Internet Electronic Journal of Molecular Design **2003**, *2*, 621–641 ISSN 1538–6414
BioChem Press http://www.biochempress.com

Internet Electronic Journal of
Molecular Design

September 2003, Volume 2, Number 9, Pages 621–641

Editor: Ovidiu Ivanciuc

Special issue dedicated to Professor Nenad Trinajstić on the occasion of the 65th birthday
Part 3

Guest Editor: Douglas J. Klein

Eigenvalues of the Bond Adjacency Matrix Extended to Application in Physicochemical Properties of Alkanes

Chenzhong Cao, Lihui Jiang, and Hua Yuan

Department of Chemistry, Xiangtan Normal University, Xiangtan, 411201, People's Republic of China

Received: May 3, 2003; Revised: June 6, 2003; Accepted: June 10, 2003; Published: September 30, 2003

Citation of the article:
C. Cao, L. Jiang, and H. Yuan, Eigenvalues of the Bond Adjacency Matrix Extended to Application in Physicochemical Properties of Alkanes, *Internet Electron. J. Mol. Des.* **2003**, *2*, 621–641, http://www.biochempress.com.

Inter*net* Electronic Journal of Molecular Design

BioChem Press
http://www.biochempress.com

Eigenvalues of the Bond Adjacency Matrix Extended to Application in Physicochemical Properties of Alkanes[#]

Chenzhong Cao,* Lihui Jiang, and Hua Yuan

Department of Chemistry, Xiangtan Normal University, Xiangtan, 411201, People's Republic of China

Received: May 3, 2003; Revised: June 6, 2003; Accepted: June 10, 2003; Published: September 30, 2003

Internet Electron. J. Mol. Des. **2003**, *2* (9), 621–641

Abstract

Motivation. Eigenvalue of bonding orbital–connection matrix is a good descriptor for expressing the relative bond energies of C–C and C–H bonds in alkane. Its application can be extended to the physicochemical properties of alkanes.

Method. Based on polarizability effect index (*PEI*), the bond adjacency matrix $\mathbf{B}_{\mathbf{CH}}$ for each C–H bond and orbital overlapping matrix $\mathbf{B}_{\mathbf{CC}}$ for whole carbon skeleton in alkane molecule, respectively, were built. The eigenvalues $X_{1\mathrm{CH}}$, $X_{2\mathrm{CH}}$ for every $\mathbf{B}_{\mathbf{CH}}$ and $X_{1\mathrm{CC}}$, …, $X_{(n-1)\mathrm{CC}}$, $X_{n\mathrm{CC}}$, …, $X_{2(n-1)\mathrm{CC}}$ for $\mathbf{B}_{\mathbf{CC}}$ were obtained. Then the parameters $SX_{1\mathrm{CH}}$ and $SX_{2\mathrm{CH}}$ were calculated by summing eigenvalues $X_{1\mathrm{CH}}$ and $X_{2\mathrm{CH}}$ of all $\mathbf{B}_{\mathbf{CH}}$, and $SX_{1\mathrm{CC}}$ and $SX_{2\mathrm{CC}}$ were computed by summing eigenvalues from $X_{1\mathrm{CC}}$ to $X_{(n-1)\mathrm{CC}}$ and from $X_{n\mathrm{CC}}$ to $X_{2(n-1)\mathrm{CC}}$, respectively. In addition, another parameter V_{ij} was developed for C–C bond and summed to get parameter SV_{ij}. This set of descriptors combined with odd–even index (*OEI*) and $N^{2/3}$ (*N* is the number of carbon atoms) can express most of the important molecular structural information of alkanes.

Results. The correlations between 10 physicochemical properties (including boiling point Bp, critical temperature Tc, critical volume Vc, critical pressure Pc, density D, refraction index n_D, vapor pressure log Pv, heat capacities Cp, heats of vaporization Hv and chromatographic retention index Ipt) and the five parameters $SX_{1\mathrm{CH}}$, $SX_{1\mathrm{CC}}$, SV_{ij}, *OEI*, and $N^{2/3}$ were carried out, and good results were obtained with Bp ($r = 0.9987$, $s = 3.06$ ˚C), Tc ($r = 0.9991$, $s = 4.03$ ˚C), Vc ($r = 0.9995$, $s = 7.66$ $\mathrm{cm^3 \cdot mol^{-1}}$), Pc ($r = 0.9941$, $s = 0.0771$ Mpa), D ($r = 0.9906$, $s = 0.0048$ $\mathrm{g \cdot cm^{-3}}$), n_D ($r = 0.9499$, $s = 0.0055$), logPv ($r = 0.9974$, $s = 0.0959$), Cp ($r = 0.9997$, $s = 2.02 \mathrm{J \cdot mol^{-1} \cdot K^{-1}}$), Hv ($r = 0.9994$, $s = 0.66$ $\mathrm{kJ \cdot mol^{-1}}$) and Ipt ($r = 0.9986$, $s = 8.09$) respectively.

Conclusions. The model equations developed by present paper can be used for estimating and predicting the 10 physicochemical properties of alkanes. This method may be extended to QSPR of other compounds.

Keywords. Polarizability effect index PEI; adjacency matrix; eigenvalue; physicochemical property; alkane; quantitative structure–property relationships; QSPR.

[#] Dedicated to Professor Nenad Trinajstić on the occasion of the 65th birthday.

* Correspondence author; phone: 86–0732–8291336; fax: 86–0732–8291001; E–mail: czcao@xtnu.edu.cn.

1 INTRODUCTION

With the development of experimental chemistry, a great amount of new compounds are synthesized every year. However, a large part of these compounds are not tested for fundamental or relevant thermodynamic and physicochemical properties or biological activities, which still remain unknown due to unavailability or not easily handling (toxic, odorous, etc.). A procedure able to predict, within a reasonable error margin, physicochemical properties and biological activities for untested compounds is required to evaluate these molecular features in a fast and inexpensive way [1]. In recent years, numerous QSPR models have been introduced into calculating the physicochemical properties with various numerical descriptors of chemical structure [2–4]. These descriptors can be grouped into three classes based on molecular descriptors used in the analysis: (*a*) based on experimentally determined physicochemical properties such as partition coefficient, chromatographic retention time, melting point, boiling point, (*b*) based on group–contribution theory, (*c*) based on calculated molecular descriptors solely from its structure, such as molecular surface area, topological indices. Topological indices are numerical representations of the chemical structure computed on the basis of molecular graph [5].

Many topological indices have been proposed since the pioneering work of Wiener. Molecular connectivity indices are the most successful, and widely applied in quantitative structure–properties and activities relationships (QSPR–QSAR) [6]. Most of the existing topological indices are related either to a vertex adjacency matrix, known simply as adjacency matrix **A**, or to a distance matrix **D** in graph [7]. A lot of studies on the determinants, eigenvalues and eigenvector of the adjacency matrixes were published [8–11]. In our previous work, we have proposed the polarizability effect index (PEI), and used it to correlate with water solubility, octanol/water partition coefficient, and boiling point for alkanes and alcohols [12–13], and further with ionization potential for haloalkanes, amines, alcohols, ethers, alkenes and alkanes [14–15]. The obtained models have good predictive power. In our recent work [16], the PEI value of alkyl was used as the main diagonal element to build matrix $\mathbf{CM}_{ij}$, and the eigenvalues $X_{1\mathrm{CC}}$, $X_{1\mathrm{CH}}$ etc. of each matrix were calculated for C–C bond and for C–H bond respectively, also a steric effect parameter S_{ij} was proposed for C–C bond. Using $X_{1\mathrm{CC}}$, $X_{1\mathrm{CH}}$ and S_{ij} as descriptors, good correlations with the Bond Dissociation Energies (BDEs) of the C–C and C–H bond were obtained for alkanes. In addition, the Heat of Atomization (HA) and Heat of Formation in Gas (HFG^0) of alkanes can be estimated well with the parameters$\Sigma X_{1\mathrm{CC}}$, $\Sigma X_{1\mathrm{CH}}$ and ΣS_{ij}. In the present report, the matrix $\mathbf{B}_{\mathbf{CC}}$ was modified, and the eigenvalues of $\mathbf{B}_{\mathbf{CH}}$ and $\mathbf{B}_{\mathbf{CC}}$ were obtained and extended their application to 10 physicochemical properties including boiling point (Bp), critical temperature (Tc), critical volume (Vc), critical pressure (Pc), density (D) at 25˚C, refraction index (n_D) at 20˚C, vapor pressure (log Pv), heat capacities (Cp), heats of vaporization (Hv) and chromatographic retention index (Ipt). Good correlations were obtained.

2 MATERIALS AND METHODS

In the preceding paper [12], the *PEI* of alkyl groups has been developed and calculated. It quantitatively indicates the relative proportion polarizability effect. The *PEI* values of some normal alkyls and the increments Δ*PEI* are cited in Table 1.

Table 1. *PEI* and Δ*PEI* values of normal alkyl $H(CH_2)_N$

N	*PEI*	Δ*PEI*	*N*	*PEI*	Δ*PEI*
1	1.0000	1.0000	6	1.2350	0.0095
2	1.1405	0.1405	7	1.2414	0.0064
3	1.1887	0.0481	8	1.2461	0.0047
4	1.2122	0.0235	9	1.2498	0.0037
5	1.2260	0.0138	10	1.2527	0.0029

2.1 SX_{1CH} and SX_{2CH}

Now, let us consider a molecule of alkane, for example: 2–methylbutane. If a H atom connects with the *i*th carbon atom (C_i), when the H–C_i bond is broken, two radicals H · and R_i · will be formed (Figure 1):

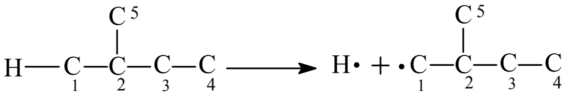

Figure 1. The breaking of the H–C_1 bond.

According to the calculation method of PEI of alkyl in paper [12,16] and values in Table 1, we can calculate the *PEI* for two radicals above as follows:

H: $PEI_H = 0$
R_1: $PEI_1 = 1.2122+0.0481 = 1.2603$

Then, PEI_H and PEI_1 were used as the main diagonal elements to build the bonding adjacency matrix $\mathbf{B_{CH}}$ of H–C_1 bond:

$$\mathbf{B}_{\mathbf{CH}} = \begin{bmatrix} \mathrm{PEI_H} & 1 \\ 1 & \mathrm{PEI_1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1.2603 \end{bmatrix}$$

The off–diagonal element “1” in matrix means that H atom and C_1 are connected with each other, *i.e.*, they are adjacent. Solving matrix $\mathbf{B_{CH}}$ by computer, we got two eigenvalues $X_{1CH} = -0.5518$ and $X_{2CH} = 1.8121$ (let $X_{1CH} < X_{2CH}$). The eigenvalues of every H–C bond adjacency matrix in a molecule are also calculated with the same method. Finally, taking sum of X_{1CH} and X_{2CH} of all $\mathbf{B_{CH}}$ respectively, we got two parameters SX_{1CH} and SX_{2CH}, in other words, let $SX_{1CH} = \Sigma X_{1CH}$ and $SX_{2CH} = \Sigma X_{2CH}$. For 2–methylbutane, there are:

$$SX_{1CH} = \Sigma X_{1CH} = (6)\times(-0.5518)+(-0.5061)+(2)\times(-0.5255)+ (3)\times(-0.5576) = -6.5407$$

$$SX_{2CH} = \Sigma X_{2CH} = (6)\times(1.8121)+1.9758+(2)\times(1.9028)+(3)\times(1.7933) = 22.0339$$

2.2 SX_{1CC} and SX_{2CC}

As far as the C–C bond is concerned, the bond adjacency matrix for every C–C bond would be built with the same method as that of matrix $\mathbf{B}_{CH}$. However, more overall consideration should be taken into molecular structure. It is known that every C–C bond is formed by the overlapping of two sp^3 hybrid orbitals of the adjacent carbon atoms in an alkane molecule, and by which, n–1 C–C bonds will be formed for the carbon skeleton consisting of n carbon atoms. Therefore, we could construct an sp^3 hybrid orbital overlapping matrix $\mathbf{B}_{CC}$ for carbon skeleton of alkane molecule. Still take 2–methylbutane for example (Figure 2).

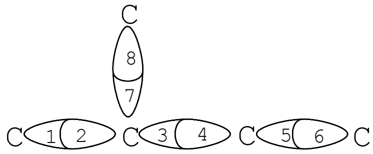

Figure 2. The overlap of sp^3 orbitals in the carbon skeleton of 2–methylbutane.

First step: the sp^3 orbital in Figure 2 is numbered at random.

Second step: let $a_{ij} = \begin{cases} 0 & \text{orbital } i \text{ and } j \text{ do not overlap} \\ 1 & \text{orbital } i \text{ and } j \text{ overlap with each other} \\ PEI_i & \text{when } i = j \end{cases}$

Then the skeleton orbital–overlapping matrix $\mathbf{B}_{CC}$ was built:

$$\mathbf{B}_{CC} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} & a_{28} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} & a_{38} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} & a_{48} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} & a_{58} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} & a_{68} \\ a_{71} & a_{72} & a_{73} & a_{74} & a_{75} & a_{76} & a_{77} & a_{78} \\ a_{81} & a_{82} & a_{83} & a_{84} & a_{85} & a_{86} & a_{87} & a_{88} \end{bmatrix} = \begin{bmatrix} PEI_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & PEI_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & PEI_3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & PEI_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & PEI_5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & PEI_6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & PEI_7 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & PEI_8 \end{bmatrix}$$

where PEI_i is the polarizability effect index of alkyl containing the ith sp^3 hybrid orbital when the interested C–C bond is broken. For instance, if the bond formed by orbital 3 and orbital 4 is disconnected, two alkyls R_3 and R_4 will be obtained (Figure 3):

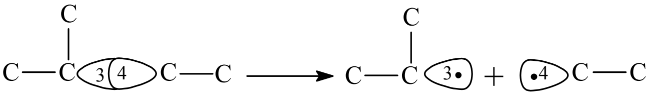

Figure 3. The disconnection of orbitals 3 and 4.

Then the PEI_3 and PEI_4 for R_3 and R_4 were calculated, respectively.

R_3: $PEI_{3'} = 1.1405+0.1405 = 1.2810$

R_4: $PEI_4 = 1.1405$

Therefore, the matrix $\mathbf{B}_{\mathbf{CC}}$ of 2–methylbutane is:

$$\mathbf{B}_{\mathbf{CC}} = \begin{bmatrix} 1.0000 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1.3292 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.2810 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1.1405 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.2368 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1.000 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1.3292 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1.000 \end{bmatrix}$$

In general, there are $2(n-1)$ eigenvalues for the matrix $\mathbf{B}_{\mathbf{CC}}$ of alkane bearing n carbon atoms ($n-1$ C–C bonds). After sorted eigenvalues from small to large, we assigned the eigenvalues from 1 to $2(n-1)$ as $X_{1\mathrm{CC}}$ to $X_{2(\mathrm{n}-1)\mathrm{CC}}$, then took sum of $X_{1\mathrm{CC}}$ to $X_{(\mathrm{n}-1)\mathrm{CC}}$ and X_{nCC} to $X_{2(\mathrm{n}-1)\mathrm{CC}}$, respectively, and got parameters $SX_{1CC} = \sum_{i=1}^{n-1} X_{iCC}$ and $SX_{2CC} = \sum_{i=n}^{2(n-1)} X_{iCC}$. For 2–methylbutane, solutions to the matrix $\mathbf{B}_{\mathbf{CC}}$ are eigenvalues: 0.1511, 2.1781, 2.2132, 0.2083, 2.1254, 0.1114, 2.1781 and 0.1511, then sorted them in order of $X_{1\mathrm{CC}}$, $X_{2\mathrm{CC}}$, …, $X_{8\mathrm{CC}}$ as follows: 0.1114, 0.1511, 0.1511, 0.2083, 2.1254, 2.1781, 2.1781, 2.2132. Therefore,

$$SX_{1CC} = \sum_{1}^{4} X_{iCC} = 0.1114+(2)\times(0.1511)+0.2083 = 0.6219$$

$$SX_{2CC} = \sum_{5}^{8} X_{iCC} = 2.1254+ (2)\times(2.1781)+2.2132 = 8.6948.$$

2.3 SV_{ij}

Steric effect plays an important role on thermodynamic properties of alkanes. Here, a parameter SV_{ij} was developed to scale steric effect of alkane molecules. We defined $SV_{ij} = \Sigma V_{ij}$, and $V_{ij} = (V_i \times V_j)^2$, V_i and V_j are the vertex degree of vertex *i* and vertex *j* in a C_i–C_j bond, respectively. As for 2–methylbutane (Figure 4), there are 4 carbon–carbon bonds C_1–C_2, C_2–C_3, C_3–C_4 and C_2–C_5, their V_{ij} are $V12 = (1\times3)^2 = 9$, $V23 = (3\times2)^2 = 36$, $V34 = (2\times1)^2 = 4$ and $V25 = (3\times1)^2 = 9$ respectively. Thus its $SV_{ij} = \Sigma V_{ij} = 9 + 36 + 4 + 9 = 58$.

Table 2. Values of parameters *SX1CH*, *SX1CC*, SV_{ij}, OEI and $N^{2/3}$ for 160 alkanes

No	Compound	SX_{1CH}	SX_{1CC}	SX_{2CH}	SX_{2CC}	SV_{ij}	OEI^{a}	$N^{2/3}$
1	methane	–2.4720	0.0000	6.4720	0.0000	0	0.0000	1.0000
2	ethane	–3.4854	0.0000	10.3284	2.0000	1	2.0000	1.5874
3	propane	–4.5074	0.1356	14.1816	4.1454	8	3.5000	2.0801
4	butane	–5.5244	0.3203	18.1136	6.3381	24	5.2222	2.5198
5	2–methylpropane	–5.5327	0.3921	18.0855	6.4509	27	4.5000	2.5198
6	pentane	–6.5340	0.5296	22.0556	8.5532	40	6.8194	2.9240
7	2–methylbutane	–6.5407	0.6219	22.0339	8.6948	58	6.4444	2.9240
8	2,2–dimehtylpropane	–6.5532	0.7552	21.9720	8.3912	64	5.0000	2.9240
9	hexane	–7.5404	0.7533	26.0200	10.7815	56	8.4967	3.3019
10	2–methylpentane	–7.5453	0.8540	26.0059	10.9341	74	7.9167	3.3019
11	3–methylpentane	–7.5440	0.8753	26.0142	10.9621	89	8.2639	3.3019
12	2,2–dimehtylbutane	–7.5706	1.0264	25.9041	11.2297	116	7.1667	3.3019
13	2,3–dimethylbutane	–7.5496	0.9654	25.9922	11.1062	117	7.8889	3.3019
14	heptane	–8.5424	0.9866	30.0004	13.0182	72	10.1183	3.6593
15	2–methylhexane	–8.5469	1.0914	29.9919	13.1763	90	9.6739	3.6593
16	3–methylhexane	–8.5434	1.1214	30.0067	13.2150	105	9.8161	3.6593
17	2,2–dimethylpentane	–8.5515	1.2794	29.9587	13.4950	132	8.5139	3.6593
18	2,3–dimethylpentane	–8.5454	1.2403	29.9991	13.3989	148	9.5833	3.6593
19	2,4–dimethylpentane	–8.5486	1.1998	29.9866	13.3406	108	8.8889	3.6593
20	3,3–dimethylpentane	–8.5480	1.3200	29.9730	13.5530	168	9.2083	3.6593
21	3–ethylpentane	–8.5400	1.1517	30.0222	13.2537	120	9.9583	3.6593
22	2,2,3–trimethylbutane	–8.5526	1.4076	29.9496	13.6997	210	8.8333	3.6593
23	octane	–9.5424	1.2266	33.9938	15.2610	88	11.7808	4.0000
24	2–methylheptane	–9.5454	1.3335	33.9886	15.4222	106	11.2400	4.0000
25	3–methylheptane	–9.5414	1.3679	33.7892	15.4661	121	11.5178	4.0000
26	4–methylheptane	–9.5392	1.3765	34.0162	15.4769	121	11.3128	4.0000
27	2,2–dimethylhexane	–9.5480	1.5296	33.9660	15.7520	148	10.3511	4.0000
28	2,3–dimethylhexane	–9.5407	1.4990	34.0139	15.6665	164	11.2156	4.0000
29	2,4–dimethylhexane	–9.5432	1.4798	34.0074	15.6364	139	10.8683	4.0000
30	2,5–dimethylhexane	–9.5482	1.4418	33.9846	15.5864	124	10.9311	4.0000
31	3,3–dimethylhexane	–9.5405	1.5865	33.9951	15.8323	184	10.6356	4.0000
32	3,4–dimethylhexane	–9.5370	1.5284	34.0286	15.7056	179	11.3578	4.0000
33	3–ethylhexane	–9.5350	1.4112	34.0364	15.5202	136	11.5906	4.0000
34	2,2,3–trimethylpentane	–9.5556	1.7023	33.9342	16.0195	241	10.4028	4.0000
35	2,2,4–trimethylpentane	–9.5487	1.6445	33.9665	15.9292	166	9.3611	4.0000
36	2,3,3–trimethylpentane	–9.5399	1.7215	33.9939	16.0494	262	10.7500	4.0000
37	2,3,4–trimethylpentane	–9.5409	1.6247	34.0137	15.8626	207	10.7778	4.0000
38	3–ethyl–2–methylpentane	–9.5347	1.5347	34.0367	15.7159	179	11.1528	4.0000
39	3–ethyl–3–methylpentane	–9.5358	1.6351	34.0176	15.9014	220	11.1250	4.0000
40	2,2,3,3–tetramethylbutane	–9.5454	1.8843	33.9426	16.3553	352	10.0000	4.0000
41	nonane	–10.5392	1.4718	33.9976	17.5080	104	13.4120	4.3267
42	2–methyloctane	–10.5423	1.5800	37.9947	17.6714	122	12.9433	4.3267
43	3–methyloctane	–10.5373	1.6165	38.0196	17.7183	137	13.1247	4.3267
44	4–methyloctane	–10.5432	1.6299	38.0322	17.7339	137	13.0553	4.3267
45	2,2–dimethylheptane	–10.5425	1.7802	37.9799	18.0076	164	11.8617	4.3267
46	2,3–dimethylheptane	–10.5341	1.7312	38.0331	17.9007	180	12.8617	4.3267
47	2,4–dimethylheptane	–10.5347	1.7439	38.0340	17.9075	155	12.3094	4.3267
48	2,5–dimethylheptane	–10.5449	1.8454	37.9993	17.9674	146	12.7194	4.3267
49	2,6–dimethylheptane	–10.5436	1.8132	37.9934	17.8876	154	12.3061	4.3267
50	3,3–dimethylheptane	–10.5337	1.8453	38.0201	18.0986	200	12.4172	4.3267
51	3,4–dimethylheptane	–10.5288	1.7963	38.0615	17.9824	195	12.9344	4.3267
52	3,5–dimethylheptane	–10.5330	1.7692	38.0476	17.9410	206	12.7922	4.3267
53	4,4–dimethylheptane	–10.5300	1.8620	38.0346	18.1210	200	12.0072	4.3267
54	3–ethylheptane	–10.5296	1.6668	38.0571	17.7806	152	13.2367	4.3267
55	4–ethylheptane	–10.5265	1.6803	38.0700	17.7960	152	13.1672	4.3267
56	2,2,3–trimethylhexane	–10.5326	1.9732	38.0240	18.3023	257	12.1150	4.3267

Table 2. (Continued)

No	Compound	SX_{1CH}	SX_{1CC}	SX_{2CH}	SX_{2CC}	SV_{ij}	OEI [a]	$N^{2/3}$
57	2,2,4–trimethylhexane	–10.5370	1.9369	38.0105	18.2401	170	11.4206	4.3267
58	2,2,5–trimethylhexane	–10.5434	1.8917	37.9811	18.1779	182	11.6883	4.3267
59	2,3,3–trimethylhexane	–10.5284	2.0003	38.0383	18.3438	278	12.2572	4.3267
60	2,3,4–trimethylhexane	–10.5280	1.9256	38.0660	18.1846	238	12.6322	4.3267
61	2,3,5–trimethylhexane	–10.5346	1.8696	38.0369	18.1035	198	12.3478	4.3267
62	2,4,4–trimethylhexane	–10.5333	1.9640	38.0252	18.2819	218	11.5628	4.3267
63	3,3,4–trimethylhexane	–10.5253	2.0291	38.0531	18.3839	293	12.3994	4.3267
64	3–ethyl–2–methylhexane	–10.5265	1.8099	38.0749	17.9978	195	12.8650	4.3267
65	4–ethyl–2–methylhexane	–10.5409	1.7826	38.0163	17.9566	170	12.7228	4.3267
66	3–ethyl–3–methylhexane	–10.5239	1.9151	38.0620	18.1955	236	12.6322	4.3267
67	4–ethyl–2–methylhexane	–10.5231	1.8390	38.0895	18.0376	210	13.0072	4.3267
68	2,2,3,3–tetramethylpentane	–10.5260	2.2165	38.0158	18.7325	404	11.7917	4.3267
69	2,2,3,4–tetramethylpentane	–10.5625	2.1052	37.9183	18.5116	300	11.4722	4.3267
70	2,2,4,4–tetramethylpentane	–10.5396	2.1074	37.9756	18.5464	224	9.7083	4.3267
71	2,3,3,4–tetramethylpentane	–10.5252	2.1414	38.0432	18.5734	356	12.1667	4.3267
72	2,2–dimethyl–3–ethylpentane	–10.5258	2.0193	38.0546	18.3641	272	11.8472	4.3267
73	2,3–dimethyl–3–ethylpentane	–10.5208	2.0373	38.0674	18.3978	314	12.5417	4.3267
74	2,4–dimethyl–3–ethylpentane	–10.5428	1.9431	38.0816	18.2064	238	12.2222	4.3267
75	3,3–diethylpentane	–10.5164	1.9716	38.0908	18.2764	272	12.9167	4.3267
76	decane	–11.5348	1.7204	42.0098	19.7590	120	15.0680	4.6416
77	2–methylnonane	–11.5372	1.8295	42.0098	19.9235	138	14.5433	4.6416
78	3–methylnonane	–11.5315	1.8670	42.0360	19.9712	153	14.7968	4.6416
79	4–methylnonane	–11.5280	1.8833	42.0529	19.9908	163	14.6309	4.6416
80	5–methylnonane	–11.5268	1.8878	42.0584	19.9958	153	14.7665	4.6416
81	2,2–dimethyloctane	–11.5370	2.0324	41.9990	20.2638	180	13.6058	4.6416
82	2,3–dimethyloctane	–11.5280	2.0092	42.0577	20.1860	196	14.5094	4.6416
83	2,4–dimethyloctane	–11.5368	2.0018	42.0309	20.1701	181	14.0927	4.6416
84	2,5–dimethyloctane	–11.5282	1.9952	42.0546	20.1600	181	14.2977	4.6416
85	2,6–dimethyloctane	–11.5330	1.9780	42.0371	20.1389	181	14.2316	4.6416
86	2,7–dimethyloctane	–11.5382	1.9391	42.0082	20.0889	156	14.1466	4.6416
87	3,3–dimethyloctane	–11.5263	2.1024	42.0471	20.3608	216	13.9686	4.6416
88	3,4–dimethyloctane	–11.5196	2.0581	42.0931	20.2492	211	14.6214	4.6416
89	3,5–dimethyloctane	–11.5221	2.0393	42.0866	20.2188	186	14.2741	4.6416
90	3,6–dimethyloctane	–11.5266	2.0185	42.0668	20.1913	186	14.5486	4.6416
91	4,4–dimethyloctane	–11.5196	2.1275	42.0714	20.3935	216	13.8297	4.6416
92	4,5–dimethyloctane	–11.5168	2.0709	42.1058	20.2653	221	14.5519	4.6416
93	3–ethyloctane	–11.5227	1.9220	42.0797	20.0391	168	14.8844	4.6416
94	4–ethyloctane	–11.5178	1.9423	42.1021	20.0625	168	14.8541	4.6416
95	2,2,3–trimethylheptane	–11.5382	2.2363	42.0053	20.5731	273	13.7056	4.6416
96	2,2,4–trimethylheptane	–11.5255	2.2090	42.0542	20.5214	213	12.8061	4.6416
97	2,2,5–trimethylheptane	–11.5308	2.1848	42.0328	20.4874	213	13.4211	4.6416
98	2,2,6–trimethylheptane	–11.5376	2.1431	42.0016	20.4321	198	12.8722	4.6416
99	2,3,3–trimethylheptane	–11.5794	2.2674	41.8604	20.6200	294	13.9833	4.6416
100	2,3,4–trimethylheptane	–11.5152	2.2022	42.1146	20.4710	254	14.1533	4.6416
101	2,3,5–trimethylheptane	–11.5210	2.1672	42.0937	20.4181	229	14.2161	4.6416
102	2,3,6–trimethylheptane	–11.5287	2.1216	42.0612	20.3567	214	14.0078	4.6416
103	2,4,4–trimethylheptane	–11.5188	2.2480	42.0808	20.5800	234	12.8789	4.6416
104	2,4,5–trimethylheptane	–11.5197	2.1754	42.1016	20.4291	229	14.0111	4.6416
105	2,4,6–trimethylheptane	–11.5273	2.1189	42.0686	20.3487	189	13.2506	4.6416
106	2,5,5–trimethylheptane	–11.5259	2.2158	42.0523	20.5346	234	13.6989	4.6416
107	3,3,4–trimethylheptane	–11.5110	2.2176	42.1065	20.5828	309	14.0561	4.6416
108	3,3,5–trimethylheptane	–11.5181	2.2648	42.0866	20.6026	249	13.5667	4.6416
109	3,4,4–trimethylheptane	–11.5096	2.3170	42.1143	20.6878	309	13.8511	4.6416
110	3,4,5–trimethylheptane	–11.5108	2.2352	42.1352	20.5164	269	14.4311	4.6416
111	3–ethyl–2–methylheptane	–11.5162	2.0741	42.1103	20.2675	211	14.4556	4.6416
112	4–ethyl–2–methylheptane	–11.5178	2.0602	42.1097	20.2417	186	14.2439	4.6416

Table 2. (Continued)

No	Compound	$SX_{1\mathrm{CH}}$	$SX_{1\mathrm{CC}}$	$SX_{2\mathrm{CH}}$	$SX_{2\mathrm{CC}}$	SV_{ij}	OEI [a]	$N^{2/3}$
113	5–ethyl–2–methylheptane	–11.5237	2.0346	42.0844	20.2094	186	14.3828	4.6416
114	3–ethyl–3–methylheptane	–11.5132	2.1826	42.1034	20.4713	252	14.3583	4.6416
115	4–ethyl–3–methylheptane	–11.5095	2.1204	42.1438	20.3283	226	14.6639	4.6416
116	5–ethyl–3–methylheptane	–11.5952	2.0804	41.8139	20.2708	201	14.5911	4.6416
117	3–ethyl–4–methylheptane	–11.5106	2.1159	42.1382	20.3233	226	14.5283	4.6416
118	4–ethyl–4–methylheptane	–11.5087	2.2037	42.1238	20.4985	252	14.0839	4.6416
119	4–propylheptane	–11.5144	1.9584	42.1193	20.0808	168	14.6883	4.6416
120	4–isopropylheptane	–11.7131	2.0915	41.3617	20.2885	211	14.5217	4.6416
121	2,2,3,3–tetrahexane	–11.5088	2.5074	42.0816	21.0432	420	13.3789	4.6416
122	2,2,3,4–tetrahexane	–11.5119	2.4178	42.1039	20.8497	331	13.4067	4.6416
123	2,2,3,5–tetrahexane	–11.5217	2.3552	42.0691	20.7556	291	13.3272	4.6416
124	2,2,4,4–tetrahexane	–11.5202	2.4382	42.0570	20.9157	276	11.9900	4.6416
125	2,2,4,5–tetrahexane	–11.5244	2.3378	42.0628	20.7238	256	12.9800	4.6416
126	2,2,5,5–tetrahexane	–11.5340	2.3523	41.9996	20.7867	240	12.5256	4.6416
127	2,3,3,4–tetrahexane	–11.5143	2.4615	42.0852	20.9238	387	13.8961	4.6416
128	2,3,3,5–tetrahexane	–11.5158	2.3895	42.0900	20.8093	312	13.2644	4.6416
129	2,3,4,4–tetrahexane	–11.5083	2.4442	42.1183	20.8919	352	13.5489	4.6416
130	2,3,4,5–tetrahexane	–11.5124	2.3348	42.1242	20.6800	297	13.9867	4.6416
131	3,3,4,4–tetrahexane	–11.5018	2.5621	42.1106	21.1259	456	13.6633	4.6416
132	2,2–dimethyl–3–ethylhexane	–11.5105	2.2189	42.1128	20.6056	288	13.6394	4.6416
133	2,2–dimethyl–4–ethylhexane	–11.5106	2.2928	42.1265	20.6334	228	13.3550	4.6416
134	2,3–dimethyl–3–ethylhexane	–11.5047	2.3488	42.1339	20.7341	330	15.0178	4.6416
135	2,3–dimethyl–4–ethylhexane	–11.5085	2.2483	42.1481	20.5318	269	14.3617	4.6416
136	2,4–dimethyl–3–ethylhexane	–11.5071	2.2569	42.1555	20.5426	269	14.1567	4.6416
137	2,4–dimethyl–4–ethylhexane	–11.5116	2.3046	42.1138	20.6605	270	13.6394	4.6416
138	2,5–dimethyl–3–ethylhexane	–11.5159	2.1932	42.1196	20.4498	229	14.0772	4.6416
139	3,3–dimethyl–4–ethylhexane	–11.5038	2.3590	42.1421	20.7433	324	13.9239	4.6416
140	3,4–dimethyl–3–ethylhexane	–11.5016	2.3772	42.1488	20.7744	345	14.2711	4.6416
141	3,3–diethylhexane	–11.5010	2.2645	42.1584	20.5847	288	14.5039	4.6416
142	3,4–diethylhexane	–11.3856	2.1626	42.5816	20.3838	241	14.7367	4.6416
143	3–isopropyl–2–methylhexane	–11.5102	2.2279	42.1406	20.5028	254	14.0144	4.6416
144	2,2,3,3,4–pentamethylpentane	–11.3029	2.6542	42.8909	21.2878	498	13.0833	4.6416
145	2,2,3,4,4–pentamethylpentane	–11.5120	2.6037	42.0751	21.1915	393	12.0417	4.6416
146	3–ethyl–2,2,3–trimethylpentane	–11.4998	2.5709	42.1180	21.1374	456	13.4583	4.6416
147	3–ethyl–2,2,4–trimethylpentane	–11.5068	2.4432	42.1266	20.8833	331	12.7917	4.6416
148	3–ethyl–2,3,4–trimethylpentane	–11.4999	2.4966	42.1467	20.9778	408	13.8333	4.6416
149	3,3–diethyl–2–methylpentane	–11.4958	2.4130	42.1706	20.8271	366	14.2083	4.6416
150	2,4–dimethyl–3–isopropylpentane	–11.5060	2.3679	42.1538	20.7237	297	13.1667	4.6416
151	undecane	–12.5296	1.9726	46.0298	22.0122	136	16.7039	4.9461
152	dodecane	–13.5220	2.2268	50.0546	24.2682	152	18.3563	5.2415
153	tridecane	–14.5138	2.4834	54.0844	26.5258	168	19.9949	5.5288
154	tetradecane	–15.5046	2.7412	58.1190	28.7854	184	21.6453	5.8088
155	pentadecane	–16.4948	3.0014	62.1580	31.0456	200	23.2855	6.0822
156	hexadecane	–17.4838	3.2623	66.1984	33.3075	216	24.9346	6.3496
157	heptadecane	–18.4720	3.5242	70.2420	35.5706	232	26.5759	6.6115
158	octadecane	–19.4602	3.7874	74.2886	37.8342	248	28.2241	6.8683
159	nonadecane	–20.4476	4.0514	78.3376	40.0986	264	29.8661	7.1204
160	eicosane	–21.4344	4.3159	82.3886	42.3641	280	31.5136	7.3681

[a] from Ref. [17]

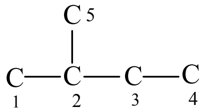

Figure 4. The skeleton of 2–methylbutane.

With the same calculation method as above, the values of parameters $SX_{1\mathrm{CH}}$, $SX_{1\mathrm{CC}}$, SV_{ij} and so on for 160 alkanes (150 alkanes with carbon atom C_1–C_{10} and 10 normal alkanes with C_{11}–C_{20}) were calculated and listed in Table 2. This paper has got two groups of parameters: one is $SX_{1\mathrm{CH}}$, $SX_{1\mathrm{CC}}$ and SV_{ij}; another is $SX_{2\mathrm{CH}}$, $SX_{2\mathrm{CC}}$ and SV_{ij}. We made regression analyses between the two sets of parameters above and the boiling points of 160 alkanes, and obtained the following correlation expressions:

$$\mathrm{Bp} = -206.875 - 44.1901\, SX_{1\mathrm{CH}} - 100.765\, SX_{1CC} + 0.299541 SV_{ij} \quad r = 0.9841,\ s = 10.72^\circ\mathrm{C},\ \mathrm{F} = 1592.5,\ \mathrm{n} = 160 \tag{1a}$$

$$\mathrm{Bp} = -62.72 + 0.336535\, SX_{2\mathrm{CH}} + 10.1031\, SX_{2\mathrm{CC}} + 0.002909 SV_{ij} \quad r = 0.9730,\ s = 13.93^\circ\mathrm{C},\ \mathrm{F} = 922.72,\ \mathrm{n} = 160 \tag{1b}$$

The equations above show good correlations but still have larger standard deviation, which reminds us that some molecular structural information is still not counted. It has been proved that the addition of N (the number of carbon atoms) and Odd–Even Index (*OEI*) can improve the correlations with the physicochemical properties of hydrocarbons [17–19]. Therefore, they were also employed in this paper.

3 RESULTS AND DISCUSSION

To build a property data set, 150 alkanes from C_1 to C_{10} and 10 straight chain alkanes from C_{11} to C_{20} were selected. The data set contains 160 boiling points (Bp) [20], 61 critical temperatures (Tc) [20], 50 critical volumes (Vc) [20], 60 critical pressures (Pc) [20], 94 densities (D) at 25 ºC [20], 149 refractive indices (n_D) at 20 ºC [20], 63 vapor pressures (log Pv) [21], 34 heat capacities (Cp) in liquid [20], 57 heats of vaporization (Hv) [20], and 64 chromatographic retention indices (Ipt) [22] of examined alkanes. The selection of these compounds has been determined by the availability of the most recently experimental values in the literature.

Table 3. The *r* and *s* of two groups of results

property	*r*(I)	*s*(I)	*r*(II)	*s*(II)
Bp	0.9987	3.06 ˚C	0.9984	3.47 ˚C
Tc	0.9989	4.73 ˚C	0.9985	5.49 ˚C
Vc	0.9995	7.66 $\mathrm{cm^3{\cdot}mol^{-1}}$	0.9991	10.40 $\mathrm{cm^3{\cdot}mol^{-1}}$
Pc	0.9900	0.1036 Mpa	0.9897	0.1063 Mpa
D	0.9907	0.0064 $\mathrm{g{\cdot}cm^{-3}}$	0.9911	0.0062 $\mathrm{g{\cdot}cm^{-3}}$
n_D	0.9499	0.0055	0.9508	0.0055
Log Pv	0.9974	0.0959	0.9965	0.1117
Cp	0.9997	2.02 $\mathrm{J{\cdot}mol^{-1}{\cdot}K^{-1}}$	0.9997	2.01 $\mathrm{J{\cdot}mol^{-1}{\cdot}K^{-1}}$
Hv	0.9994	0.66 $\mathrm{kJ{\cdot}mol^{-1}}$	0.9978	1.27 $\mathrm{kJ{\cdot}mol^{-1}}$
Ipt	0.9986	8.09	0.9984	8.81

Now, the two groups of parameters in this paper are: I. $SX_{1\mathrm{CH}}$, $SX_{1\mathrm{CC}}$, SV_{ij}, *OEI* and $N^{2/3}$; II. $SX_{2\mathrm{CH}}$, $SX_{2\mathrm{CC}}$, SV_{ij}, *OEI* and $N^{2/3}$. The linear correlations between 10 physicochemical properties and these two groups of parameters were performed for alkanes. The obtained correlation coefficient (r) and the standard deviation (s) were listed in Table 3.

From Table 3, we may draw a conclusion that most of the correlations with the group I of parameters are somewhat better than those with the group II of parameters. Therefore, we take the group I of parameters to correlate with the physicochemical properties of alkanes. The expressions may be (2a) or (2b):

$$y = a_0 + a_1 SX_{1\mathrm{CH}} + a_2 SX_{1\mathrm{CC}} + a_3 SV_{ij} + a_4 OEI + a_5 N^{2/3} \tag{2a}$$

$$\ln(\mathrm{opt} - y) = b_0 + b_1 SX_{1\mathrm{CH}} + b_2 SX_{1\mathrm{CC}} + b_3 SV_{ij} + b_4 OEI + b_5 N^{2/3} \tag{2b}$$

Here a_0 and b_0 are the intercept terms, a_i and b_i are the regression coefficients, "opt" is an optimizing value of some property. The best regression expressions of 10 physicochemical properties were obtained and listed below. As is well known, an excellent QSPR model should have not only good estimation ability for any internal sample, but should also have good prediction ability for an external sample. The most usual method to prove that a model has excellent prediction ability is a cross–validation (CV). Therefore every QSPR model was estimated with CV, the obtained corresponding correlation coefficient q^2 and *PRESS* etc. were also listed in Eqs. (3)–(12):

$$\begin{gathered} \mathrm{Bp} = -245.915 + 16.6783 SX_{1\mathrm{CH}} - 42.1725 SX_{1\mathrm{CC}} + 0.1179 SV_{ij} + 5.3772 OEI + 126.1505 N^{2/3} \\ r = 0.9987,\ \mathrm{r_a}^2 = 0.9972,\ s = 3.06^\circ\mathrm{C},\ \mathrm{F} = 12100.5,\ n = 160 \\ q^2 = 0.997,\ PRESS = 3.22^\circ\mathrm{C},\ \mathrm{F} = 10920.72(\mathrm{df}5,154),\ \text{cross–validation(CV)} \end{gathered} \tag{3}$$

$$\begin{gathered} \ln(995 - \mathrm{Tc}) = 6.8635 + 0.001018 SX_{1\mathrm{CH}} + 0.1353 SX_{1\mathrm{CC}} - 0.00054 SV_{ij} - 0.01753 OEI - 0.1768 N^{2/3} \\ r = 0.9991,\ \mathrm{r_a}^2 = 0.9981,\ s = 4.03^\circ\mathrm{C},\ \mathrm{F} = 6403.20,\ n = 61 \\ q^2 = 0.996,\ PRESS = 6.28^\circ\mathrm{C},\ \mathrm{F} = 2513.18(\mathrm{df}5,55),\ \mathrm{CV} \end{gathered} \tag{4}$$

$$\begin{gathered} \mathrm{Vc} = -229.595 - 229.726 SX_{1\mathrm{CH}} - 235.22 SX_{1\mathrm{CC}} + 0.3023 SV_{ij} - 20.8066 OEI - 240.902 N^{2/3} \\ r = 0.9995,\ \mathrm{r_a}^2 = 0.9986,\ s = 7.66\mathrm{cm}^3\cdot\mathrm{mol}^{-1},\ \mathrm{F} = 8543.96,\ n = 50 \\ q^2 = 0.999,\ PRESS = 8.43\ \mathrm{cm}^3\cdot\mathrm{mol}^{-1},\ \mathrm{F} = 7042.20(\mathrm{df}5,44),\ \mathrm{CV} \end{gathered} \tag{5}$$

$$\begin{gathered} \ln(\mathrm{Pc} - 0.9) = 2.0248 + 0.4569 SX_{1\mathrm{CH}} + 0.2588 SX_{1\mathrm{CC}} + 0.00051 SV_{ij} + 0.05345 OEI + 0.4583 N^{2/3} \\ r = 0.9941,\ \mathrm{r_a}^2 = 0.9871,\ s = 0.0771\mathrm{Mpa},\ \mathrm{F} = 904.47,\ n = 60 \\ q^2 = 0.935,\ PRESS = 0.1514\mathrm{Mpa},\ \mathrm{F} = 154.35(\mathrm{df}5,54),\ \mathrm{CV} \end{gathered} \tag{6}$$

$$\begin{gathered} \ln(0.9 - \mathrm{D}) = -0.3253 - 0.2806 SX_{1\mathrm{CH}} + 0.01938 SX_{1\mathrm{CC}} - 0.00126 SV_{ij} - 0.05397 OEI - 0.8026 N^{2/3} \\ r = 0.9906,\ \mathrm{r_a}^2 = 0.9803,\ s = 0.0048\mathrm{g}\cdot\mathrm{cm}^{-3},\ \mathrm{F} = 924.49,\ n = 94 \\ q^2 = 0.969,\ PRESS = 0.0066\mathrm{g}\cdot\mathrm{cm}^{-3},\ \mathrm{F} = 543.26(\mathrm{df}5,88),\ \mathrm{CV} \end{gathered} \tag{7}$$

$$\begin{gathered} \mathrm{n_D} = 1.1848 + 0.02241 SX_{1\mathrm{CH}} - 0.03737 SX_{1\mathrm{CC}} + 0.000177 SV_{ij} + 0.000466 OEI + 0.1127 N^{2/3} \\ r = 0.9499,\ \mathrm{r_a}^2 = 0.8990,\ s = 0.0055,\ \mathrm{F} = 264.37,\ n = 149 \\ q^2 = 0.879,\ PRESS = 0.0060,\ \mathrm{F} = 207.31(\mathrm{df}5,143),\ \mathrm{CV} \end{gathered} \tag{8}$$

$$\begin{gathered} \log\mathrm{Pv} = 9.4487 + 0.9288 SX_{1\mathrm{CH}} + 1.7147 SX_{1\mathrm{CC}} - 0.00305 SV_{ij} - 0.02673 OEI + 0.2913 N^{2/3} \\ r = 0.9974,\ \mathrm{r_a}^2 = 0.9944,\ s = 0.0959,\ \mathrm{F} = 2191.31,\ n = 63 \\ q^2 = 0.993,\ PRESS = 0.1110,\ \mathrm{F} = 1646.10(\mathrm{df}5,57),\ \mathrm{CV} \end{gathered} \tag{9}$$

$$\begin{gathered} \mathrm{Cp} = 4.0364 - 39.2301 SX_{1\mathrm{CH}} - 7.4779 SX_{1\mathrm{CC}} - 0.00916 SV_{ij} + 3.1347 OEI - 37.7583 N^{2/3} \\ r = 0.9997,\ \mathrm{r_a}^2 = 0.9992,\ s = 2.02\mathrm{J}\cdot\mathrm{mol}^{-1}\cdot\mathrm{K}^{-1},\ \mathrm{F} = 8565.69,\ n = 34 \\ q^2 = 0.999,\ PRESS = 2.37\ \mathrm{J}\cdot\mathrm{mol}^{-1}\cdot\mathrm{K}^{-1},\ \mathrm{F} = 6187.51(\mathrm{df}5,28),\ \mathrm{CV} \end{gathered} \tag{10}$$

$$\mathrm{Hv} = -37.7075 - 18.0227SX_{1\mathrm{CH}} - 36.4903SX_{1\mathrm{CC}} + 0.07941SV_{ij} - 1.1393OEI - 10.3182N^{2/3}$$
$$r = 0.9994,\ \mathrm{r_a}^2 = 0.9987,\ s = 0.66\mathrm{kJ{\cdot}mol^{-1}},\ \mathrm{F} = 8562.06,\ n = 57 \qquad (11)$$
$$q^2 = 0.997,\ PRESS = 1.05\ \mathrm{kJ{\cdot}mol^{-1}},\ \mathrm{F} = 3396.10(\mathrm{df}5{,}51),\ \mathrm{CV}$$

$$\mathrm{Ipt} = -75.6355 - 111.69SX_{1\mathrm{CH}} - 137.421SX_{1\mathrm{CC}} + 0.3251SV_{ij} + 32.3281OEI - 110.349N^{2/3}$$
$$r = 0.9986,\ \mathrm{r_a}^2 = 0.9970,\ s = 8.09,\ \mathrm{F} = 4247.75,\ n = 64 \qquad (12)$$
$$q^2 = 0.997,\ PRESS = 9.14,\ \mathrm{F} = 3329.71(\mathrm{df}5{,}58),\ \mathrm{CV}$$

Table 4. Experimental and Estimated Boiling Point ($\mathrm{Bp_{exp}}$, $\mathrm{Bp_{calc}}$) for 160 Alkanes

No	$\mathrm{Bp_{exp}}$ [a]	$\mathrm{Bp_{calc}}$ [b]	No	$\mathrm{Bp_{exp}}$ [a]	$\mathrm{Bp_{calc}}$ [b]	No	$\mathrm{Bp_{exp}}$ [a]	$\mathrm{Bp_{calc}}$ [b]	No	$\mathrm{Bp_{exp}}$ [a]	$\mathrm{Bp_{calc}}$ [b]
1	–161.5	–160.99	42	143.2	141.43	83	156.0	159.92	124	153.8	141.69
2	–88.6	–92.92	43	144.2	142.72	84	158.5	161.44	125	147.9	148.82
3	–42.1	–44.64	44	142.4	141.68	85	160.4	161.73	126	137.4	143.72
4	–0.5	–2.77	45	132.7	132.12	86	159.9	159.88	127	164.6	164.15
5	–11.7	–9.47	46	140.5	141.59	87	161.2	159.31	128	153.1	154.92
6	36.0	33.03	47	132.9	135.13	88	163.4	164.21	129	161.6	158.98
7	27.8	29.13	48	136.0	131.82	89	159.4	160.15	130	156.2	159.39
8	9.4	16.24	49	135.2	131.92	90	160.8	162.42	131	170.0	167.00
9	68.7	65.39	50	137.3	136.75	91	157.5	157.62	132	156.1	161.38
10	60.2	60.06	51	140.6	141.09	92	162.2	164.52	133	147.0	149.66
11	63.2	62.82	52	136.0	142.70	93	166.5	166.24	134	166.0	168.37
12	49.7	53.29	53	135.2	133.91	94	163.7	165.30	135	162.0	161.82
13	57.9	60.21	54	143.0	143.10	95	157.8	158.77	136	162.0	160.38
14	98.5	94.53	55	141.2	142.20	96	148.3	148.22	137	161.1	155.63
15	90.0	89.77	56	133.6	136.48	97	150.8	152.46	138	154.1	157.77
16	92.0	91.09	57	126.5	123.94	98	148.9	149.39	139	162.9	161.36
17	79.2	80.48	58	124.0	128.59	99	160.2	160.75	140	162.1	164.98
18	89.7	89.87	59	137.7	138.65	100	161.0	160.76	141	166.3	164.27
19	80.4	83.07	60	139.1	139.10	101	160.7	159.53	142	163.9	166.20
20	86.0	86.80	61	131.4	135.11	102	156.0	158.44	143	165.0	159.02
21	93.5	92.41	62	130.7	129.28	103	151.0	149.56	144	166.1	168.27
22	80.8	85.97	63	140.5	140.02	104	156.5	158.10	145	159.3	148.92
23	125.6	121.53	64	138.0	140.19	105	147.6	151.55	146	169.5	165.56
24	117.6	116.19	65	133.8	137.38	106	152.8	155.21	147	155.3	152.50
25	118.9	118.07	66	140.6	139.38	107	161.9	166.15	148	169.5	165.05
26	117.7	116.64	67	140.0	141.55	108	155.7	154.33	149	172.0	165.70
27	106.8	108.05	68	140.2	141.93	109	161.1	160.88	150	157.1	153.69
28	115.6	116.00	69	133.0	132.03	110	162.5	162.71	151	195.9	191.74
29	109.5	111.95	70	122.2	113.87	111	163.0	162.70	152	216.3	212.50
30	109.1	112.04	71	141.5	141.46	112	158.0	159.17	153	235.4	232.08
31	111.9	111.55	72	133.8	134.98	113	159.7	160.90	154	253.5	250.77
32	117.7	117.36	73	144.7	142.99	114	163.8	162.49	155	270.6	268.48
33	118.6	118.51	74	136.7	135.91	115	165.0	163.75	156	286.8	285.47
34	110.0	111.89	75	146.3	142.89	116	158.2	160.67	157	302.0	301.69
35	99.2	99.99	76	174.1	169.87	117	165.0	163.19	158	316.3	317.25
36	114.8	115.69	77	167.1	164.53	118	160.8	160.20	159	329.7	332.17
37	113.5	113.41	78	167.9	166.17	119	157.5	163.79	160	343.0	346.55
38	115.6	116.03	79	165.7	165.83	120	158.9	159.04			
39	118.2	116.46	80	165.1	165.21	121	160.3	163.41			
40	106.4	115.31	81	155.0	155.89	122	157.0	156.79			
41	150.8	146.44	82	164.3	163.76	123	148.4	154.12			

[a] From Ref. [20]
[b] Calculated with Eq. (3)

Table 5. Experimental and Estimated Critical Temperature (Tc_{exp}, Tc_{calc}) for 61 Alkanes

No	Tc_{exp} [a]	Tc_{calc} [b]	No	Tc_{exp} [a]	Tc_{calc} [b]	No	Tc_{exp} [a]	Tc_{calc} [b]	No	Tc_{exp} [a]	Tc_{calc} [b]
1	190.56	195.33	17	520.50	521.42	33	565.50	566.27	76	617.70	616.98
2	305.32	300.11	18	537.30	536.65	34	563.50	564.85	108	609.50	605.43
3	369.83	366.16	19	519.80	523.47	35	543.80	542.31	121	623.00	626.84
4	425.12	421.78	20	536.40	533.66	36	573.50	571.19	126	581.40	591.57
5	407.80	409.77	21	540.60	538.20	37	566.40	564.26	151	639.00	638.30
6	469.70	466.29	22	531.10	535.65	38	567.10	565.80	152	658.00	657.88
7	460.40	461.35	23	568.70	567.26	39	576.50	569.28	153	675.00	675.75
8	433.80	439.52	24	559.70	561.14	40	567.80	576.92	154	693.00	692.30
9	507.60	504.80	25	563.60	564.75	41	594.60	593.31	155	708.00	707.51
10	497.70	497.87	26	561.70	562.69	42	582.80	588.01	156	723.00	721.68
11	504.60	503.47	27	549.80	552.66	45	576.70	578.42	157	736.00	734.83
12	489.00	491.01	28	563.50	564.86	58	569.80	574.92	158	747.00	747.11
13	500.00	501.70	29	553.50	557.47	68	607.50	606.50	160	768.00	769.31
14	540.20	537.81	30	550.00	556.65	69	592.60	587.86			
15	530.40	532.19	31	562.00	560.03	70	574.60	557.28			
16	535.20	535.22	32	568.80	567.71	71	607.50	602.88			

[a] From Ref. [20] [b] Calculated with Eq. (4)

Table 6. Experimental and Estimated Critical Volume (Vc_{exp}, Vc_{cal}) for 50 Alkanes

No	Vc_{exp} [a]	Vc_{calc} [b]	No	Vc_{exp} [a]	Vc_{calc} [b]	No	Vc_{exp} [a]	Vc_{calc} [b]	No	Vc_{exp} [a]	Vc_{calc} [b]
1	98.6	97.39	14	428.0	430.46	27	478.0	469.80	40	461.0	454.75
2	145.5	147.37	15	421.0	421.53	28	468.0	462.17	41	555.0	555.41
3	200.0	202.47	16	404.0	415.24	29	472.0	466.93	76	624.0	620.17
4	255.0	255.74	17	416.0	415.20	30	482.0	471.18	151	689.0	686.82
5	259.0	256.69	18	393.0	405.58	31	443.0	459.66	152	754.0	754.31
6	311.0	312.67	19	418.0	418.19	32	466.0	455.98	153	823.0	823.32
7	306.0	305.74	20	414.0	401.28	33	455.0	465.25	154	894.0	893.34
8	307.0	309.13	21	416.0	408.91	34	436.0	457.96	155	966.0	964.46
9	368.0	370.15	22	398.0	402.23	35	468.0	468.98	156	1034.0	1036.40
10	368.0	365.10	23	492.0	491.90	36	455.0	448.97	157	1103.0	1109.40
11	368.0	357.10	24	488.0	484.14	37	460.0	454.76	158	1189.0	1183.19
12	358.0	358.66	25	464.0	473.88	38	442.0	458.24			
13	361.0	353.46	26	476.0	475.62	39	455.0	447.85			

[a] From Ref. [20] [b] Calculated with Eq. (5)

Table 7. Experimental and Estimated Critical Pressure (Pc_{exp}, Pc_{calc}) for 60 Alkanes

No	Pc_{exp} [a]	Pc_{calc} [b]	No	Pc_{exp} [a]	Pc_{calc} [b]	No	Pc_{exp} [a]	Pc_{calc} [b]	No	Pc_{exp} [a]	Pc_{calc} [b]
1	4.599	4.771	17	2.770	2.812	33	2.610	2.642	108	2.320	2.287
2	4.872	4.451	18	2.910	2.926	34	2.730	2.743	121	2.510	2.503
3	4.248	4.042	19	2.740	2.790	35	2.570	2.558	126	2.190	2.226
4	3.796	3.700	20	2.950	2.945	36	2.820	2.821	151	1.980	1.940
5	3.640	3.638	21	2.890	2.896	37	2.730	2.723	152	1.820	1.790
6	3.370	3.363	22	2.950	2.991	38	2.700	2.697	153	1.680	1.659
7	3.380	3.387	23	2.490	2.532	39	2.810	2.779	154	1.570	1.546
8	3.196	3.277	24	2.500	2.543	40	2.870	2.910	155	1.480	1.448
9	3.025	3.060	25	2.550	2.598	41	2.290	2.309	156	1.400	1.364
10	3.040	3.065	26	2.540	2.585	42	2.310	2.324	157	1.340	1.293
11	3.120	3.136	27	2.530	2.582	45	2.350	2.346	158	1.290	1.231
12	3.100	3.096	28	2.630	2.668	68	2.740	2.737	160	1.070	1.135
13	3.150	3.170	29	2.560	2.603	69	2.600	2.536			
14	2.740	2.780	30	2.490	2.575	70	2.490	2.348			
15	2.740	2.800	31	2.650	2.671	71	2.720	2.694			
16	2.810	2.848	32	2.690	2.712	76	2.110	2.113			

[a] From Ref. [20] [b] Calculated with Eq. (6)

Table 8. Experimental and Estimated Density (D_{exp}, D_{calc}) at 25°C for 94 Alkanes

No	D_{exp} [a]	D_{calc} [b]	No	D_{exp} [a]	D_{calc} [b]	No	D_{exp} [a]	D_{calc} [b]	No	D_{exp} [a]	D_{calc} [b]
3	0.4930	0.5041	63	0.7414	0.7420	102	0.7305	0.7349	127	0.7656	0.7662
4	0.5730	0.5684	65	0.7195	0.7187	103	0.7308	0.7290	128	0.7449	0.7480
5	0.5510	0.5552	66	0.7371	0.7327	104	0.7373	0.7383	129	0.7586	0.7578
8	0.5852	0.5895	68	0.7530	0.7575	105	0.7190	0.7226	130	0.7456	0.7513
9	0.6548	0.6470	73	0.7508	0.7475	106	0.7362	0.7362	131	0.7789	0.7760
10	0.6500	0.6440	81	0.7208	0.7238	107	0.7527	0.7544	132	0.7447	0.7471
11	0.6598	0.6535	83	0.7226	0.7287	109	0.7535	0.7526	134	0.7599	0.7653
12	0.6444	0.6446	84	0.7264	0.7310	110	0.7519	0.7499	135	0.7516	0.7494
23	0.6986	0.6944	86	0.7202	0.7239	111	0.7398	0.7390	136	0.7514	0.7478
28	0.6912	0.7065	87	0.7351	0.7352	112	0.7322	0.7319	137	0.7525	0.7433
29	0.6962	0.6964	88	0.7410	0.7403	113	0.7318	0.7329	138	0.7368	0.7390
30	0.6901	0.6931	89	0.7329	0.7320	114	0.7463	0.7461	139	0.7598	0.7560
32	0.7151	0.7117	90	0.7324	0.7343	115	0.7466	0.7439	140	0.7596	0.7625
35	0.6877	0.6856	91	0.7312	0.7342	116	0.7368	0.7345	141	0.7575	0.7544
42	0.7095	0.7103	92	0.7432	0.7418	117	0.7468	0.7427	142	0.7472	0.7525
43	0.7170	0.7158	93	0.7359	0.7341	118	0.7472	0.7439	143	0.7436	0.7436
44	0.7160	0.7147	94	0.7343	0.7340	119	0.7321	0.7326	144	0.7767	0.7851
47	0.7115	0.7115	95	0.7385	0.7435	120	0.7354	0.7304	145	0.7636	0.7529
54	0.7225	0.7205	96	0.7237	0.7235	121	0.7609	0.7681	146	0.7780	0.7747
55	0.7241	0.7200	97	0.7243	0.7291	122	0.7513	0.7528	148	0.7735	0.7697
56	0.7257	0.7319	98	0.7200	0.7204	123	0.7336	0.7442	149	0.7755	0.7657
59	0.7345	0.7377	99	0.7450	0.7480	124	0.7424	0.7292	150	0.7545	0.7447
60	0.7354	0.7329	100	0.7447	0.7446	125	0.7316	0.7340			
62	0.7201	0.7181	101	0.7413	0.7400	126	0.7148	0.7259			

[a] From Ref. [20] [b] Calculated with Eq. (7)

From the equations above, we can see that most of the regression equations have high correlation coefficient (r) and low standard deviation (s). Using the expressions above, we can estimate the physicochemical properties of alkanes. The experimental and estimated properties are shown in Tables 4–13, respectively. The numberings of compounds in all tables are consistent with those in Table 2.

As can be seen from the results above, there is no remarkable estimated deviation for most QSPR models. The estimated values are quite close to the experimental ones. As we know, boiling point has received a lot of research and many papers were produced [13,17–19,23–25]. But the result in this paper (r = 0.9986, s = 3.06˚C) is better than most previous results. And there are only two compound deviations being larger than 10˚C, they are No. 124 (Bp_{exp} = 153.8˚C, ΔBp = –12.11˚C) and No. 145 (Bp_{exp} = 159.3˚C, ΔBp = –10.38˚C). Also the deviation obtained using the CV technique is only 3.26˚C. In general, nonlinear regression model is better than linear one for boiling point. However, in the present study, it is found that the correlation coefficient r = 0.9987 and standard deviation s = 3.02˚C for nonlinear model are not improved greatly than those for the linear model when opt = 950. Therefore, the linear regression equation was used here because of its simple calculation.

The critical properties express the individual features of each substance and are determined by the molecular structural differences between the substances. Thus, critical properties should be primary targets for calculation of structure–property relationships. In the literature [26], Katritzky *et* *al*. have reviewed the estimation methods for critical temperature in detail and developed one–parameter for 76 hydrocarbon and three–parameter model for 165 diverse compounds, the obtained results are r^2 = 0.953, s = 13.9K and r^2 = 0.955, s = 14.8K, respectively. In this paper, three QSPR models between Tc, Pc, and Vc with five parameters are developed. Their correlation coefficients are all larger than 0.99. Here, it should be noted that the logarithm models of Tc, Pc give better results than linear ones. Stelian Grigoras [4] have studied critical temperature (Tc) and critical molar volume (Vc) with molecular surface interactions (MSI) and obtained r = 0.982, s = 16.4°C and r = 0.990, s = 12.44cm^3·mol^{-1}. This work, got the results r = 0.9991, s = 4.03°C and r = 0.9995, s = 7.66cm^3·mol^{-1} for Tc and Vc respectively.

Table 9. Experimental and Estimated Refraction Index (n_{Dexp}, n_{Dcalc}) for 149 Alkanes

No	n_{Dexp} [a]	n_{Dcalc} [b]	No	n_{Dexp} [a]	n_{Dcalc} [b]	No	n_{Dexp} [a]	n_{Dcalc} [b]	No	n_{Dexp} [a]	n_{Dcalc} [b]
4	1.3326	1.3397	48	1.4033	1.3989	86	1.4086	1.4111	124	1.4208	1.4131
6	1.3575	1.3584	49	1.4011	1.4014	87	1.4165	1.4158	125	1.4132	1.4136
7	1.3537	1.3578	50	1.4087	1.4086	88	1.4182	1.4170	126	1.4055	1.4098
9	1.3749	1.3737	51	1.4108	1.4099	89	1.4139	1.4131	127	1.4298	1.4328
10	1.3715	1.3727	52	1.4083	1.4127	90	1.4139	1.4139	128	1.4196	1.4219
11	1.3765	1.3748	53	1.4076	1.4079	91	1.4144	1.4149	129	1.4267	1.4273
12	1.3688	1.3728	54	1.4093	1.4072	92	1.4190	1.4183	130	1.4204	1.4217
13	1.3750	1.3761	55	1.4096	1.4068	93	1.4156	1.4145	131	1.4368	1.4415
14	1.3878	1.3864	56	1.4106	1.4138	94	1.4151	1.4139	132	1.4197	1.4244
15	1.3848	1.3853	57	1.4033	1.3993	95	1.4168	1.4205	133	1.4131	1.4109
16	1.3887	1.3870	58	1.3997	1.4031	96	1.4092	1.4107	134	1.4270	1.4277
17	1.3822	1.3851	59	1.4141	1.4166	97	1.4101	1.4118	135	1.4226	1.4203
18	1.3919	1.3900	60	1.4144	1.4125	98	1.4078	1.4103	136	1.4225	1.4199
19	1.3815	1.3841	61	1.4051	1.4073	99	1.4202	1.4222	137	1.4235	1.4180
20	1.3909	1.3903	62	1.4074	1.4070	100	1.4195	1.4191	138	1.4157	1.4150
21	1.3934	1.3887	63	1.4178	1.4184	101	1.4169	1.4159	139	1.4269	1.4258
22	1.3864	1.3942	64	1.4106	1.4094	102	1.4131	1.4147	140	1.4267	1.4290
23	1.3974	1.3970	65	1.4063	1.4056	103	1.4142	1.4132	141	1.4258	1.4233
24	1.3949	1.3959	66	1.4140	1.4127	104	1.4160	1.4155	142	1.4190	1.4215
26	1.3979	1.3971	67	1.4134	1.4111	105	1.4071	1.4100	143	1.4195	1.4182
27	1.3935	1.3955	68	1.4236	1.4307	106	1.4149	1.4146	144	1.4361	1.4496
28	1.4011	1.4000	69	1.4147	1.4155	107	1.4236	1.4283	145	1.4307	1.4278
30	1.3925	1.3948	70	1.4069	1.4016	108	1.4170	1.4155	146	1.4420	1.4411
31	1.4001	1.4000	71	1.4222	1.4252	109	1.4235	1.4245	148	1.4333	1.4355
32	1.4041	1.4017	72	1.4123	1.4147	110	1.4229	1.4208	149	1.4343	1.4315
33	1.4018	1.3987	73	1.4221	1.4219	111	1.4174	1.4164	150	1.4246	1.4203
34	1.4030	1.4053	74	1.4131	1.4114	112	1.4137	1.4124	151	1.4398	1.4196
35	1.3915	1.3939	75	1.4206	1.4172	113	1.4134	1.4133	152	1.4216	1.4247
36	1.4075	1.4089	76	1.4102	1.4134	114	1.4208	1.4196	153	1.4256	1.4289
37	1.4042	1.4027	77	1.4099	1.4122	115	1.4206	1.4176	154	1.4290	1.4322
38	1.4040	1.4015	78	1.4125	1.4137	116	1.4164	1.4127	155	1.4315	1.4347
39	1.4078	1.4049	79	1.4123	1.4149	117	1.4207	1.4177	156	1.4345	1.4365
40	1.4695	1.4182	80	1.4116	1.4130	118	1.4210	1.4188	157	1.4369	1.4377
41	1.4054	1.4059	81	1.4082	1.4116	119	1.4135	1.4133	158	1.4390	1.4383
42	1.4031	1.4048	82	1.4146	1.4159	120	1.4153	1.4114	160	1.4425	1.4378
45	1.4016	1.4042	83	1.4091	1.4132	121	1.4282	1.4368			
46	1.4088	1.4095	84	1.4112	1.4137	122	1.4216	1.4244			
47	1.4034	1.4044	85	1.4097	1.4142	123	1.4142	1.4194			

[a] From Ref. [20]   [b] Calculated with Eq. (8)

Table 10. Experimental and Estimated Vapor Pressure (logPv $^{a}_{exp}$, logPv $^{b}_{calc}$) for 63 Alkanes

No	logPv$_{exp}$ ^{a}	logPv$_{calc}$ ^{b}	No	logPv$_{exp}$ ^{a}	logPv$_{calc}$ ^{b}	No	logPv$_{exp}$ ^{a}	logPv$_{calc}$ ^{b}	No	logPv$_{exp}$ ^{a}	logPv$_{calc}$ ^{b}
2	6.622	6.617	20	4.042	4.079	37	3.558	3.618	81	2.686	2.657
3	5.979	5.983	21	3.890	3.925	38	3.503	3.545	108	2.746	2.863
4	5.387	5.388	22	4.135	4.107	39	3.486	3.592	121	2.730	2.771
5	5.546	5.513	23	3.272	3.270	41	2.764	2.768	126	3.072	3.054
6	4.835	4.835	24	3.439	3.410	42	2.927	2.908	132	3.177	2.671
7	4.962	4.942	25	3.417	3.420	43	2.921	2.924	136	3.126	2.783
9	4.307	4.300	26	3.436	3.442	44	2.959	2.944	151	1.745	1.772
10	4.449	4.429	27	3.657	3.640	49	3.094	3.226	152	1.252	1.279
11	4.402	4.412	28	3.495	3.522	58	3.347	3.292	153	0.755	0.789
12	4.631	4.593	29	3.607	3.572	68	3.103	3.184	154	0.271	0.300
13	4.496	4.486	30	3.606	3.547	70	3.427	3.590	155	–0.184	–0.187
14	3.783	3.782	31	3.581	3.627	75	2.988	3.146	156	–0.700	–0.673
15	3.943	3.914	32	3.461	3.526	76	2.258	2.268			
16	3.913	3.919	33	3.428	3.452	77	2.400	2.412			
17	4.147	4.135	34	3.631	3.644	78	2.421	2.429			
18	3.962	3.996	35	3.818	3.808	79	2.490	2.434			
19	4.118	4.064	36	3.556	3.618	80	2.468	2.470			

^{a} From Ref. [21] ^{b} Calculated with Eq. (9)

Table11. Experimental and Estimated Heat of Capacity (Cp$_{exp}$, Cp$_{calc}$) for 34 Alkanes

No	Cp$_{exp}$ ^{a}	Cp$_{calc}$ ^{b}	No	Cp$_{exp}$ ^{a}	Cp$_{calc}$ ^{b}	No	Cp$_{exp}$ ^{a}	Cp$_{calc}$ ^{b}	No	Cp$_{exp}$ ^{a}	Cp$_{calc}$ ^{b}
4	140.9	139.37	17	221.1	217.26	35	239.1	243.13	151	344.9	345.18
6	167.2	167.01	19	224.2	219.13	36	245.6	245.68	152	375.8	376.09
7	164.8	165.24	21	219.6	222.40	37	247.3	247.03	153	406.7	407.22
9	195.6	195.66	22	213.5	216.63	41	284.4	284.21	156	501.6	501.97
10	193.7	193.12	23	254.6	254.30	68	271.5	270.29			
11	190.7	193.86	24	252.0	251.76	70	266.3	266.76			
12	191.9	190.09	25	250.2	252.08	75	278.2	276.48			
13	189.7	191.97	26	251.1	251.29	76	314.4	314.56			
14	224.7	224.67	30	249.2	249.93	77	313.3	312.03			
15	222.9	222.50	31	246.6	247.07	80	314.4	311.75			

^{a} From Ref. [20] ^{b} Calculated with Eq. (10)

Table 12. Experimental and Estimated Heat of Vaporization (Hv$_{exp}$, Hv$_{calc}$) for 57 Alkanes

No	Hv$_{exp}$ ^{a}	Hv$_{calc}$ ^{b}	No	Hv$_{exp}$ ^{a}	Hv$_{calc}$ ^{b}	No	Hv$_{exp}$ ^{a}	Hv$_{calc}$ ^{b}	No	Hv$_{exp}$ ^{a}	Hv$_{calc}$ ^{b}
2	5.16	6.531	18	34.26	34.122	33	39.64	38.966	78	49.71	49.393
3	14.79	13.765	19	32.88	33.273	34	36.91	38.406	80	49.34	48.584
4	21.02	20.126	20	33.03	33.276	35	35.14	35.622	83	47.13	47.596
5	19.23	18.717	21	35.22	34.607	36	37.27	38.695	151	56.58	56.864
6	26.43	25.965	22	32.05	33.925	37	37.75	37.846	152	61.52	61.814
7	24.85	24.574	23	41.49	41.807	38	38.52	38.368	153	66.68	66.765
8	21.84	22.057	24	39.67	40.006	39	37.99	38.011	154	71.73	71.716
9	31.56	31.400	25	39.83	39.553	40	42.90	40.854	155	76.77	76.648
10	29.89	29.904	26	39.69	39.433	41	46.55	46.866	156	81.35	81.585
11	30.28	29.899	27	37.28	37.245	58	40.16	39.777	157	86.47	86.537
12	27.68	28.259	28	38.78	38.516	61	41.41	40.944	158	91.44	91.485
13	29.12	29.363	29	37.76	37.672	70	38.49	37.429	160	101.81	101.417
14	36.57	36.681	30	37.85	37.886	75	42.00	42.122			
15	34.87	34.874	31	37.53	37.568	76	51.42	51.873			
17	32.42	32.753	32	38.97	38.405	77	49.63	49.962			

^{a} From Ref. [20] ^{b} Calculated with Eq. (11)

In Table 8, the estimated values of D at 25°C are very close to the experimental ones. There is no large deviation in all estimated values. The largest deviation is No. 28 (D_{exp} = 0.6912g·cm^{-3}, ΔD = 0.0153 g·cm^{-3}), but the relative error 2.2% indicates that it is also within the reasonable margin. Maybe one would note that the relationship between the densities (D) with the group II parameters is better than that with the group I (see Table 3). However, the nonlinear model with group I has better result. While using the CV technique, results with q^2 = 0.969 and *PRESS* = 0.0066g·cm^{-3} were obtained. Karelson and Perkson [27] have developed a correlation equation for the normal densities of organic liquid compounds containing various heteroatoms and obtained results with r^2 = 0.9749 and s = 0.046g·cm^{-3} with intrinsic density ρ_R and the total molecular electrostatic interaction energy per atom $E_{el·stat}$.

The correlation of refraction index (n_D) with group I parameters gets relatively low correlation coefficient r = 0.9499 which reminds us that perhaps some important molecular information about the condensed structure is still not counted. It is worth mentioning that we obtained satisfactory result with r = 0.9889, s = 0.0025 when three outliers (No. 40, n_D = 1.4695, Δn_D = –0.0513; No. 144, n_D = 1.4361, Δn_D = 0.0135 and No. 151, n_D = 1.4398, Δn_D = –0.0202) are excluded.

Vapor pressure has stimulated interests of many chemical engineers and environmental protecting scientists. Substances with very low vapor pressure are less likely to vaporize. Vapor pressure can be used to derive other relevant physical property such as Henry's law constant. However, it is very difficult to determine accurately because some substance has very low vaporization. Therefore, people often predict vapor pressure on the basis of molecular structure when experimental data are deficient. McClelland and Jurs [21] has created a 10–descriptor model for the prediction of the logPv of diverse organic compounds and obtained result with rms (root mean square) = 0.33 log unit. In the present study, the five–descriptor equation attained a satisfactory correlation (r = 0.9974) and low standard deviation (s = 0.0959) for the logarithm of vapor pressure. This provides a new efficient and reliable way to estimate vapor pressure for alkanes.

The two thermodynamic properties Cp and Hv are investigated in this paper. The high correlation coefficient (r = 0.9997) and low standard deviation (s = 2.02J·mol^{-1}·K^{-1}) of eq.10 indicate the excellent relationship between the heat capacities (Cp) and the group I parameters. There is no large deviation in all estimated values. The heat of vaporization (Hv) is also perfectly correlated with five parameters SX_{1CH}, SX_{1CC}, SV_{ij}, OEI and $N^{2/3}$ (r = 0.9994, s = 0.66kJ·mol^{-1}). The largest s is No. 40 (Hv_{exp} = 42.9kJ·mol^{-1}, ΔHv = –2.05kJ·mol^{-1}), but the relative error is –4.8%, which is also within reasonable margin. And the significant cross–validated correlation coefficient q^2 = 0.999, deviation *PRESS* = 2.37J·mol^{-1}·K^{-1} for Cp and q^2 = 0.997, *PRESS* = 1.05 kJ·mol^{-1} for

Hv indicate the predictive potential of the models for untested compounds of these two class, respectively.

Table 13. Experimental and Estimated Chromatographic Retention Index for 64 Alkanes

No	Ipt$_{exp}$ ^{a}	Ipt$_{calc}$ ^{b}	No	Ipt$_{exp}$ ^{a}	Ipt$_{calc}$ ^{b}	No	Ipt$_{exp}$ ^{a}	Ipt$_{calc}$ ^{b}	No	Ipt$_{exp}$ ^{a}	Ipt$_{calc}$ ^{b}
3	300.0	295.37	21	683.0	677.07	42	864.8	865.35	61	814.9	830.15
4	400.0	395.93	22	631.4	636.20	43	871.4	870.52	62	812.2	798.15
5	354.2	364.62	23	800.0	789.65	44	863.7	867.09	63	847.0	839.74
6	500.0	492.17	24	764.1	763.66	45	818.9	816.55	64	847.0	853.19
7	466.1	473.96	25	772.1	772.35	46	856.3	859.87	65	827.3	845.83
9	600.0	591.55	26	765.6	764.29	47	823.1	832.21	67	856.3	858.29
10	562.0	565.36	27	721.6	721.92	48	835.5	829.73	76	1000.0	990.19
11	578.6	578.39	28	757.9	758.46	49	830.1	823.25	77	964.0	964.36
12	528.5	533.91	29	731.7	742.03	50	839.7	836.28	78	970.5	971.64
13	557.7	563.62	30	729.7	744.96	51	858.8	857.56	79	961.6	966.90
14	700.0	689.59	31	739.2	734.17	52	837.6	860.73	80	959.8	967.28
15	662.9	667.18	32	767.6	763.48	54	870.3	871.24	96	889.9	879.12
16	672.2	672.14	33	775.0	772.91	56	823.1	827.34	108	913.2	906.92
17	620.5	618.01	35	688.0	680.06	57	791.6	782.09			
18	665.0	662.47	37	747.6	741.04	58	785.5	801.57			
19	625.8	632.95	38	759.7	755.73	59	839.7	834.57			
20	650.5	646.19	41	900.0	889.17	60	850.2	843.91			

^{a} From Ref. [22] ^{b} Calculated with Eq. (12)

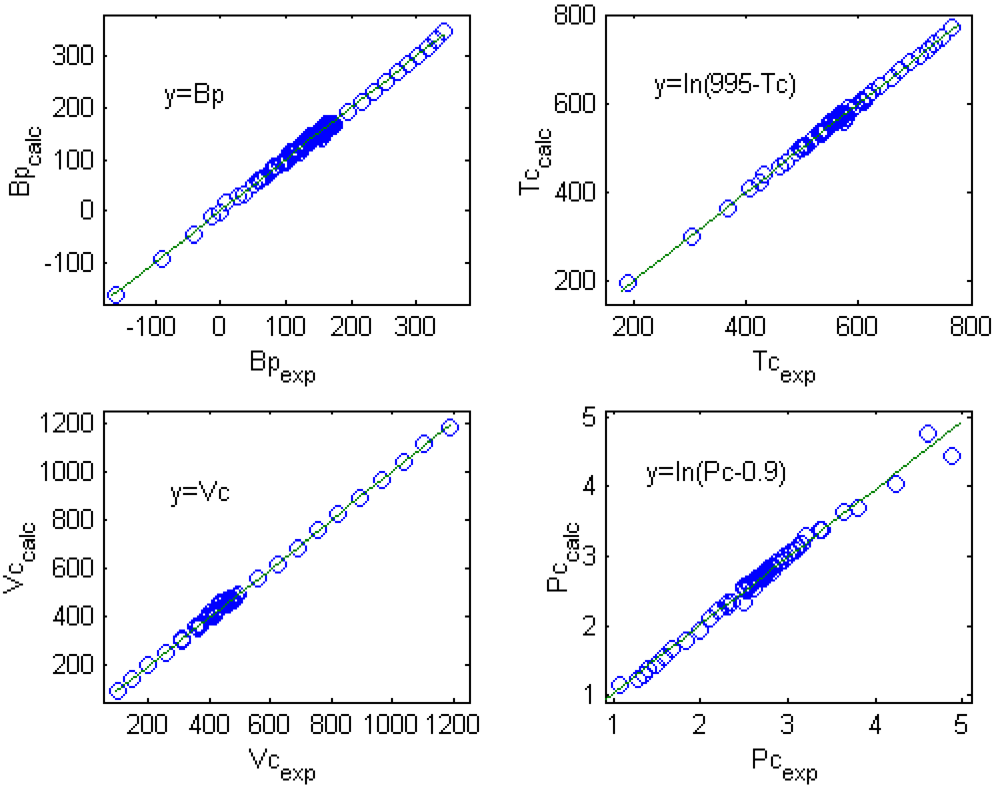

Figure 5. Plot of the property (y) estimated vs. observed for alkane 10 properties.

The estimation of chromatographic retention index is a widely investigated topic in QSPR studies. Among the obtained model parameters, topological index are frequently used because they can be derived directly from the molecular structure. Several authors have estimated retention indices using topological descriptor [28–31]. Liu *et al*. [22] have developed a QSPR model between the μ vector (containing 10 elements) and chromatographic retention indices of 64 alkanes and obtained r = 0.9992, s = 5.938. In the present study, the relation of chromatographic retention indices of the same sample with five parameters obtained result with r = 0.9986, s = 8.09. When the parameter $N^{2/3}$ is replaced by N (the number of carbon atom), the result is somewhat improved (r = 0.9987, s = 7.76). Though the result is not as good as that of literature, the method and the parameters used in this paper are simple and the result is satisfactory. To intuitively express the relationship of property to the group I parameters, the plots of the estimated properties (y_{calc}) by QSPR models vs. the experimental properties (y_{exp}) are shown in Figure 5.

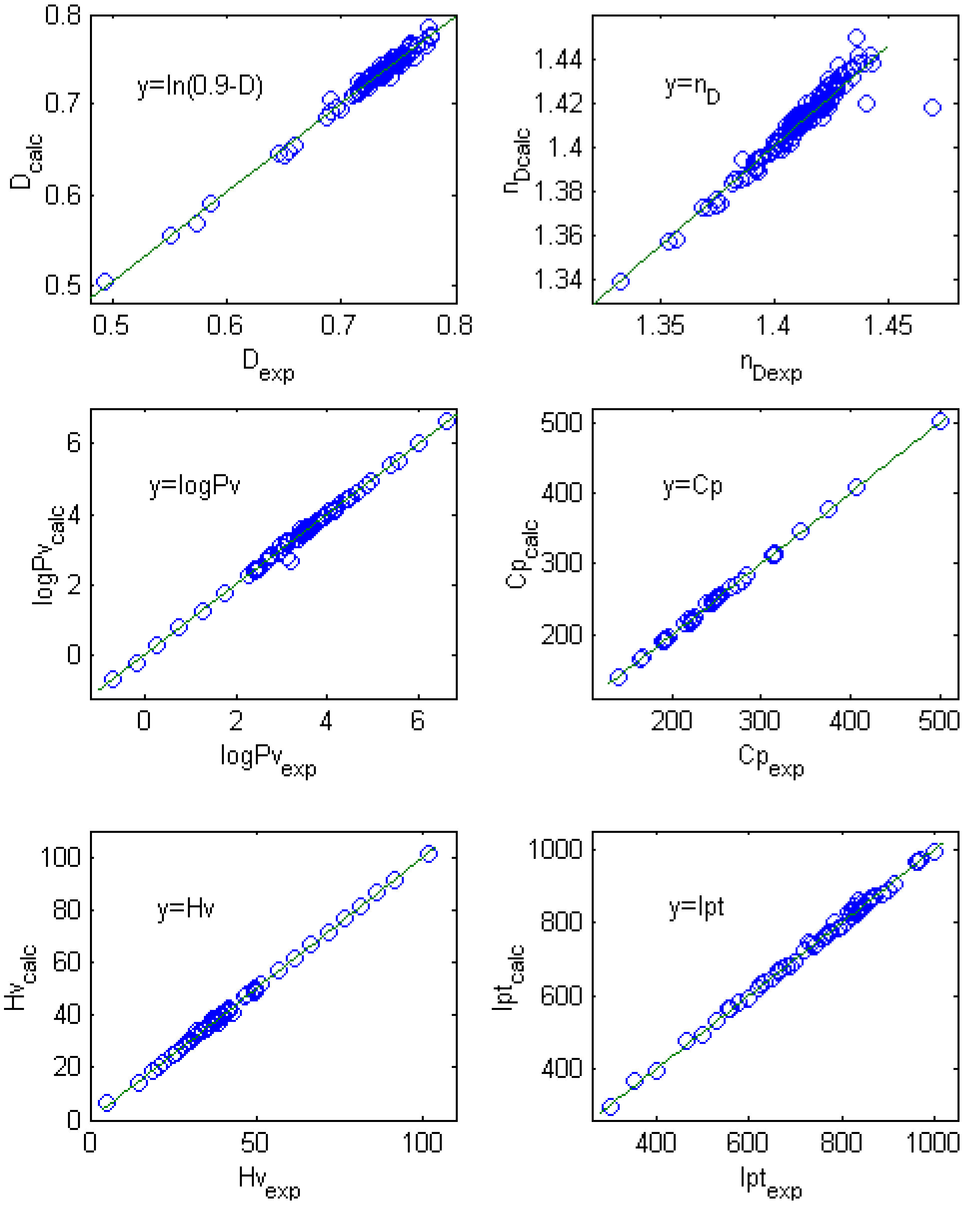

Figure 5. (Continued).

Most of QSPR research only investigated one or several properties correlation with some parameters or descriptors. In this paper, we have obtained good correlations between eigenvalues of bond adjacency matrix and the 10 physicochemical properties of alkanes. This study demonstrates that the theoretical predicative models for physicochemical properties of alkanes based on PEI, have powerful predicative capability, general applicability and are easy to use. In the literature [32], O.Invaciuc et al. have developed several QSPR models for Bp (r = 0.9939, s = 2.97°C), D (r = 0.9902, s = 3.77 kg·m^{-3}), n$_D^{25}$ (r = 0.9840, s = 0.0025) and Cp (r = 0.9883, s = 3.92 J·mol^{-1}·K^{-1}) with three independent variables by selecting the best combination from the 45 structural descriptors of alkanes C_6–C_{10}. We also noted that Liu *et al.* [33] have constructed several QSPR models between molecular electronegative distance vector (MEDV) and physical properties or thermodynamic functions of alkanes , and obtained the estimated rms values of Bp, D, n$_D$, Tc, Pc, Cp and Hv for alkanes are respectively 3.22K, 0.0286 g·cm^{-3}, 0.0021, 5.30 K, 0.58 atm, 3.81 J·mol^{-1}·K^{-1} and 0.22 kJ·mol^{-1}, and the estimated r is 0.9978, 0.9940, 0.9884, 0.9956, 0.9909, 0.9886 and 0.9991. In our present paper, s is 3.06°C, 0.0048 g·cm^{-3}, 0.0055, 4.03°C, 0.0771 Mpa, 2.02 J·mol^{-1}·K^{-1} and 0.66 kJ·mol^{-1}, and r is 0.9986, 0.9906, 0.9499, 0.9991, 0.9941, 0.9997 and 0.9994, respectively. Known from the results above, most of results in the current paper are better than those in the literature [33], and the number of parameters employed by this paper is only half of that employed by Liu *et al*. [33].

4 CONCLUSIONS

The model equations developed by present paper can be used for estimating and predicting the 10 physicochemical properties of alkanes, and their reliability have been confirmed by cross–validation with the results, Bp (q^2 = 0.997, *PRESS* = 3.22 °C), Tc (q^2 = 0.996, *PRESS* = 6.28 °C), Vc (q^2 = 0.999, *PRESS* = 8.43 cm^3·mol^{-1}), Pc (q^2 = 0.935, *PRESS* = 0.1514 Mpa), D (q^2 = 0.969, *PRESS* = 0.0066 g·cm^{-3}), n_D (q^2 = 0.879, *PRESS* = 0.0060), logPv (q^2 = 0.993, *PRESS* = 0.1110), Cp (q^2 = 0.999, *PRESS* = 2.37 J·mol^{-1}·K^{-1}), Hv (q^2 = 0.997, *PRESS* = 1.05 kJ·mol^{-1}) and Ipt (q^2 = 0.997, *PRESS* = 9.14). The predicted values in good agreement with the experimental ones and the good correlations indicate that the eigenvalues of bond adjacency matrix combining with odd–even index and the number of carbon atoms can express most of the important molecular structural information of alkanes. This method may be extended to QSPR of other compounds. Further research along these lines will continue and is in progress.

Acknowledgment

The project was supported by the National Natural Science Foundation of China No. 20172043 (NSFC) and Hunan Province Education Commission (HPEC).

5 REFERENCES

[1] X. Gironés, L. Amat, D. Robert, and R. Carbó–Dorca, Use of Electron–electron Repulsion Energy as a Molecular Descriptor in QSAR and QSPR Studies. *J. Comput.–Aided Mol. Des.* **2000**, *14*, 477–485.

[2] G. Hua, S.Veerabahu, and L. Pil, Estimation of Aqueous Solubility of Organic Compounds with QSPR Approach. *Pharmacol. Res.* **2002**, *4*, 497–503.

[3] C. Yin, X. Liu, W. Guo, T. Lin, X. Wang, and L. Wang, Prediction and Application in QSPR of Aqueous Solubility of Sulfur–containing Aromatic Esters Using GA–based MLR with Quantum Descriptors. *Water Res.* **2002**, *36*, 2975–2982.

[4] S. Grigoras, A Structural Approach to Calculated Physical Properties of Pure Organic Substances: The Critical Temperature, Critical Volume and Related Properties. *J. Comput. Chem.* **1990**, *4*, 493–510.

[5] O. Ivanciuc, T. Ivanciuc, and A. T. Balaban, Quantitative Structure–Property Relationship Study of Normal Boiling Points Halogen–/Oxygen–/Sulfur–Cotaining Organic Compounds Using the CODESSA Program. *Tetrahedron* **1998**, *54*, 9129–9142.

[6] N. Guevara, Fragmental graph. A Novel Approach to Generate a New Family of Descriptors Applications to QSPR Studies, *J. Mol. Struct.* (*Theochem*) **1999**, *493*, 29–36.

[7] E. Estrada, Edge Adjacency Relationships and a Novel Topological Index Related to Molecular Volume, *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 31–33.

[8] D. Cvetković, and P. W. Fowler, A Group–Theoretical Bound for the Number of Main Eigenvalues of a Graph, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 638–641.

[9] H. P. Schultz, Topological Organic Chemistry. 2. Graph Theory, Matrix Determinants and Eigenvalus, and Topological Indices of Alkanes, *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 27–209.

[10] A. T. Balaban, D. Ciubotariu, and M. Medeleanu, Topological Indices and Real Number Vertex Invariants Based on Graph Eigenvalues or Eigenvectors, *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 517–523.

[11] F. Torrens, Computing the Permanent of the Adjacency Matrix for Fullerenes, *Internet Electron. J. Mol. Des.* **2002**, *1*, 351–359, http://www.biochempress.com.

[12] C. Cao and Z. Li, On Molecular Polarizability. 1. Relationship to Water Solubility of Alkanes and Alcohols. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1–7.

[13] C. Cao, S. Liu, and Z. Li, On Molecular Polarizability. 2. Relationship to the Boiling Point of Alkanes and Alcohols. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1105–1111.

[14] C. Cao, H. Yuan, S. Liu, and R. Zeng, On Molecular Polarizability. 3. Relationship to the Ionization Potential of Haloalkanes, Amines, Alcohols, and Ethers. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1010–1014.

[15] C. Cao and H. Yuan, On Molecular Polarizability. 4. Evaluation of the Ionization Potential for Alkanes and Alkenes with Poarizability . *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 667–672.

[16] C. Cao and H. Yuan, A New Approach of Evaluating Bond Dissociation Energy form Eigenvalue of Bond Orbital–Connection Matrix for C–C, C–H Bonds in Alkane. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 600–608.

[17] C. Cao and H. Yuan, Topological Indices Based on Vertex, Distance, and Ring: On the Boiling Points of Paraffines and Cycloalkanes. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 867–877.

[18] H. Yuan and C. Cao, Topological Indices Based on Vertex, Edge, Ring and Distance: Application to Various Physico–Chemical Properties of Diverse Hydrocarbons. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 501–512.

[19] C. Cao and H. Yuan, A Modified Distance Matrix to Distinguish *Cis*/*Trans* Isomers of Cycloalkanes, *Internet Electron. J. Mol. Des.* **2002**, *1*, 401–409, http://www.biochempress.com.

[20] CRC. Handbook of Chemistry and Physics, 81st ed.; David, R. L. Ed.; KRC Press: **2000–2001**.

[21] H. E. McClelland and P. C. Jurs, Quantitatively Structure–Property Relationships for the Prediction of Vapor Pressure of Organic Compounds from Molecular Structures. *J. Chem. Inf. Comput. Sci.* **2000**, 40, 967–975.

[22] S. Liu, H. Liu, Zh. Xia, C. Cao, and Z. Li, Molecular Distance–Edge Vector (μ) and Chromatographic Retention Index of Alkanes. *J. Chin. Chem. Soc.* **2000**, *47*, 455–460.

[23] M. D. Wessel and P. C. Jurs, Prediction of Normal Boiling Points of Hydrocarbons from Molecular Structure. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 68–76.

[24] G. Espinosa, D. Yaffe, Y. Cohen, A. Arenas, and F. Giralt, Neural Network Based Quantitative Structural Property Relations (QSPRs) for Predicting Boiling Points of Aliphatic Hydrocarbons. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 859–879.

[25] R. Gerta and R. Christoph, On Topological Indices, Boiling Points, and Cycloalkanes. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 788–802.

[26] A. R. Katritzky, M. Lan, and M. Karelson, Relationships of Critical Temperatures to Calculated Molecular Properties. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 293–299.

[27] M. Karelson and A. Perkson, QSPR Prediction of Densities of Organic Liquids. *Comput. Chem.* **1999**, *23*, 49–59.

[28] M. Pompe and M. Novic, Prediction of Gas–Chromatographic Retention Indices Using Topological Descriptors.

J. Chem. Inf. Comput. Sci. **1999**, *39*, 59–67.

[29] Ch. Duvenbeck and P. Zinn, List Operations on Chemical Graph.3. Development of Vertex and Edge Models for Fitting Retention Index Data. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 211–219.

[30] Ch. Duvenbeck and P. Zinn, List Operations on Chemical Graph. 4. Using Edge Models for Prediction of Retention Index Data. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 220–230.

[31] B. S. Junkes, R. D. M. C. Amboni, R. A. Yunes, and V. E. F. Heinzen, Semiempirical Topological Index: A Novel Molecular Descriptor for Quantitative Structure–Retention Relationship Studies, *Internet Electron. J. Mol. Des.* **2003**, *2*, 33–49, http://www.biochempress.com.

[32] O. Ivanciuc, T. Ivanciuc, and A. T. Balaban, Quantitative Structure–Property Relationship Evaluation of Structural Descriptors Derived from the Distance and Reverse Wiener Matrices, *Internet Electron. J. Mol. Des.* **2002**, *1*, 467–487, http://www.biochempress.com.

[33] S. Liu, S. Cai, C. Cao, and Z. Li, Molecular Electronegative Distance Vector (MEDV) Related to 15 Properties of Alkanes. *J. Chem. Inf. Comput. Sci.* **2000**, 40, 1337–1348.

Biographies

Chenzhong Cao is professor of organic chemistry at the Department of Chemistry, Xiangtan Normal University , People’s Republic of China.