i\text{thvertex} = (Z - Z^v)/Z^v$$
$$\theta_i = \text{VEMcountofthe}i\text{thvertex} \quad (3)$$

$$= 8 - (2h + 1.5v + n), \text{whenunsaturationisnotpresent} \quad (4)$$

$$= 0.5v + 2\pi, \text{whenunsaturationispresent} \quad (5)$$

h=number of hydrogen atom(s) bonded

v=number of sigma bonds (other than hydrogen)

n=number of nonbonded electrons attached to the atom

π=number of pi bonds associated with the atom

In Eq. (3), Z and Z^V represent the atomic number and the valence electron number, respectively. Obviously, $1/\lambda$ roughly corresponds to the strength of the positive field of the atomic core. The VEM vertex count (θ) is obtained by subtracting the valence electron localized (VEL) count from 8, as a total of eight electrons constitute the valence electronic environment of a bonded atom. While formulating the VEM vertex count equation, it has been considered that an atom enjoys, besides its own, 50% of the other electron in a σ-bond with a non-hydrogen atom. In the case of a σ-bond with a hydrogen atom, the electron pair is predominantly enjoyed by the atom to which it is bonded (considering a graph theoretical self-loop). Further, σ- and π-electrons are given unequal weights in the formalism, thus considering higher mobility of the latter type.

In the case of a heteroatom, the VEM edge weight of an edge incident upon the heteroatom is assigned a negative value. The composite topochemical index may be partitioned into two factors, viz. the first order skeletal index (T_R) and the functionality index (F). The skeletal index T_R is the topochemical index of the reference alkane, which can be obtained by replacing the heteroatom with carbon and removing the multiple bonds that may be present. T_R may further be divided into a simple branchedness index B and the constitutional parameter *vertex count* (N_V).

The derived indices F and B are easily obtained from the following formulae:

$$F = T_R - T(T_R\text{isthefirstorderVEMmolecularindexofthe referencealkane}) \quad (6)$$

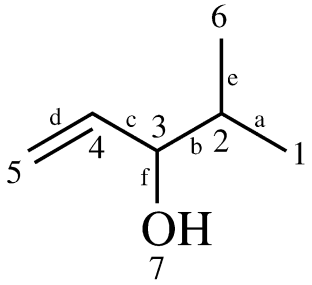

Fig. 1 2-methyl-4-penten-3-ol

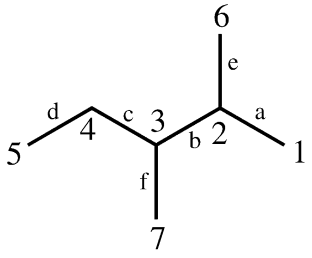

Fig. 2 Reference alkane for 2-methyl-4-penten-3-ol

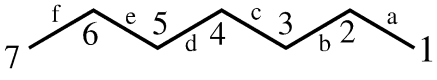

Fig. 3 Normal alkane for 2-methyl-4-penten-3-ol

$$B = T_N - T_R(\text{foracyclicmolecules, } T_N\text{isthetopochemicalindexofthecorrespondingnormalalkane}) \quad (7)$$

The calculation of TAU indices may be illustrated here with an example of 2-methyl-4-penten-3-ol. The hydrogen-suppressed graphical representation of the compound is shown in the chart in Fig. 1 where the atoms (1 though 7) and edges (a through f) have been marked arbitrarily.

The vertex count (V_i) values of the vertices 1 through 7 are calculated to be 1, 1/3, 1/3, 1/6, 1/5, 1 and 2/3 respectively (see Eqs. 2, 3, 4, 5) (see the chart in Fig. 1). Again, the edge counts (E_{ij}) of the edges a though f are calculated to be 0.577, 0.333, 0.236, 0.183, 0.577 and –0.471 respectively (see Eq. 1). The composite index T of the compound is 1.436.

Again, the reference alkane for the above compound is obtained by replacing the heteroatom with carbon (after satisfying valency) and removing the double bond (see the chart in Fig. 2).

The vertex count (V_i) values of the vertices 1 through 7 are calculated to be 1, 1/3, 1/3, 1/2, 1, 1 and 1, respectively. Again, the edge counts (E_{ij}) of the edges a though f are calculated to be 0.577, 0.333, 0.408, 0.707, 0.577 and 0.577, respectively. The skeletal index T_R of the compound is 3.181. The functionality of 2-methyl-4-penten-3-ol is calculated to be 1.745 (Eq. 6).

The structure of the corresponding normal alkane can be seen in the chart in Fig. 3.

The vertex count (V_i) values of the vertices 1 through 7 are calculated to be 1, 1/2, 1/2, 1/2, 1/2, 1/2 and 1, respectively. Again, the edge counts (E_{ij}) of the edges a though f are calculated to be 0.707, 0.5, 0.5, 0.5, 0.5 and 0.707, respectively. The composite index for normal alkane (T_N) of the compound is 3.414. The

Table 1 Topological indices of diverse functional aliphatic compounds

Sl. no.	Compound name	Descriptors				
		$^1\chi^V$	I	T	T_R	T_N
1	Methanol	0.447	3.238	−0.817	1.000	1.000
2	Ethanol	1.023	6.555	0.130	1.414	1.414
3	*n*-Propanol	1.523	10.315	0.630	1.914	1.914
4	2-Propanol	1.413	7.679	0.683	1.731	1.914
5	*n*-Butanol	2.023	14.404	1.130	2.414	2.414
6	2-Methyl propanol	1.879	11.768	0.985	2.270	2.414
7	2-Butanol	1.951	14.177	1.221	2.269	2.414
8	2-Methyl-2-propanol	1.724	7.622	1.092	2.000	2.414
9	*n*-Pentanol	2.523	18.755	1.630	2.914	2.914
10	2-Pentanol	2.451	18.528	1.721	2.769	2.914
11	3-Pentanol	2.489	14.314	1.759	2.807	2.914
12	2-Methyl-1-butanol	2.417	18.528	1.523	2.807	2.914
13	3-Methyl-1-butanol	2.379	16.120	1.485	2.769	2.914
14	2-Methyl-2-butanol	2.284	15.291	1.652	2.561	2.914
15	3-Methyl-2-butanol	2.324	15.893	1.593	2.641	2.914
16	*n*-Hexanol	3.023	23.325	2.130	3.414	3.414
17	*n*-Heptanol	3.523	28.081	2.630	3.914	3.914
18	*n*-Octanol	4.023	33.001	3.130	4.414	4.414
19	2-Ethyl-1-hexanol	3.955	32.774	3.061	4.345	4.414
20	*n*-Nonanol	4.523	38.006	3.630	4.914	4.914
21	*n*-Decanol	5.023	43.261	4.130	5.414	5.414
22	Ethane	1.000	1.954	1.000	1.000	1.000
23	Propane	1.414	5.582	1.414	1.414	1.414
24	*n*-Butane	1.914	7.765	1.914	1.914	1.914
25	2-Methylpropane	1.732	6.026	1.731	1.731	1.914
26	*n*-Pentane	2.414	12.034	2.414	2.414	2.414
27	2-Methylbutane	2.270	13.613	2.269	2.269	2.414
28	2,2-Dimethylpropane	2.000	5.559	2.000	2.000	2.414
29	*n*-Hexane	2.914	14.729	2.914	2.914	2.914
30	2-Methylpentane	2.770	18.114	2.769	2.769	2.914
31	3-Methylpentane	2.808	16.308	2.807	2.807	2.914
32	2,2-Dimethylbutane	2.561	13.968	2.561	2.561	2.914
33	2,3-Dimethylbutane	2.643	9.458	2.641	2.641	2.914
34	*n*-Heptane	3.414	19.426	3.414	3.414	3.414
35	2-Methylhexane	3.270	22.811	3.269	3.269	3.414
36	3-Methylhexane	3.308	25.219	3.307	3.307	3.414
37	3-Ethylpentane	3.346	15.200	3.345	3.345	3.414
38	2,2-Dimethylpentane	3.061	18.665	3.061	3.061	3.414
39	2,3-Dimethylpentane	3.181	22.584	3.179	3.179	3.414
40	2,4-Dimethylpentane	3.126	14.155	3.124	3.124	3.414
41	3,3-Dimethylpentane	3.121	17.768	3.121	3.121	3.414
42	2,2,3-Trimethylbutane	2.943	16.029	2.943	2.943	3.414
43	*n*-Octane	3.914	22.487	3.914	3.914	3.914
44	2-Methylheptane	3.770	27.679	3.769	3.769	3.914
45	3-Methylheptane	3.808	30.087	3.807	3.807	3.914
46	4-Methylheptane	3.808	24.066	3.807	3.807	3.914
47	3-Ethylhexane	3.846	25.873	3.845	3.845	3.914
48	2,2-Dimethylhexane	3.561	23.532	3.561	3.561	3.914
49	2,3-Dimethylhexane	3.681	27.451	3.679	3.679	3.914
50	2,4-Dimethylhexane	3.664	27.451	3.662	3.662	3.914
51	2,5-Dimethylhexane	3.626	17.216	3.624	3.624	3.914
52	3,3-Dimethylhexane	3.621	26.849	3.621	3.621	3.914
53	3,4-Dimethylhexane	3.719	22.033	3.717	3.717	3.914
54	2-Methyl-3-ethylpentane	3.719	23.237	3.717	3.717	3.914
55	3-Methyl-3-ethylpentane	3.682	19.238	3.682	3.682	3.914
56	2,2,3-Trimethylpentane	3.481	23.305	3.481	3.481	3.914
57	2,2,4-Trimethylpentane	3.417	20.897	3.416	3.416	3.914
58	2,3,3-Trimethylpentane	3.504	24.214	3.503	3.503	3.914
59	2,3,4-Trimethylpentane	3.553	18.795	3.551	3.551	3.914
60	2,2,3,3-Tetramethylbutane	3.250	8.923	3.250	3.250	3.914
61	*n*-Nonane	4.414	27.506	4.414	4.414	4.414
62	4-Methyloctane	4.308	32.697	4.307	4.307	4.414
63	2,2-Dimethylheptane	4.061	28.550	4.061	4.061	4.414
64	2,2,3-Trimethylhexane	3.981	28.323	3.981	3.981	4.414
65	2,2,4-Trimethylhexane	3.954	28.323	3.954	3.954	4.414
66	2,2,5-Trimethylhexane	3.916	25.915	3.916	3.916	4.414
67	2,3,3-Trimethylhexane	4.004	29.232	4.003	4.003	4.414
68	2,3,5-Trimethylhexane	4.037	29.834	4.034	4.034	4.414

Table 1 (continued)

Sl. no.	Compound name	Descriptors				
		$^1\chi^V$	I	T	T_R	T_N
69	2,4,4-Trimethylhexane	3.977	29.232	3.976	3.976	4.414
70	3,3,4-Trimethylhexane	4.042	29.232	4.041	4.041	4.414
71	2,2-Dimethyl-3-ethylpentane	4.019	24.109	4.019	4.019	4.414
72	2,4-Dimethyl-3-ethylpentane	4.091	23.814	4.089	4.089	4.414
73	3,3-Diethylpentane	4.243	17.418	4.243	4.243	4.414
74	2,2,3,3-Tetramethylpentane	3.811	25.086	3.811	3.811	4.414
75	2,2,3,4-Tetramethylpentane	3.854	25.688	3.853	3.853	4.414
76	2,2,4,4-Tetramethylpentane	3.707	13.942	3.707	3.707	4.414
77	2,3,3,4-Tetramethylpentane	3.887	20.576	3.885	3.885	4.414
78	*n*-Decane	4.914	30.853	4.914	4.914	4.914
79	3,3,5-Trimethylheptane	4.515	36.794	4.514	4.514	4.914
80	2,2,3,3-Tetramethylhexane	4.311	30.239	4.311	4.311	4.914
81	2,2,5,5-Tetramethylhexane	4.207	17.289	4.207	4.207	4.914
82	*n*-Undecane	5.414	36.128	5.414	5.414	5.414
83	2-Methyldecane	5.270	43.126	5.269	5.269	5.414
84	*n*-Dodecane	5.914	39.709	5.914	5.914	5.914
85	2,2,4,4,6-Pentamethylheptane	5.123	38.266	5.123	5.123	5.914
86	2,2,4,6,6-Pentamethylheptane	5.101	27.724	5.100	5.100	5.914
87	*n*-Hexadecane	7.914	58.605	7.914	7.914	7.914
88	2-Methylpentadecane	7.770	71.021	7.769	7.769	7.914
89	Ethylene	0.500	1.659	0.200	1.000	1.000
90	Propylene	0.986	6.555	0.591	1.414	1.414
91	1-Butene	1.524	10.315	1.179	1.914	1.914
92	*trans*-2-Butene	1.488	6.475	0.984	1.914	1.914
93	2-Methylpropene	1.354	7.077	0.925	1.731	1.914
94	1-Pentene	2.024	14.404	1.679	2.414	2.414
95	*trans*-2-Pentene	2.026	14.177	1.572	2.414	2.414
96	2-Methyl-1-butene	1.914	13.575	1.522	2.269	2.414
97	3-Methyl-1-butene	1.896	11.768	1.573	2.269	2.414
98	2-Methyl-2-butene	1.866	10.939	1.319	2.269	2.414
99	1-Hexene	2.524	18.755	2.179	2.914	2.914
100	*trans*-2-Hexene	2.526	18.528	2.072	2.914	2.914
101	*trans*-3-Hexene	2.564	13.110	2.159	2.914	2.914
102	2-Methyl-1-pentene	2.414	17.926	2.022	2.769	2.914
103	3-Methyl-1-pentene	2.434	18.528	2.111	2.807	2.914
104	4-Methyl-1-pentene	2.379	16.120	2.034	2.769	2.914
105	2-Methyl-2-pentene	2.404	15.291	1.907	2.769	2.914
106	3-Methyl-*trans*-2-pentene	2.427	17.699	1.916	2.807	2.914
107	4-Methyl-*trans*-2-pentene	2.399	15.893	1.966	2.769	2.914
108	2-Ethyl-1-butene	2.475	13.712	2.118	2.807	2.914
109	2,3-Dimethyl-1-butene	2.297	15.291	1.920	2.641	2.914
110	3,3-Dimethyl-1-butene	2.197	11.973	1.887	2.561	2.914
111	2,3-Dimethyl-2-butene	2.250	6.634	1.656	2.641	2.914
112	1-Heptene	3.024	23.325	2.679	3.414	3.414
113	5-Methyl-1-hexene	2.879	20.689	2.534	3.269	3.414
114	3-Methyl-*trans*-3-hexene	2.965	22.268	2.503	3.307	3.414
115	2,4-Dimethyl-1-pentene	2.770	19.860	2.377	3.124	3.414
116	4,4-Dimethyl-1-pentene	2.670	16.543	2.325	3.061	3.414
117	2,4-Dimethyl-2-pentene	2.777	17.225	2.301	3.124	3.414
118	4,4-Dimethyl-*trans*-2-pentene	2.699	16.316	2.280	3.061	3.414
119	3-Methyl-2-ethyl-1-butene	2.858	19.860	2.516	3.179	3.414
120	2,3,3-Trimethyl-1-butene	2.604	15.714	2.236	2.943	3.414
121	1-Octene	3.524	28.081	3.179	3.914	3.914
122	2,2,-Dimethyl-*trans*-3-hexene	3.237	21.072	2.867	3.561	3.914
123	2-Methyl-3-ethyl-1-pentene	3.373	22.810	2.996	3.717	3.914
124	2,4,4-Trimethyl-2-pentene	3.077	17.835	2.615	3.416	3.914
125	1-Decene	4.524	38.066	4.179	4.914	4.914
126	Dimethyl ether	0.816	3.317	−1.154	1.414	1.414
127	Methyl ethyl ether	1.404	9.485	−0.278	1.914	1.914
128	Diethyl ether	1.992	9.360	0.598	2.414	2.414
129	Methyl-*n*-propyl ether	1.904	13.575	0.222	2.414	2.414
130	Methyl-*sec*-propyl ether	1.799	10.939	0.244	2.269	2.414
131	Methyl-*tert*-butyl ether	2.112	11.144	0.634	2.561	2.914
132	Di-*n*-propyl ether	2.992	16.475	1.598	3.414	3.414
133	Di-*sec*-propyl ether	2.781	11.204	1.642	3.124	3.414
134	Isopropyl-*tert*-butyl ether	3.095	17.835	2.033	3.416	3.914
135	Di-*n*-butyl ether	3.992	24.345	2.598	4.414	4.414
136	Di-*sec*-butyl ether	3.857	23.890	2.718	4.201	4.414

Table 1 (continued)

Sl. no.	Compound name	Descriptors				
		$^1\chi^V$	I	T	T_R	T_N
137	Di-*tert*-butyl ether	3.408	10.781	2.423	3.707	4.414
138	Methanethiol	1.341	3.238	−1.826	1.000	1.000
139	Ethanethiol	1.655	6.555	−0.584	1.414	1.414
140	1-Propanethiol	2.155	10.315	−0.084	1.914	1.914
141	2-Propanethiol	1.929	7.679	0.101	1.731	1.914
142	1-Butanethiol	2.655	14.404	0.416	2.414	2.414
143	2-Butanethiol	2.467	14.177	0.639	2.269	2.414
144	2-Methyl-1-propanethiol	2.511	11.768	0.271	2.269	2.414
145	2-Methyl-2-propanethiol	2.171	7.622	0.587	2.000	2.414
146	1-Pentanethiol	3.155	18.755	0.916	2.914	2.914
147	3-Methyl-1-butanethiol	3.011	16.120	0.771	2.769	2.914
148	2-Methyl-1-butanethiol	3.049	18.528	0.809	2.807	2.914
149	1-Hexanethiol	3.655	23.325	1.416	3.414	3.414
150	1-Heptanethiol	4.155	28.081	1.916	3.914	3.914
151	1-Decanethiol	5.665	43.261	3.416	5.414	5.414

Table 2 Relations of heat of formation (H_f) of alcohols with various indices. Model equation: $H_f=\sum\beta_i x_i+\alpha$

Eq. no.	Type of index	Regression coefficient(s) and constant[a]					Statistics		
		β_1 s.e.	β_2 s.e.	β_3 s.e.	β_4 s.e.	α s.e.	R^2_a (r or R)	s (F (df))[b]	$AVRES$ (n)
8	MCI	9.110 $^1\chi^v$				50.465	0.885	3.749	2.772
		0.730				2.014	(0.944)	(155.659 (1, 19))	(21)
9	MN	0.920 I				56.260	0.791	5.070	3.554
		0.105				2.253	(0.895)	(76.496 (1, 19))	(21)
10	TAU	9.231 T				57.982	0.937	2.785	2.032
		0.535				1.082	(0.970)	(297.549 (1, 19))	(21)
11	TAU	8.974 T_R				47.487	0.855	4.220	3.094
		0.823				2.552	(0.928)	(118.791 (1, 19))	(21)
12	TAU	22.084 B	5.008 N_V			40.470	0.982	1.468	1.051
		2.702	0.150			1.059	(0.992)	(560.658 (2, 18))	(21)
13	TAU	4.729 N_I	14.161N_B			51.596	0.931	2.917	1.664
		0.289	1.331			1.555	(0.968)	(135.222 (2, 18))	(21)
14	TAU	4.909 N_V	4.230 N_B			40.850	0.974	1.785	1.169
		0.180	0.674			1.279	(0.988)	(376.203 (2, 18))	(21)
15	TAU	4.939 N_I	22.122 N_X	13.706 N_Y		50.444	0.982	1.471	1.164
		0.148	1.277	0.674		0.665	(0.992)	(372.280 (3, 17))	(21)

[a] s.e.=standard error; t values of the regression coefficients and constants are significant at 95% level [$df=n-np-i$; np=no. of predictor variables; i=1 if intercept is present; i=0, otherwise]
[b] F values are significant at 99% level [$df=np$, $n-np-i$]

branchedness of 2-methyl-4-penten-3-ol is calculated to be 0.233 (Eq. 7).

The vertex count (N_V) of the hydrogen-suppressed molecular formula is purely an atopological parameter because it may be obtained directly from the molecular formula. Not even the structural formula is needed for obtaining the value of N_V. Obviously, any index showing better correlation with physicochemical or biological activity than that shown by N_V will have significance in the context of QSAR/QSPR studies. N_V can be partitioned into N_P (number of methyl carbons), N_I (number of methylene carbons) and N_B (number of branched carbons). N_B may further be factored into N_X (number of quaternary carbons) and N_Y (number of tertiary carbons). The integer index values are easily obtained from the structure of the reference alkane and these values in the above example of 2-methyl-4-penten-3-ol are: N_V=7, N_B=2, N_X=0, N_Y=2, N_I=1 and N_P=4. During development of QSAR equations with TAU parameters, the above mentioned hierarchical relations among various TAU parameters are followed. For obvious reasons, B and N_B (both represent branchedness) or N_P and N_B (both have interrelation) [19] or N_V and N_I (N_I may be considered as a trimmed counterpart of N_V) [19] are not used in the same equation. It may be mentioned here that all these TAU indices are derived by sequentially partitioning the first-order composite index T into different factors. In this paper, multifactorial TAU relations have been compared to those with molecular connectivity and molecular negentropy only to show the statistical acceptance of TAU relations in the perspectives of other well accepted topological indices.

The first-order VEM molecular index T_R is considered as the index for intrinsic lipophilicity while N_B, N_X and N_Y represent shape parameters [9, 10, 16, 17]. The functionality contribution and bulk parameter are represented by F and N_V, respectively [9, 10, 16, 17].

Multiple linear regression analyses were done using the program *RRR98* developed by one of the authors [20]. The statistical quality of the equations [21] was judged by examining the parameters like R_a^2 (adjusted R^2, i.e., explained variance), r or R (correlation coefficient), F (variance ratio) with df (degree of freedom), s (standard error of estimate) and $AVRES$ (average of absolute values of residuals). The significance of the regression coefficients and constants (intercepts) was judged by the "t" test. In

Table 3 Relations of heat of formation (H_f) of alkanes with various indices. Model equation: $H_f=\sum\beta_i x_i+\alpha$

Eq. no.	Type of index	Regression coefficient(s) and constant[a]					Statistics		
		β_1 s.e.	β_2 s.e.	β_3 s.e.	β_4 s.e.	α s.e.	R^2_a (r or R)	s (F (df))	*AVRES* (n)
16	MCI	9.934 $^1\chi^v$ 0.391				15.747 1.524	0.907 (0.953)	3.656 (645.474 (1, 65))	3.044 (67)
17	MN	0.914 I 0.068				31.203 1.777	0.731 (0.857)	6.225 (180.02 (1, 65))	4.479 (67)
18	TAU	9.934 T 0.391				15.755 1.523	0.907 (0.953)	3.655 (645.537 (1, 65))	3.044 (67)
19	TAU	4.844 N_V 0.058	6.259 B 0.636			11.384 0.480	0.992 (0.996)	1.083 (4015.679 (2, 64))	0.785 (67)
20	TAU	4.847 N_I 0.066	16.448 N_X 0.268	10.208 N_Y 0.218		21.316 0.374	0.989 (0.995)	1.230 (2071.78 (3, 63))	0.942 (67)
21	TAU	4.689 N_I 0.213	12.132 N_B 0.658			22.508 1.444	0.890 (0.945)	3.981 (267.469 (2, 64))	3.164 (67)

[a] Obs. = Observed (Ref. [8]; Calc. = Calculated, [b] From Eq.15, [c] From Eq. 19, [d] From Eq. 29, [e] From Eq. 37, [f] From Eq. 41, [g] From Eq. 50, [h] From Eq. 57, [i] From Eq. 65, [j] From Eq. 74

Table 4 Relations of heat of formation (H_f) of the composite set (alcohols and alkanes) with various indices. Model equation: $H_f=\sum\beta_i x_i+\alpha$

Eq. no.	Type of index	Regression coefficient(s) and constant[a]					Statistics		
		β_1 s.e.	β_2 s.e.	β_3 s.e.	β_4 s.e.	α s.e.	R^2_a (r or R)	s (F (df))	*AVRES* (n)
22	MCI	5.184 $^1\chi^v$ 1.132				39.867 4.140	0.187 (0.443)	13.250 (20.970 (1, 86))	10.686 (88)
23	MN	0.736 I 0.116				41.197 2.905	0.311 (0.565)	12.194 (40.300 (1, 86))	9.967 (88)
24	TAU	2.422 T 1.064				49.854 3.768	0.046 (0.238)	14.351 (5.176 (1, 86))*	11.305 (88)
25	TAU	6.637 T_R 1.112				34.288 4.139	0.285 (0.541)	12.427 (35.607 (1, 86))	10.065 (88)
26	TAU	9.532 T_R 0.483	22.217 F 1.091			17.628 1.902	0.877 (0.938)	5.154 (310.962 (2, 85))	3.728 (88)
27	TAU	25.210 F 0.805	9.316 B 1.962	4.746 N_V 0.164		11.647 1.342	0.943 (0.972)	3.520 (477.219 (3, 84))	1.836 (88)
28	TAU	23.300 F 1.077	51.561 B 3.074	4.397 N_I 0.217		26.203 1.317	0.894 (0.947)	4.795 (244.305 (3, 84))	3.455 (88)
29	TAU	25.163 F 0.849	4.743 N_I 0.172	16.868 N_X 0.773	10.722 N_Y 0.606	21.266 0.965	0.938 (0.970)	3.655 (330.754 (4, 83))	1.971 (88)

the case that the intercept of an equation was statistically insignificant and omission of the same did not affect the quality of the equation, exclusion of the intercept gave a statistically more acceptable equation. The robustness of the best equations under different series was checked with the "leave-one-out" technique [22, 23] using the programs *KRPRES1* and *KRPRES2* [20]

Results and discussion

The calculated topological index values of some diverse functional acyclic organic compounds are shown in Table 1. Tables 2, 3 and 4 show the relations of the heat of formation data of aliphatic hydrocarbons and alcohols to various indices. The relations of heat of atomization data of aliphatic hydrocarbons, alcohols, ethers and thiols with different indices are shown in Tables 5, 6, 7, 8, 9 and 10. Regression coefficients and F ratios of all accepted equations are significant at the 99% confidence level (unless marked with *).

Table 2 shows that first order valence molecular connectivity and molecular negentropy can explain 88.5% (Eq. 8) and 79.1% (Eq. 9), respectively, of the variance of the heat of formation data of alcohols while composite topochemical index (T) is capable of explaining 93.7% of the variance (Eq. 10). When the composite index is partitioned into B and N_V, the resultant relation (Eq. 12) explains 98.2% of the variance. The best relation (Eq. 15) involving TAU parameters (N_I, N_X and N_Y) explains 98.2% of the variance. Specific contributions of branchedness (B), shape (N_X or N_Y) and size parameters

Table 5 Relations of heat of atomization (H_a) of alcohols with various indices. Model equation: $H_a=\sum\beta_i x_i+\alpha$

Eq. no.	Type of index	Regression coefficient(s) and constant[a]					Statistics		
		β_1 s.e.	β_2 s.e.	β_3 s.e.	β_4 s.e.	α s.e.	R^2_a (r or R)	s (F (df))	*AVRES* (n)
30	MCI	544.985 $^1\chi^v$ 9.809				278.667 27.052	0.994 (0.997)	50.361 (3086.66 (1, 19))	37.156 (21)
31	MN	58.307 I 3.967				592.409 82.413	0.915 (0.959)	183.049 (216.074 (1, 19))	131.723 (21)
32	TAU	536.950T 11.684				735.840 23.621	0.991 (0.996)	60.796 (2112.069 (1, 19))	44.341 (21)
33	TAU	543.777T_R 14.576				80.383 45.174	0.986 (0.993)	74.723 (1391.704 (1, 19))	53.305 (21)
34	TAU	543.507T_R 8.209	−314.221 F 48.539			463.722 64.450	0.996 (0.998)	42.081 (2214.989 (2, 18))	34.752 (21)
35	TAU	−7.394 F 1.162	14.641 B 1.845	279.958N_V 0.070		−60.166 1.518	0.999 (0.999)	0.641 (6.39×10^6 (3, 17))	0.421 (21)
36	TAU	−293.539 F 107.058	269.726N_I 7.417	514.172N_B 41.182		932.103 137.218	0.986 (0.994)	73.997 (473.830 (3, 17))	48.245 (21)
37	TAU	−7.809 F 0.995	279.903N_I 0.061	845.252N_X 0.704	561.727N_Y 0.334	500.496 1.249	0.999 (0.999)	0.577 (5.91×10^6 (4, 16))	0.358 (21)

Table 6 Relations of heat of atomization (H_a) of alkanes with various indices. Model equation: $H_a=\sum\beta_i x_i+\alpha$

Eq. no.	Type of index	Regression coefficient(s) and constant[a]					Statistics		
		β_1 s.e.	β_2 s.e.	β_3 s.e.	β_4 s.e.	α s.e.	R^2_a (r or R)	s (F (df))	*AVRES* (n)
38	MCI	590.427 $^1\chi^v$ 24.617				207.849 82.050	0.930 (0.965)	107.807 (575.262 (1, 42))	87.323 (44)
39	MN	45.631 I 6.485				1267.165 130.598	0.530 (0.736)	279.988 (49.513 (1, 42))	208.523 (44)
40	TAU	590.557T 24.619				207.881 82.037	0.930 (0.965)	107.792 (575.435 (1, 42))	87.289 (44)
41	TAU	280.169N_V 0.106	6.330 B 0.805			116.732 0.712	0.999 (0.999)	0.910 (4.33×10^6 (2, 41))	0.682 (44)
42	TAU	281.309N_P 0.278	280.105N_I 0.131	279.694N_B 0.406		114.772 1.250	0.999 (0.999)	1.048 (2.18×10^6 (3, 40))	0.802 (44)
43	TAU	280.105N_I 0.131	842.312N_X 0.358	561.003N_Y 0.263		677.390 0.477	0.999 (0.999)	1.048 (2.18×10^6 (3, 40))	0.802 (44)

(N_V or N_I) are explored from the relations involving TAU indices. Positive coefficients of T_R, N_V and B indicate that the heat of formation increases with increase in the values of skeletal index, molecular bulk and branchedness. The calculated heat of formation data according to the Eq. 15 are shown in Table 11.

In the case of alkanes, first order valence molecular connectivity and molecular negentropy can explain 90.7% (Eq. 16) and 73.1% (Eq. 17) respectively of the variance while the composite topochemical index (T) explains (Eq. 18) to the same extent as molecular connectivity does. However, when T is partitioned into B and N_V, the resultant relation (Eq. 19) explains 99.2% of the variance. Heat of formation values of alkanes increase with molecular bulk and branchedness as evidenced from TAU relations. The calculated heat of formation data according to the Eq. 19 are shown in Table 11.

For the composite set, only 18.7% (Eq. 22) and 31.1% (Eq. 23) of the variances of heat of formation data are explained by first-order valence molecular connectivity and molecular negentropy respectively while the composite topochemical index (Eq. 24) gives a further inferior relation. However, when the composite index is partitioned into F, B and N_V, the resultant relation (Eq. 27) explains 94.3% of the variance. Specific contributions of branchedness (B), functionality (F), shape (N_X or N_Y) and size parameters (N_V or N_I) are also explored from the relations involving TAU indices. The calculated heat of formation data according to Eq. 29 are shown in Table 11.

From the relations of the heat of formation data with TAU indices in the case of the composite set, it appears that H_f increases with an increase in intrinsic lipophilicity (T_R), branchedness (B), molecular bulk (N_V) and functionality (F).

Table 7 Relations of heat of atomization (H_a) of alkenes with various indices. Model equation, $H_a=\sum\beta_i x_i+\alpha$

Eq. no.	Type of index	Regression coefficient (s) and constant[a]					Statistics		
		β_1 s.e.	β_2 s.e.	β_3 s.e.	β_4 s.e.	α s.e.	R^2_a (r or R)	s (F (df))	$AVRES$ (n)
44	MCI	578.038 $^1\chi^v$ 17.243				284.143 43.656	0.969 (0.985)	75.125 (1123.76 (1, 35))	59.537 (37)
45	MN	57.161 I 5.193				766.956 90.168	0.769 (0.881)	204.636 (121.168 (1, 35))	161.416 (37)
46	TAU	566.927T 20.469				537.450 44.108	0.955 (0.978)	90.295 (767.098 (1, 35))	73.331 (37)
47	TAU	601.568T_R 4.725					0.962 (0.981)	83.015 (1.62×10^4 (1, 36))	66.070 (37)
48	TAU	277.217N_V 0.407	20.558 B 12.156				0.999 (0.999)	9.743 (5.89×10^5 (2, 35))	4.773 (37)
49	TAU	279.926N_I 12.554	603.017N_B 29.264			552.316 50.307	0.941 (0.972)	103.364 (289.051 (2, 34))	70.180 (37)
50	TAU	280.312N_I 0.986	844.025N_X 3.984	565.912N_Y 2.351		539.487 3.230	0.999 (0.999)	8.115 (3.31×10^4 (3, 33))	3.472 (37)
51	TAU	269.743N_P 1.977	280.311N_I 0.986	304.537N_X 6.918	296.168N_Y 4.119		0.999 (0.999)	8.115 (4.25×10^5 (4, 33))	3.471 (37)

Table 8 Relations of heat of atomization (H_a) of ethers with various indices. Model equation: $H_a=\sum\beta_i x_i+\alpha$

Eq. no.	Type of index	Regression coefficient(s) and constant[a]					Statistics		
		β_1 s.e.	β_2 s.e.	β_3 s.e.	β_4 s.e.	α s.e.	R^2_a (r or R)	s (F (df))	$AVRES$ (n)
52	MCI	576.478 $^1\chi^v$ 33.208				274.284 89.218	0.965 (0.984)	109.421 (301.365 (1, 10))	80.892 (12)
53	MN	74.547 I 18.513				714.222 273.096	0.580 (0.786)	377.101 (16.215 (1, 10))	244.939 (12)
54	TAU	464.014T 20.711				1209.351 33.645	0.979 (0.990)	85.332 (501.967 (1, 10))	61.557 (12)
55	TAU	588.722T_R 12.193					0.950 (0.975)	129.604 (2331.262 (1, 11))	104.743 (12)
56	TAU	483.531T_R 46.267	−411.548 F 112.138			1055.882 323.795	0.977 (0.990)	88.833 (231.704 (2, 9))	59.536 (12)
57	TAU	−9.683 F 3.558	7.323 B 3.463	279.873N_V 0.537		−56.677 7.486	0.999 (0.999)	1.809 (3.80×10^5 (3, 8))	1.087 (12)
58	TAU	−457.969 F 290.238	217.335N_I 40.806	458.139N_B 137.098		1654.773 570.644	0.929 (0.974)	154.858 (49.152 (3, 8))	101.998 (12)
59	TAU	−7.570 F 3.580	279.965N_I 0.502	842.864N_X 2.089	562.159N_Y 1.535	498.034 6.836	0.999 (0.999)	1.675 (3.32×10^5 (4, 7))	0.895 (12)

In the case of the heat of atomization data (Tables 5, 6, 7, 8, 9 and 10) molecular connectivity can explain 99.4% of the variance for alcohols (Eq. 30), 93.0% for alkanes (Eq. 38), 96.9% for alkenes (Eq. 44), 96.5% for ethers (Eq. 52), 97.8% for thiols (Eq. 60) and 90.3% for the composite set (Eq. 68). On the other hand, molecular negentropy gives inferior relations for all types of compounds. It can explain only 91.5% variance for alcohols (Eq. 31), 53.0% for alkanes (Eq. 39), 76.9% for alkenes (Eq. 45), 58.0% for ethers (Eq. 53), 94.5% for thiols (Eq. 61) and 67.8% for the composite set (Eq. 69). The composite topochemical index (T) explains 99.1% of variance for alcohols (Eq. 32), 93.0% for alkanes (Eq. 40), 95.5% for alkenes (Eq. 46), 97.9% for ethers (Eq. 54), 96.4% for thiols (Eq. 62) and 84.1% for the composite set (Eq. 70). When the composite index was suitably factored into different components, the TAU indices could explain more than 99.5% of the variance in all the cases: up to 100% for alcohols, alkanes, alkenes, ethers and thiols (Eqs. 35/37, 41/42/43, 48/50/51, 57/59, 65/66/67) and up to 99.7% for the composite set (Equations 73/74). Specific contributions of branchedness (B), functionality (F), shape (N_X or N_Y) and size parameters (N_V or N_I) are also explored from the relations involving TAU indices. The calculated heat of atomization data according to the

Table 9 Relations of heat of atomization (H_a) of thiols with various indices. Model equation: $H_a=\sum\beta_i x_i+\alpha$

Eq. no.	Type of index	Regression coefficient(s) and constant[a]					Statistics		
		β_1 s.e.	β_2 s.e.	β_3 s.e.	β_4 s.e.	α s.e.	R^2_a (r or R)	s (F (df))	*AVRES* (n)
60	MCI	548.249 $^1\chi^v$ 22.936				−117.282 69.376	0.978 (0.990)	92.382 (571.371 (1, 12))	64.161 (14)
61	MN	58.017 I 3.865				504.904 72.912	0.945 (0.974)	144.832 (225.352 (1, 12))	92.933 (14)
62	TAU	508.168T 27.207				1114.349 35.715	0.964 (0.983)	117.461 (348.854 (1, 12))	86.292 (14)
63	TAU	552.735T_R 6.812					0.987 (0.993)	71.478 (6584.161 (1, 13))	54.482 (14)
64	TAU	543.203T_R 11.490	−184.210 F 40.958			387.790 89.788	0.994 (0.998)	46.039 (1168.937 (2, 11))	36.349 (14)
65	TAU	13.266 B 2.163	280.113N_V 0.117			−109.242 0.736	0.999 (0.999)	0.925 (2.91×10^6 (2, 11))	0.469 (14)
66	TAU	280.103N_V 0.129	2.995 N_B 0.555			−109.389 0.824	0.999 (0.999)	1.018 (2.40×10^6 (2, 11))	0.649 (14)
67	TAU	280.110N_I 0.092	845.580N_X 0.848	562.770N_Y 0.478		450.790 0.367	0.999 (0.999)	0.730 (3.12×10^6 (3, 10))	0.395 (14)

Table 10 Relations of heat of atomization (H_a) of the composite set (aliphatic hydrocarbons, alcohols, ethers, thiols) with various indices. Model equation: $H_a=\sum\beta_i x_i+\alpha$

Eq. no.	Type of index	Regression coefficient(s) and constant[a]					Statistics		
		β_1 s.e.	β_2 s.e.	β_3 s.e.	β_4 s.e.	α s.e.	R^2_a (r or R)	s (F (df))	*AVRES* (n)
68	MCI	562.823 $^1\chi^v$ 16.324				245.205 47.864	0.903 (0.951)	170.367 (1188.691 (1, 126))	114.092 (128)
69	MN	57.841 I 3.529				816.393 66.655	0.678 (0.825)	310.965 (268.616 (1, 126))	245.279 (128)
70	TAU	393.952T 15.197				962.286 38.041	0.841 (0.918)	218.677 (671.975 (1, 126))	168.126 (128)
71	TAU	606.012T_R 15.261				14.383 47.155	0.925 (0.962)	149.699 (1576.789 (1, 126))	122.001 (128)
72	TAU	569.533T_R 10.342	−161.560 F 12.358			253.389 35.793	0.968 (0.984)	97.686 (1936.923 (2, 125))	77.629 (128)
73	TAU	−111.110 F 4.263	2.542 B 17.710	279.987N_V 1.664		96.560 11.301	0.997 (0.998)	31.522 (1.28×10^4 (3, 124))	27.233 (128)
74	TAU	−110.455 F 4.365	279.675N_I 1.701	844.064N_Y 6.897	559.925N_Y 4.812	656.145 7.434	0.997 (0.998)	31.570 (9541.036 (4, 123))	27.113 (128)

best equations under individual series and composite set are shown in Table 11.

From the relations of the heat of atomization data with TAU indices in the case of the composite set, it appears that H_a increases with an increase in intrinsic lipophilicity (T_R), branchedness (B) and molecular bulk (N_V) and decreases with an increase in functionality.

The PRESS statistics of the best equations under individual series and composite set for each thermochemical data are shown in Table 12. These results prove the stability and predictive potential of the equations.

The present study shows that, although the composite topochemical index T does not always provide a better model for molecular thermochemical properties of heterofunctional acyclic compounds in comparison to molecular connectivity and negentropy, the TAU scheme can generate statistically comparable relations when the composite index is partitioned into different components like skeletal index, size and shape factors, branchedness and functionality. Moreover, TAU indices can unravel specific contributions of molecular bulk (size), functionality, branchedness and shape parameters to the molecular thermochemical properties of diverse functional compounds. The diagnostic feature of the TAU scheme and its suitability in describing QSPR relations in comparison to molecular negentropy and molecular connectivity are revealed from the present study. However, further studies need be done on a wide range of physicochemical

Table 11 Observed and calculated molecular thermochemical data

Sl. no.	Compound	Heat of formation (H_f)			Heat of atomization (H_a)		
		Obs.[a]	Calc.	Calc.[d]	Obs.[a]	Calc.	Calc.[j]
1	Methanol	48.07	50.444[b]	66.987	486.93	486.306[e]	455.449
2	Ethanol	56.24	55.383[b]	58.317	770.20	770.372[e]	793.996
3	*n*-Propanol	61.17	60.322[b]	63.060	1050.23	1050.275[e]	1073.670
4	2-Propanol	65.12	64.150[b]	58.358	1054.18	1054.038[e]	1100.314
5	*n*-Butanol	65.79	65.260[b]	67.802	1329.95	1330.178[e]	1353.345
6	2-Methyl propanol	67.84	69.088[b]	69.064	1332.00	1332.091[e]	1353.810
7	2-Butanol	69.98	69.088[b]	63.101	1334.14	1333.941[e]	1379.988
8	2-Methyl-2-propanol	74.72	72.566[b]	60.981	1338.88	1338.658[e]	1399.916
9	*n*-Pentanol	70.66	70.199[b]	72.545	1609.92	1610.080[e]	1633.020
10	2-Pentanol	75.18	74.027[b]	67.843	1614.44	1613.844[e]	1659.663
11	3-Pentanol	75.21	74.027[b]	67.843	1614.47	1613.844[e]	1659.663
12	2-Methyl-1-butanol	72.19	74.027[b]	73.782	1611.45	1612.001[e]	1633.596
13	3-Methyl-1-butanol	72.02	74.027[b]	73.782	1611.23	1612.001[e]	1633.596
14	2-Methyl-2-butanol	75.35	77.504[b]	65.749	1618.33	1618.553[e]	1679.481
15	3-Methyl-2-butanol	79.07	77.855[b]	69.080	1615.61	1615.765[e]	1660.239
16	*n*-Hexanol	75.65	75.138[b]	77.287	1890.01	1889.983[e]	1912.694
17	*n*-Heptanol	79.09	80.077[b]	82.030	2168.55	2169.886[e]	2192.369
18	*n*-Octanol	85.30	85.015[b]	86.772	2449.86	2449.789[e]	2472.044
19	2-Ethyl-1-hexanol	87.31	88.844[b]	88.009	2451.87	2451.710[e]	2472.620
20	*n*-Nonanol	91.12	89.954[b]	91.515	2730.78	2729.692[e]	2751.718
21	*n*-Decanol	94.81	94.893[b]	96.257	3009.57	3009.595[e]	3031.393
22	Ethane	20.24	21.072[c]	21.266	–	–	–
23	Propane	24.82	25.916[c]	26.008	955.49	957.237[f]	935.820
24	*n*-Butane	30.15	30.759[c]	30.751	1236.31	1237.406[f]	1215.495
25	2-Methylpropane	32.15	31.905[c]	31.987	1238.31	1238.564[f]	1216.071
26	*n*-Pentane	35.00	35.603[c]	35.493	1516.65	1517.574[f]	1495.169
27	2-Methylbutane	36.92	36.511[c]	36.730	1518.57	1518.492[f]	1495.745
28	2,2-Dimethylpropane	40.27	38.194[c]	38.133	1521.32	1520.195[f]	1500.210
29	*n*-Hexane	39.96	40.447[c]	40.236	1797.10	1797.743[f]	1774.844
30	2-Methylpentane	41.66	41.354[c]	41.472	1798.80	1798.661[f]	1775.420
31	3-Methylpentane	41.02	41.116[c]	41.472	1798.16	1798.420[f]	1775.420
32	2,2-Dimethylbutane	44.35	42.656[c]	42.876	1801.49	1799.977[f]	1779.884
33	2,3-Dimethylbutane	42.49	42.155[c]	42.709	1799.63	1799.471[f]	1775.996
34	*n*-Heptane	44.89	45.290[c]	44.978	2077.52	2077.911[f]	2054.518
35	2-Methylhexane	46.60	46.198[c]	46.215	2079.23	2078.829[f]	2055.095
36	3-Methylhexane	45.96	45.960[c]	46.215	2078.60	2078.589[f]	2055.095
37	3-Ethylpentane	45.34	45.722[c]	46.215	2077.97	2078.348[f]	2055.095
38	2,2-Dimethylpentane	49.29	47.500[c]	47.618	2081.91	2080.146[f]	2059.559
39	2,3-Dimethylpentane	46.65	46.761[c]	47.452	2080.26	2079.399[f]	2055.670
40	2,4-Dimethylpentane	48.30	47.105[c]	47.452	2080.92	2079.747[f]	2055.670
41	3,3-Dimethylpentane	48.17	47.124[c]	47.618	2080.81	2079.766[f]	2059.559
42	2,2,3-Trimethylbutane	48.96	48.238[c]	48.855	2081.59	2080.893[f]	2060.135
43	*n*-Octane	49.82	50.134[c]	49.721	2357.94	2358.080[f]	2334.193
44	2-Methylheptane	51.50	51.042[c]	50.957	2359.62	2358.998[f]	2334.769
45	3-Methylheptane	50.82	50.804[c]	50.957	2358.94	2358.757[f]	2334.769
46	4-Methylheptane	50.69	50.804[c]	50.957	2358.81	2358.757[f]	2334.769
47	3-Ethylhexane	50.40	50.566[c]	50.957	2358.52	2358.517[f]	2334.769
48	2,2-Dimethylhexane	53.71	52.343[c]	52.361	2361.83	2360.315[f]	2339.234
49	2,3-Dimethylhexane	51.13	51.605[c]	52.194	2359.25	2359.568[f]	2335.345
50	2,4-Dimethylhexane	52.44	51.711[c]	52.194	2360.56	2359.675[f]	2335.345
51	2,5-Dimethylhexane	53.21	51.949[c]	52.194	2361.33	2359.916[f]	2335.345
52	3,3-Dimethylhexane	52.61	51.968[c]	52.361	2360.73	2359.935[f]	2339.234
53	3,4-Dimethylhexane	50.91	51.367[c]	52.194	2359.03	2359.327[f]	2335.345
54	2-Methyl-3-ethylpentane	50.48	51.367[c]	52.194	2358.60	2359.327[f]	2335.345
55	3-Methyl-3-ethylpentane	51.38	51.586[c]	52.361	2359.50	2359.549[f]	2339.234
56	2,2,3-Trimethylpentane	52.61	52.844[c]	53.597	2360.73	2360.821[f]	2339.809
57	2,2,4-Trimethylpentane	53.57	53.251[c]	53.597	2361.69	2361.233[f]	2339.809
58	2,3,3-Trimethylpentane	51.73	52.707[c]	53.597	2359.85	2360.682[f]	2339.809
59	2,3,4-Trimethylpentane	51.97	52.406[c]	53.431	2360.09	2360.378[f]	2335.921
60	2,2,3,3-Tetramethylbutane	53.99	54.290[c]	55.001	2362.11	2362.283[f]	2344.274
61	*n*-Nonane	54.54	54.978[c]	54.463	2638.35	2638.248[f]	2613.868
62	4-Methyloctane	56.19	55.648[c]	55.700	–	–	–
63	2,2-Dimethylheptane	58.83	57.187[c]	57.103	–	–	–
64	2,2,3-Trimethylhexane	57.70	57.688[c]	58.340	–	–	–
65	2,2,4-Trimethylhexane	58.12	57.857[c]	58.340	–	–	–
66	2,2,5-Trimethylhexane	60.53	58.095[c]	58.340	–	–	–

Table 11 continued)

Sl. no.	Compound	Heat of formation (H_f)			Heat of atomization (H_a)		
		Obs.[a]	Calc.	Calc.[d]	Obs.[a]	Calc.	Calc.[j]
67	2,3,3-Trimethylhexane	57.31	57.550[c]	58.340	–	–	–
68	2,3,5-Trimethylhexane	57.97	57.356[c]	58.173	–	–	–
69	2,4,4-Trimethylhexane	57.47	57.719[c]	58.340	–	–	–
70	3,3,4-Trimethylhexane	56.42	57.312[c]	58.340	–	–	–
71	2,2-Dimethyl-3-ethylpentane	55.37	57.450[c]	58.340	–	–	–
72	2,4-Dimethyl-3-ethylpentane	54.48	57.012[c]	58.173	–	–	–
73	3,3-Diethylpentane	55.81	56.048[c]	57.103	2639.05	2639.331[f]	2618.908
74	2,2,3,3-Tetramethylpentane	57.07	58.752[c]	59.743	2640.31	2642.066[f]	2323.948
75	2,2,3,4-Tetramethylpentane	56.81	58.489[c]	59.577	2640.25	2641.800[f]	2620.060
76	2,2,4,4-Tetramethylpentane	58.16	59.403[c]	59.743	2641.44	2642.724[f]	2623.948
77	2,3,3,4-Tetramethylpentane	56.68	58.289[c]	59.577	2640.07	2641.597[f]	2620.060
78	*n*-Decane	59.67	59.822[c]	59.206	–	–	–
79	3,3,5-Trimethylheptane	62.22	62.325[c]	63.082	–	–	–
80	2,2,3,3-Tetramethylhexane	62.08	63.596[c]	64.486	–	–	–
81	2,2,5,5-Tetramethylhexane	67.29	64.246[c]	64.486	–	–	–
82	*n*-Undecane	64.58	64.665[c]	63.948	–	–	–
83	2-Methyldecane	66.06	65.573[c]	65.185	–	–	–
84	*n*-Dodecane	69.49	69.509[c]	68.691	–	–	–
85	2,2,4,4,6-Pentamethylheptane	72.10	74.459[c]	75.207	–	–	–
86	2,2,4,6,6-Pentamethylheptane	75.45	74.604[c]	75.207	–	–	–
87	*n*-Hexadecane	89.21	88.884[c]	87.661	–	–	–
88	2-Methylpentadecane	90.87	89.791[c]	88.898	–	–	–
89	Ethylene	–	–	–	537.75	539.487[g]	567.781
90	Propylene	–	–	–	820.42	819.799[g]	844.915
91	1-Butene	–	–	–	1100.60	1100.110[g]	1134.310
92	*trans*-2-Butene	–	–	–	1103.39	1100.110[g]	1112.771
93	2-Methylpropene	–	–	–	1104.66	1105.399[g]	1127.044
94	1-Pentene	–	–	–	1380.83	1380.422[g]	1413.985
95	*trans*-2-Pentene	–	–	–	1383.43	1380.422[g]	1402.166
96	2-Methyl-1-butene	–	–	–	1384.05	1385.711[g]	1413.235
97	3-Methyl-1-butene	–	–	–	1382.11	1385.711[g]	1418.868
98	2-Methyl-2-butene	–	–	–	1385.62	1385.711[g]	1390.813
99	1-Hexene	–	–	–	1660.55	1660.733[g]	1693.659
100	*trans*-2-Hexene	–	–	–	1663.48	1660.733[g]	1681.841
101	*trans*-3-Hexene	–	–	–	1663.61	1660.733[g]	1691.450
102	2-Methyl-1-pentene	–	–	–	1664.79	1666.023[g]	1692.910
103	3-Methyl-1-pentene	–	–	–	1662.43	1666.023[g]	1698.543
104	4-Methyl-1-pentene	–	–	–	1662.85	1666.023[g]	1694.235
105	2-Methyl-2-pentene	–	–	–	1666.58	1666.023[g]	1680.208
106	3-Methyl-*trans*-2-pentene	–	–	–	1665.69	1666.023[g]	1677.004
107	4-Methyl-*trans*-2-pentene	–	–	–	1665.30	1666.023[g]	1686.724
108	2-Ethyl-1-butene	–	–	–	1663.99	1666.023[g]	1699.316
109	2,3-Dimethyl-1-butene	–	–	–	1665.79	1671.312[g]	1696.358
110	3,3-Dimethyl-1-butene	–	–	–	1665.11	1663.824[g]	1705.438
111	2,3-Dimethyl-2-butene	–	–	–	1667.02	1671.312[g]	1667.198
112	1-Heptene	–	–	–	1940.51	1941.045[g]	1973.334
113	5-Methyl-1-hexene	–	–	–	1941.40	1946.334[g]	1973.910
114	3-Methyl-*trans*-3-hexene	–	–	–	1944.06	1946.334[g]	1966.289
115	2,4-Dimethyl-1-pentene	–	–	–	1945.73	1951.623[g]	1973.160
116	4,4-Dimethyl-1-pentene	–	–	–	1945.14	1944.135[g]	1978.264
117	2,4-Dimethyl-2-pentene	–	–	–	1946.90	1951.623[g]	1964.766
118	4,4-Dimethyl-*trans*-2-pentene	–	–	–	1946.92	1944.135[g]	1973.293
119	3-Methyl-2-ethyl-1-butene	–	–	–	1994.71	1951.623[g]	1982.439
120	2,3,3-Trimethyl-1-butene	–	–	–	1946.13	1949.424[g]	1982.043
121	1-Octene	–	–	–	2220.21	2221.356[g]	2253.009
122	2,2,-Dimethyl-*trans*-3-hexene	–	–	–	2226.53	2224.447[g]	2262.578
123	2-Methyl-3-ethyl-1-pentene	–	–	–	2224.77	2231.935[g]	2255.707
124	2,4,4-Trimethyl-2-pentene	–	–	–	2225.87	2229.736[g]	2251.335
125	1-Decene	–	–	–	2780.48	2781.979[g]	2812.358
126	Dimethyl ether	–	–	–	757.95	758.077 [h]	652.171
127	Methyl ethyl ether	–	–	–	1040.78	1041.590[h]	973.377
128	Diethyl ether	–	–	–	1324.42	1325.104[h]	1294.583
129	Methyl-*n*-propyl ether	–	–	–	1320.98	1321.463[h]	1253.052
130	Methyl-*sec*-propyl ether	–	–	–	1324.40	1324.142[h]	1272.074
131	Methyl-*tert*-butyl ether	–	–	–	1608.86	1606.487[h]	1567.037
132	Di-*n*-propyl ether	–	–	–	1884.11	1884.850[h]	1853.932

Table 11 (conttinued)

Sl. no.	Compound	Heat of formation (H_f)			Heat of atomization (H_a)		
		Obs.[a]	Calc.	Calc.[d]	Obs.[a]	Calc.	Calc.[j]
133	Di-*sec*-propyl ether	–	–	–	1890.46	1890.207[h]	1891.976
134	Isopropyl-*tert*-butyl ether	–	–	–	2174.86	2172.562[h]	2187.050
135	Di-*n*-butyl ether	–	–	–	2444.28	2444.595[h]	2413.281
136	Di-*sec*-butyl ether	–	–	–	2450.72	2449.380[h]	2451.215
137	Di-*tert*-butyl ether	–	–	–	2451.56	2454.924[h]	2482.124
138	Methanethiol	–	–	–	450.35	450.985[i]	343.999
139	Ethanethiol	–	–	–	731.05	731.099[i]	715.131
140	1-Propanethiol	–	–	–	1011.29	1011.212[i]	994.805
141	2-Propanethiol	–	–	–	1013.29	1013.640[i]	1036.029
142	1-Butanethiol	–	–	–	1291.23	1291.326[i]	1274.480
143	2-Butanethiol	–	–	–	1293.34	1293.249[i]	1315.703
144	2-Methyl-1-propanethiol	–	–	–	1293.42	1293.249[i]	1275.056
145	2-Methyl-2-propanethiol	–	–	–	1296.37	1296.818[i]	1344.137
146	1-Pentanethiol	–	–	–	1571.59	1571.439[i]	1554.155
147	3-Methyl-1-butanethiol	–	–	–	1572.76	1573.362[i]	1554.731
148	2-Methyl-1-butanethiol	–	–	–	1575.65	1572.858[i]	1554.731
149	1-Hexanethiol	–	–	–	1851.34	1851.552[i]	1833.829
150	1-Heptanethiol	–	–	–	2131.28	2131.666[i]	2113.504
151	1-Decanethiol	–	–	–	2971.50	2972.006[i]	2952.528

[a] Obs. = Observed (Ref. [8]; Calc. = Calculated, [b] From Eq.15, [c] From Eq. 19, [d] From Eq. 29, [e] From Eq. 37, [f] From Eq. 41, [g] From Eq. 50, [h] From Eq. 57, [i] From Eq. 65, [j] From Eq. 74

Table 12 Summary of PRESS statistics

Eq. no.	15	19	29	37	41	50	57	65	74
[a] PRESS	79.16	83.38	1,322.42	14.94	40.84	2,765.59	83.73	14.21	134,373.8
Average [b] Pres	1.57	0.83	2.12	0.54	0.74	3.92	1.77	0.64	28.25
[c] Q^2	0.967	0.991	0.929	1.000	1.000	1.000	1.000	1.00	0.996
[d] SDEP	1.94	1.94	3.88	0.84	0.96	8.65	2.64	1.01	32.40

[a] PRESS=predicted residual sum of squares
[b] Pres=predicted residuals
[c] Q^2=cross-validated R^2
[d] SDEP=standard deviation of error of predictions

properties of more diverse functional chemical compounds to establish the utility of TAU scheme in QSPR studies.

Acknowledgement The authors are grateful to Sri Dipak Kumar Pal for guidance and inspiration.

References

1. Trinajstic N (1983) Chemical graph theory. CRC Press, Boca Raton, FL
2. Hansen PJ, Jurs PC (1988) J Chem Educ 65:574–580
3. Rouvray DH (1995) J Mol Struct (Theochem) 336:101–104
4. Basak SC, Gute BD (1997) SAR QSAR Environ Res 7:1–21
5. Devillers J, Balaban AT (1999) Topological indices and related descriptors in QSAR and QSPR. Gordon and Breach, The Netherlands
6. Motoc I, Balaban AT (1981) Rev Roum Chim 26:593–600
7. Balasubramanian K, Basak SC (1998) J Chem Inf Comput Sci 38:367–373
8. Kier LB, Hall LH (1976) Molecular connectivity in chemistry and drug research. Academic Press, New York
9. Roy K, Pal DK, De AU, Sengupta C (1999) Indian J Chem 38B:664–671
10. Roy K, Pal DK, De AU, Sengupta C (2001) Indian J Chem 40B:129–135
11. Kier LB, Hall LH (1983) J Pharm Sci 72:1170–1173
12. Kier LB, Hall LH (1986) Molecular connectivity in structure–activity analysis. Research Studies Press, Letchworth, UK
13. Shannon CE (1948) Bell Syst Tech J 27:379–423
14. Shannon CE, Weaver W (1949) The mathematical theory of communications. University of Illinois Press, Urbana Ill.
15. Kier LB (1980) J Pharm Sci 69:807–810
16. Pal DK, Sengupta C, De AU (1988) Indian J Chem 27B:734–739
17. Pal DK, Sengupta C, De AU (1989) Indian J Chem 28B:261–267
18. Pal DK, Sengupta M, Sengupta C, De AU (1990) Indian J Chem 29B:451–454
19. Pal DK, Purkayastha SK, Sengupta C, De AU (1992) Indian J Chem 31B:109–114
20. The GW-BASIC programs *RRR98*, *KRPRES1* and *KRPRES2* were developed by Kunal Roy (1998) and standardized using known data sets
21. Snedecor GW, Cochran WG (1967) Statistical methods. Oxford & IBH Publishing, New Delhi, pp 381–418
22. Kier LB, Hall LH (1992) Atom description in QSAR models: Development and use of an atom level index. In: Testa B (ed) Advances in drug research, vol 22. Academic Press, New York, pp 1–38
23. Wold S, Eriksson L (1995) Validation tools. In: van de Waterbeemd H (ed) Chemometric methods in molecular design. VCH, Weinheim, pp 309–318