AN IMPROVED TOPOLOGICAL DESCRIPTOR ED_m' AND ITS APPLICATIONCHANGMING NIE^{1,*}, YAXIN WU¹, RONGYAN WU² and SONGNIAN WEN²¹School of Chemistry and Chemical Engineering, University of South China, Hengyang, Hunan Province 421001, P. R. China²College of electrical Engineering, University of South China., Hengyang, Hunan Province 421001, P. R. China

(Received: June 30, 2010 - Accepted: January 26, 2011)

ABSTRACT

In this paper, on the basis of the topological index ED_m derived from ionicity index matrix, improved distance matrix and branching degree matrix, we proposed the new topological descriptor ED_m' by introducing the bond angle into hidden hydrogen graph of molecules and using the geometric distance instead of the sum of bond length between two vertexes. The ED_m' describes the molecular structure more accurately, and realizes unique characterization to cis-trans isomers. The quantitative structure-property relationship (QSPR) models, with correlation coefficients (R) in the range of 0.984-1.000 for the boiling point ($b.p.$), the standard enthalpy of formation (ΔH_m°), the molar refraction (R_m) and the molar volume (V_m) of some cis-trans isomers for alkenes, are subsequently developed by index ED_m' . Moreover, the good stability and predictive ability of the models were demonstrated by LOO (leave-one-out) method and RSP (random sampling prediction) method, which further manifests the index ED_m' has high potential of wide applications in QSPR study.

Key words: cis/trans isomers, ionicity index matrix, improved distance matrix, branching degree matrix

1. INTRODUCTION

Since the Wiener index W introduced in 1947¹, many chemists such as Balaban A.T.²⁻⁵, Randic M.⁶⁻¹⁰, Gutman I.¹¹⁻¹⁴, Trinajstić N.¹⁵⁻¹⁸, and Estrada E.¹⁹⁻²¹ have taken lots of time and gotten many brilliant achievements in quantitative structure-property/activity relationships (QSPR/QSAR) study. These studies have played significant roles in various fields such as drug design, physicochemical property/ biological activity forecast, environmental pollutant judgment, etc²²⁻³⁰. Recently, our group has proposed some new topological descriptor, such as ED_m index³¹, PE index³² and Nt index³³, which have been applied to QSPR study for series of compounds.

Stereochemistry is one of the crucial parts in organic chemistry, and the cis-trans isomerism that belongs to diastereoisomeric category, is of practical importance in stereochemistry³⁴. The cis-trans isomers have diverse chemical and physical properties, such as biological activity and inhibiting ability. Many experimental results demonstrated that stereo structural character has a great influence on chemical and vital processes. For instance, the cis-trans isomerism for alkenes often determines whether a reaction could go on wheels or not. Chiral features of drug molecules usually have decisive effect on drug effect³⁵. So it is very necessary to explore the QSPR/QSAR models for cis/trans-isomers.

The traditional topological indices, although denote the molecular structural features effectively, are not able to differentiate the stereo structural characteristics of molecules³⁵. More and more attention has been given to develop new topological descriptors adapted to cis/trans isomers³⁵⁻³⁹. Chenzhong Cao and Hua Yuan presented the modified vertex degree-distance index (MVDI) and modified edge degree-distance index (MEDII), and used them together with the odd-even index (OEI) and the number of carbon atoms (N) to characterize the molecular structural information for the cis- and trans-isomers of alkenes³⁵. Golbraikh, A. et al introduced several series of novel ZE-isomerism descriptors using ZE-isomerism correction, which is added to the vertex degrees of atoms connected by double bonds in Z and E configurations³⁹.

Based on the previous work^{31-33,40-45}, we proposed the index ED_m' by revising ED_m index³¹. The new topological descriptor ED_m' possesses good distinguishing ability to the molecular structures for cis-trans isomers. Superior to reported work, the ED_m' break with the tradition methods that all depend on using correction factor.

2. Theories and methods

For alkene, the double bonded carbons and the four attached atoms lie in a plane. C=C is made up of a σ bond and a π bond. The π bond is formed by sideways overlap of p orbitals and without symmetry axis. Since double bond can not be rotated, two alkenes with the same substituents in different spatial location are different compounds, and known as cis-trans isomers.

In modern quantum mechanics theory, the double bonded carbons in ethylene are sp^2 -hybridized and have three equivalent orbitals, which lie in a plane at angles of 120° to one another. In this paper, we will use the C-C-C angle equal to 120° for sp^2 hybridization and approximate it to 110° for sp^3 hybridization, and consider all the saturated carbon chains as stable zigzag staggered conformations. We introduce the bond angle into hidden hydrogen graph of molecules, use the geometric distance instead of the sum of bond

length between two atoms, and put forward the topological index ED_m' based on branching degree matrix G , improved distance matrix S' , and ionicity index matrix Q .

$$ED_m' = GS_m'Q \quad (m=1, 2, 3) \quad (1)$$

$$G = [g_1 \quad g_2 \quad \cdots \quad g_n] \quad (2)$$

$$S_m' = \begin{bmatrix} (s_m)_{11}' & (s_m)_{12}' & \cdots & (s_m)_{1n}' \\ (s_m)_{21}' & (s_m)_{22}' & \cdots & (s_m)_{2n}' \\ \vdots & \vdots & \ddots & \vdots \\ (s_m)_{n1}' & (s_m)_{n2}' & \cdots & (s_m)_{nn}' \end{bmatrix} \quad (m=1, 2, 3) \quad (3)$$

$$Q = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} \quad (4)$$

Substituting Eq. (2)-(4) into Eq. (1), we obtain

$$ED_m' = [g_1 \quad g_2 \quad \cdots \quad g_n] \cdot \begin{bmatrix} (s_m)_{11}' & (s_m)_{12}' & \cdots & (s_m)_{1n}' \\ (s_m)_{21}' & (s_m)_{22}' & \cdots & (s_m)_{2n}' \\ \vdots & \vdots & \ddots & \vdots \\ (s_m)_{n1}' & (s_m)_{n2}' & \cdots & (s_m)_{nn}' \end{bmatrix} \cdot \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} \quad (m=1, 2, 3) \quad (5)$$

2.1 Branching degree matrix

In Eq. (5), branching degree⁴⁰ g_i is figured as

$$g_i = (\sum k)^{-0.5} \quad (6)$$

Where $\sum k$ is the sum of the number of bonded non-hydrogen single bonds of atoms numbered i . Double bond is regarded as two single bonds and triple bond as three single bonds when calculated. For primary, secondary, tertiary, and quaternary carbons, obviously, g_i values are 1, 0.7071, 0.5774 and 0.5, respectively.

Table 1. Some branching degree matrices G .

No.	Compound*	G
1	cis-2-butene	[1 0.5774 0.5774 1]
2	trans-2-butene	[1 0.5774 0.5774 1]
3	cis-2-pentene	[1 0.5774 0.5774 0.7071 1]
4	trans-2-pentene	[1 0.5774 0.5774 0.7071 1]
5	cis-2-hexene	[1 0.5774 0.5774 0.7071 0.7071 1]
6	trans-2-hexene	[1 0.5774 0.5774 0.7071 0.7071 1]
7	cis-3-hexene	[1 0.7071 0.5774 0.5774 0.7071 1]
8	trans-3-hexene	[1 0.7071 0.5774 0.5774 0.7071 1]
9	cis-4-m-2-pentene	[1 0.5774 0.5774 0.5774 1 1]
10	trans-4-m-2-pentene	[1 0.5774 0.5774 0.5774 1 1]
11	cis-2-heptene	[1 0.5774 0.5774 0.7071 0.7071 0.7071 1]
12	trans-2-heptene	[1 0.5774 0.5774 0.7071 0.7071 0.7071 1]
13	cis-3-heptene	[1 0.7071 0.5774 0.5774 0.7071 0.7071 1]
14	trans-3-heptene	[1 0.7071 0.5774 0.5774 0.7071 0.7071 1]
15	cis-4-m-2-hexene	[1 0.5774 0.5774 0.5774 0.7071 1 1]
16	trans-4-m-2-hexene	[1 0.5774 0.5774 0.5774 0.7071 1 1]
17	cis-5-m-2-hexene	[1 0.5774 0.5774 0.7071 0.5774 1 1]
18	trans-5-m-2-hexene	[1 0.5774 0.5774 0.7071 0.5774 1 1]
19	cis-2-m-3-hexene	[1 0.5774 0.5774 0.5774 0.7071 1 1]
20	trans-2-m-3-hexene	[1 0.5774 0.5774 0.5774 0.7071 1 1]

* m = methyl

2.2 Improved distance matrix

For molecular graph of n vertexes, the distance matrix is a $n \times n$ symmetric matrix. We depict the improved distance matrix as

$$S'_m = \left[(s_m)_{ij} \right] = \begin{cases} 1/(l_m)_{ij}, & m \text{ bond distance between the vertex } i \text{ and } j \\ 0, & \text{otherwise} \end{cases} \quad (m=1, 2, 3) \quad (7)$$

In Eq. (7), $(l_m)_{ij}$ is the ratio of the geometric distance between the two vertexes i and j to the bond-length of C—C, and defined as the relative distance. The actual bond lengths of C—C, C=C, C≡C are 154, 134, 120 nm⁴⁶. For example, Fig. 1 is the hidden hydrogen graph of cis-2-pentene.

Figure 1. Hidden hydrogen graph of cis-2-pentene. $(l_m)_{ij}$ in Figure 1

$$\left. \begin{aligned} (l_1)_{12}=1, (l_1)_{23}=134/154=0.87, (l_1)_{34}=(l_1)_{45}=(l_1)_{12}=1; \\ \angle 123=120^\circ \quad \angle 345=110^\circ \\ (l_1)_{12}=1 \quad \Rightarrow (l_2)_{13}=1.6209, (l_1)_{34}=1 \quad \Rightarrow (l_2)_{35}=1.6383, (l_2)_{24}=(l_2)_{13}=1.6209; \\ (l_1)_{23}=0.8701 \quad (l_1)_{45}=1 \end{aligned} \right\} \Rightarrow (l_2)_{35}=1.6383, (l_2)_{24}=(l_2)_{13}=1.6209;$$

$$\left. \begin{aligned} \angle 123=120^\circ \quad \angle 234=120^\circ \\ \angle 234=120^\circ \quad \angle 345=110^\circ \\ (l_1)_{12}=(l_1)_{34}=1 \quad (l_1)_{45}=(l_1)_{34}=1 \end{aligned} \right\} \Rightarrow (l_3)_{14}=1.8701, (l_3)_{25}=2.4545.$$

$$\left. \begin{aligned} \angle 123=120^\circ \\ \angle 234=120^\circ \\ (l_1)_{12}=(l_1)_{34}=1 \quad (l_1)_{23}=0.8701 \end{aligned} \right\} \Rightarrow (l_3)_{14}=1.8701.$$

The value of $(l_3)_{25}$ is calculated as follows (see Fig.1.b2):

$$\left. \begin{aligned} \angle 364=90^\circ \\ \angle 436=60^\circ \quad \Rightarrow l_{46}=0.8660; \\ (l_1)_{34}=1 \end{aligned} \right\} \Rightarrow l_{58}=l_{46}-l_{57}=0.8660-0.1737=0.6923$$

$$\left. \begin{aligned} \angle 645=\angle 754=(110^\circ - \angle 346)=80^\circ \\ \angle 574=90^\circ \\ (l_1)_{45}=1 \end{aligned} \right\} \Rightarrow l_{57}=0.1737;$$

$$\left. \begin{aligned} \angle 645=\angle 754=(110^\circ - \angle 346)=80^\circ \\ \angle 574=90^\circ \\ (l_1)_{45}=1 \end{aligned} \right\} \Rightarrow l_{68}=l_{47}=0.9848;$$

$$\left. \begin{aligned} \angle 364=90^\circ \\ \angle 436=60^\circ \quad \Rightarrow l_{36}=0.5 \quad ; (l_1)_{23}=0.8701 \\ (l_1)_{34}=1 \end{aligned} \right\} \Rightarrow l_{28}=(l_1)_{23}+l_{36}+l_{68}=0.8701+0.5+0.9848=2.3549$$

$$\left. \begin{aligned} l_{58}=0.6923 \\ l_{28}=2.3549 \quad \Rightarrow (l_3)_{25}=2.4545; \\ \angle 285=90^\circ \end{aligned} \right\}$$

The improved distance matrix S'_m of cis-2-pentene

$$S'_1 = \begin{bmatrix} 0 & 1/(a_1)_{12} & 0 & 0 & 0 \\ 1/(a_1)_{21} & 0 & 1/(a_1)_{23} & 0 & 0 \\ 0 & 1/(a_1)_{32} & 0 & 1/(a_1)_{34} & 0 \\ 0 & 0 & 1/(a_1)_{43} & 0 & 1/(a_1)_{45} \\ 0 & 0 & 0 & 1/(a_1)_{54} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0.87 & 0 & 0 \\ 0 & 0.87 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$S'_2 = \begin{bmatrix} 0 & 0 & 1/(a_2)_{13} & 0 & 0 \\ 0 & 0 & 0 & 1/(a_2)_{24} & 0 \\ 1/(a_2)_{31} & 0 & 0 & 0 & 1/(a_2)_{35} \\ 0 & 1/(a_2)_{42} & 0 & 0 & 0 \\ 0 & 0 & 1/(a_2)_{53} & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1/1.6209 & 0 & 0 \\ 0 & 0 & 0 & 1/1.6209 & 0 \\ 1/1.6209 & 0 & 0 & 0 & 1/1.6383 \\ 0 & 1/1.6209 & 0 & 0 & 0 \\ 0 & 0 & 1/1.6383 & 0 & 0 \end{bmatrix}$$

$$S'_3 = \begin{bmatrix} 0 & 0 & 0 & 1/(a_3)_{14} & 0 \\ 0 & 0 & 0 & 0 & 1/(a_3)_{25} \\ 0 & 0 & 0 & 0 & 0 \\ 1/(a_3)_{41} & 0 & 0 & 0 & 0 \\ 0 & 1/(a_3)_{52} & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1/1.8701 & 0 \\ 0 & 0 & 0 & 0 & 1/2.4545 \\ 0 & 0 & 0 & 0 & 0 \\ 1/1.8701 & 0 & 0 & 0 & 0 \\ 0 & 1/2.4545 & 0 & 0 & 0 \end{bmatrix}$$

For trans-2-pentene, the calculations are as follows.

Figure 2. Hidden hydrogen graph of trans-2-pentene. $(l_m)_{ij}$ in Figure 2

$$\left. \begin{aligned} (l_1)_{12}=1, (l_1)_{23}=134/154=0.87, (l_1)_{34}=(l_1)_{45}=(l_1)_{12}=1; \\ \angle 123=120^\circ \quad \angle 345=110^\circ \\ (l_1)_{12}=1 \quad \Rightarrow (l_2)_{13}=1.6209, (l_1)_{34}=1 \quad \Rightarrow (l_2)_{35}=1.6383, (l_2)_{24}=(l_2)_{13}=1.6209; \\ (l_1)_{23}=0.8701 \quad (l_1)_{45}=1 \end{aligned} \right\} \Rightarrow (l_2)_{35}=1.6383, (l_2)_{24}=(l_2)_{13}=1.6209;$$

$$\left. \begin{aligned} \angle 123=120^\circ \quad \angle 234=120^\circ \\ \angle 234=120^\circ \quad \angle 345=110^\circ \\ (l_1)_{12}=(l_1)_{34}=1 \quad (l_1)_{45}=(l_1)_{34}=1 \end{aligned} \right\} \Rightarrow (l_3)_{14}=2.5490, (l_3)_{25}=2.4545.$$

$$\left. \begin{aligned} \angle 123=120^\circ \\ \angle 234=120^\circ \\ (l_1)_{12}=(l_1)_{34}=1 \quad (l_1)_{23}=0.8701 \end{aligned} \right\} \Rightarrow (l_3)_{14}=2.5490, (l_3)_{25}=2.4545.$$

Values of S'_1 and S'_2 of trans-2-pentene are identical to that of cis-2-pentene, so we only list the S'_3 of trans-2-pentene.

$$S'_3 = \begin{bmatrix} 0 & 0 & 0 & 1/(a_3)_{14} & 0 \\ 0 & 0 & 0 & 0 & 1/(a_3)_{25} \\ 0 & 0 & 0 & 0 & 0 \\ 1/(a_3)_{41} & 0 & 0 & 0 & 0 \\ 0 & 1/(a_3)_{52} & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1/2.5490 & 0 \\ 0 & 0 & 0 & 0 & 1/2.4545 \\ 0 & 0 & 0 & 0 & 0 \\ 1/2.5490 & 0 & 0 & 0 & 0 \\ 0 & 1/2.4545 & 0 & 0 & 0 \end{bmatrix}$$

2.3 Ionicity index matrix

In Eq. (5), q_i is defined as

$$q_i = (X_{i,INI})^{0.5} \quad (8)$$

$X_{i,INI}$ is the ionicity index⁴³ of the atom i in molecules and is defined as

$$X_{i,INI} = 1/2 (X_{iE} - X_{iA}) / X_{iA} \cdot 1/2 \cdot 100 \quad (9)$$

where X_{iA} is Pauling electronegativity of atom i , and X_{iE} is the equilibrium electronegativity⁴² of atom i in a molecule. The equilibrium electronegativity X_{iE} is figured as

$$X_{iE} = (X_{iA} + \sum l X_{iG}) / (1 + \sum l) \quad (10)$$

In Eq. (10), $\sum l$ is the sum of atoms or branching groups directly attached to the atom i , and $\sum X_{iG}$ is the sum of electronegativities of atoms or groups directly attached to the atom i . X_G^{44} is given by

$$X_G = \left\{ \frac{1}{n_{1l}} \sum_{l=1}^{n_{1l}} \frac{1}{n_{2l}} \left[\sum_{l=1}^{n_{2l}} \dots \left(\frac{1}{n_{kl}} \sum_{l=1}^{n_{kl}} X_{kl} \right) \dots \right] \right\} \quad (11)$$

As shown in Figure 3, the group can be divided into k layers, and the left atoms of the dotted line labeled 1, 2, 3, ... k are known as the ground atoms. The right suffix of atoms l is the numeration of the ground atom. $n_{1l}, n_{2l}, \dots, n_{kl}$ are the sum of the ground atom l and other atoms or groups directly attached to the ground atoms in the grade 1, 2, 3, ... k . X_{kl} are the sum of electronegativities of the ground atom l and other atoms or branching groups directly attached to the ground atoms in the grade 1, 2, 3, ... k .

Figure 3. Schematics of group structure.

For instance, Figure 4 is the structural graph of cis-2-pentene.

Figure 4. Plot of molecular structure of cis-2-pentene.

X_G of $-\text{CH}_3$, $-\text{CH}_2\text{CH}_3$, $=\text{CHCH}_3$, $-\text{CH}=\text{CHCH}_3$, $=\text{CHCH}_2\text{CH}_3$, $-\text{CH}=\text{CHCH}_2\text{CH}_3$ and $-\text{CH}_2\text{CH}=\text{CHCH}_3$

$$X_{-\text{CH}_3} = \frac{1}{4} \times (X_C + 3 \times X_H) = \frac{1}{4} \times (2.55 + 3 \times 2.2) = 2.2875$$

$$X_{-\text{CH}_2\text{CH}_3} = \frac{1}{4} \times (X_C + 2 \times X_H + X_{-\text{CH}_3}) = \frac{1}{4} \times (2.55 + 2 \times 2.2 + 2.2875) = 2.3094$$

$$X_{-\text{CHCH}_3} = \frac{1}{3} \times (X_C + X_H + X_{-\text{CH}_3}) = \frac{1}{3} \times (2.55 + 2.2 + 2.2875) = 2.3458$$

$$X_{-\text{CH}=\text{CHCH}_3} = \frac{1}{3} \times (X_C + X_H + X_{-\text{CHCH}_3}) = \frac{1}{3} \times (2.55 + 2.2 + 2.3458) = 2.3653$$

$$X_{-\text{CHCH}_2\text{CH}_3} = \frac{1}{3} \times (X_C + X_H + X_{-\text{CH}_2\text{CH}_3}) = \frac{1}{3} \times (2.55 + 2.2 + 2.3094) = 2.3531$$

$$X_{-\text{CH}=\text{CHCH}_2\text{CH}_3} = \frac{1}{3} \times (X_C + X_H + X_{-\text{CHCH}_2\text{CH}_3}) = \frac{1}{3} \times (2.55 + 2.2 + 2.3531) = 2.3677$$

$$X_{-\text{CH}_2\text{CH}=\text{CHCH}_3} = \frac{1}{4} \times (X_C + 2 \times X_H + X_{-\text{CH}=\text{CHCH}_3}) = \frac{1}{4} \times (2.55 + 2 \times 2.2 + 2.3653) = 2.3288$$

Where X_C and X_H are the Pauling electronegativities of carbon atom and hydrogen atom.

X_E of carbon atoms labeled 1, 2, 3, 4 and 5 in Fig. 4

$$X_{1,E} = \frac{1}{5} \times [X_C + 3 \times X_H + X_{-\text{CH}=\text{CHCH}_2\text{CH}_3}] = \frac{1}{5} \times [2.55 + 3 \times 2.2 + 2.3677] = 2.3035$$

$$X_{2,E} = \frac{1}{4} \times [X_{-\text{CH}_3} + X_C + X_H + X_{-\text{CHCH}_2\text{CH}_3}] = \frac{1}{4} \times [2.2875 + 2.55 + 2.2 + 2.3531] = 2.3477$$

$$X_{3,E} = \frac{1}{4} \times [X_{-\text{CHCH}_3} + X_C + X_H + X_{-\text{CH}_2\text{CH}_3}] = \frac{1}{4} \times [2.3458 + 2.55 + 2.2 + 2.3094] = 2.3513$$

$$X_{4,E} = \frac{1}{5} \times [X_{-\text{CH}=\text{CHCH}_3} + X_C + 2 \times X_H + X_{-\text{CH}_3}] = \frac{1}{5} \times [2.3653 + 2.55 + 2 \times 2.2 + 2.2875] = 2.3206$$

$$X_{5,E} = \frac{1}{5} \times [X_C + 3 \times X_H + X_{-\text{CH}_2\text{CH}=\text{CHCH}_3}] = \frac{1}{5} \times [2.55 + 3 \times 2.2 + 2.3288] = 2.2958$$

X_{INI} of carbon atoms labeled 1, 2, 3, 4 and 5 in Fig. 4

$$X_{1,INI} = |(X_{1,E} - X_C) / X_C| \times 100 = |(2.3035 - 2.55) / 2.55| \times 100 = 9.6667$$

$$X_{2,INI} = |(X_{2,E} - X_C) / X_C| \times 100 = |(2.3477 - 2.55) / 2.55| \times 100 = 7.9333$$

$$X_{3,INI} = |(X_{3,E} - X_C) / X_C| \times 100 = |(2.3513 - 2.55) / 2.55| \times 100 = 7.7922$$

$$X_{4,INI} = |(X_{4,E} - X_C) / X_C| \times 100 = |(2.3206 - 2.55) / 2.55| \times 100 = 8.9961$$

$$X_{5,INI} = |(X_{5,E} - X_C) / X_C| \times 100 = |(2.2958 - 2.55) / 2.55| \times 100 = 9.9686$$

Q of cis-2-pentene

$$Q = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \end{bmatrix} = \begin{bmatrix} (X_{1,INI})^{-0.5} \\ (X_{2,INI})^{-0.5} \\ (X_{3,INI})^{-0.5} \\ (X_{4,INI})^{-0.5} \\ (X_{5,INI})^{-0.5} \end{bmatrix} = \begin{bmatrix} (9.6667)^{-0.5} \\ (7.9333)^{-0.5} \\ (7.7922)^{-0.5} \\ (8.9961)^{-0.5} \\ (9.9686)^{-0.5} \end{bmatrix} = \begin{bmatrix} 0.3216 \\ 0.3550 \\ 0.3582 \\ 0.3334 \\ 0.3167 \end{bmatrix}$$

Some ionicity index matrices Q are given in Table 2.

Table 2. Some ionicity index matrices Q .

No	Compound	Q
1	cis-2-butene	$[0.3214 \ 0.3534 \ 0.3534 \ 0.3214]^T$
2	trans-2-butene	$[0.3214 \ 0.3534 \ 0.3534 \ 0.3214]^T$
3	cis-2-pentene	$[0.3216 \ 0.3550 \ 0.3582 \ 0.3334 \ 0.3167]^T$
4	trans-2-pentene	$[0.3216 \ 0.3550 \ 0.3582 \ 0.3334 \ 0.3167]^T$
5	cis-2-hexene	$[0.3218 \ 0.3554 \ 0.3595 \ 0.3366 \ 0.3282 \ 0.3155]^T$
6	trans-2-hexene	$[0.3218 \ 0.3554 \ 0.3595 \ 0.3366 \ 0.3282 \ 0.3155]^T$
7	cis-3-hexene	$[0.3154 \ 0.3316 \ 0.3599 \ 0.3599 \ 0.3316 \ 0.3154]^T$
8	trans-3-hexene	$[0.3154 \ 0.3316 \ 0.3599 \ 0.3599 \ 0.3316 \ 0.3154]^T$
9	cis-4-m-2-pentene	$[0.3220 \ 0.3566 \ 0.3633 \ 0.3469 \ 0.3194 \ 0.3194]^T$
10	trans-4-m-2-pentene	$[0.3220 \ 0.3566 \ 0.3633 \ 0.3469 \ 0.3194 \ 0.3194]^T$
11	cis-2-heptene	$[0.3191 \ 0.3549 \ 0.3598 \ 0.3356 \ 0.3305 \ 0.3265 \ 0.3145]^T$
12	trans-2-heptene	$[0.3191 \ 0.3549 \ 0.3598 \ 0.3356 \ 0.3305 \ 0.3265 \ 0.3145]^T$
13	cis-3-heptene	$[0.3154 \ 0.3317 \ 0.3602 \ 0.3612 \ 0.3350 \ 0.3274 \ 0.3146]^T$
14	trans-3-heptene	$[0.3154 \ 0.3317 \ 0.3602 \ 0.3612 \ 0.3350 \ 0.3274 \ 0.3146]^T$
15	cis-4-m-2-hexene	$[0.3192 \ 0.3561 \ 0.3645 \ 0.3493 \ 0.3299 \ 0.3151 \ 0.3182]^T$
16	trans-4-m-2-hexene	$[0.3192 \ 0.3561 \ 0.3645 \ 0.3493 \ 0.3299 \ 0.3151 \ 0.3182]^T$
17	cis-5-m-2-hexene	$[0.3191 \ 0.3552 \ 0.3607 \ 0.3382 \ 0.3408 \ 0.3169 \ 0.3169]^T$
18	trans-5-m-2-hexene	$[0.3191 \ 0.3552 \ 0.3607 \ 0.3382 \ 0.3408 \ 0.3169 \ 0.3169]^T$
19	cis-2-m-3-hexene	$[0.3177 \ 0.3458 \ 0.3655 \ 0.3616 \ 0.3319 \ 0.3154 \ 0.3177]^T$
20	trans-2-m-3-hexene	$[0.3177 \ 0.3458 \ 0.3655 \ 0.3616 \ 0.3319 \ 0.3154 \ 0.3177]^T$

T: transposed matrix; m: methyl.

2.4 Some ED_m' values

We calculated the ED_1' , ED_2' and ED_3' values of 44 cis-trans isomers for alkenes by Eq. (5)-(11), and listed them on Table 3.

Table 3. Topological index values of ED_1' , ED_2' , ED_3' for some cis-trans isomers.

No.	Compound	ED_1'	ED_2'	ED_3'	No.	Compound	ED_1'	ED_2'	ED_3'
1	cis-2-butene	1.5469	0.6650	0.3437	23	cis-3-m-3-hexene	3.0229	1.8869	1.2703
2	trans-2-butene	1.5469	0.6650	0.2522	24	trans-3-m-3-hexene	3.0229	1.8869	1.2831
3	cis-2-pentene	2.0173	0.9395	0.5190	25	cis-4,4-2m-2-pentene	2.9804	2.7651	0.9409
4	trans-2-pentene	2.0173	0.9395	0.4392	26	trans-4,4-2m-2-pentene	2.9804	2.7651	0.8666
5	cis-2-hexene	2.4856	1.2238	0.7019	27	cis-3,4-2m-2-pentene	3.0363	2.3030	1.6924
6	trans-2-hexene	2.4856	1.2238	0.6216	28	trans-3,4-2m-2-pentene	3.0363	2.3030	1.7084
7	cis-3-hexene	2.4787	1.2118	0.6924	29	cis-2-octene	3.4139	1.7904	1.0780
8	trans-3-hexene	2.4787	1.2118	0.6256	30	trans-2-octene	3.4139	1.7904	0.9981
9	cis-4-m-2-pentene	2.4930	1.6481	0.7258	31	cis-3-octene	3.4142	1.7797	1.0638
10	trans-4-m-2-pentene	2.4930	1.6481	0.6499	32	trans-3-octene	3.4142	1.7797	0.9965
11	cis-2-heptene	2.9462	1.5051	0.8879	33	cis-4-octene	3.4149	1.7778	1.0585
12	trans-2-heptene	2.9462	1.5051	0.8080	34	trans-4-octene	3.4149	1.7778	0.991
13	cis-3-heptene	2.9467	1.4948	0.8754	35	cis-5,5-2m-2-hexene	3.4387	3.0582	1.1273
14	trans-3-heptene	2.9467	1.4948	0.8082	36	trans-5,5-2m-2-hexene	3.4387	3.0582	1.0465
15	cis-4-m-2-hexene	2.9540	1.8611	1.1661	37	cis-2,2-2m-3-hexene	3.4493	3.0418	1.1196
16	trans-4-m-2-hexene	2.9540	1.8611	1.0901	38	trans-2,2-2m-3-hexene	3.4493	3.0418	1.0595
17	cis-5-m-2-hexene	2.9486	1.9322	0.9083	39	cis-2,4-2m-3-hexene	3.4913	2.5941	1.4600
18	trans-5-m-2-hexene	2.9486	1.9322	0.8279	40	trans-2,4-2m-3-hexene	3.4913	2.5941	1.4735
19	cis-2-m-3-hexene	2.9541	1.9196	0.9005	41	cis-4-m-3-e-2-pentene	3.4889	2.5242	2.0710
20	trans-2-m-3-hexene	2.9541	1.9196	0.8384	42	trans-4-m-3-e-2-pentene	3.4889	2.5242	2.0749
21	cis-3-m-2-hexene	3.0206	1.8936	1.2585	43	cis-5-decene	4.3501	2.3477	1.4353
22	trans-3-m-2-hexene	3.0206	1.8936	1.2699	44	trans-5-decene	4.3501	2.3477	1.3677

m: methyl.

The descriptor ED_m' possesses all the traits that the ED_m has: (1) The ED_m' indicates the size of the molecule, and the more atoms correspond to the bigger values of ED_m' ; (2) The index ED_m' reveals the molecular branching character. The bigger degree of branching for the isomers, the bigger $\sum ED_m'$; (3) The ED_m' shows better structural selectivity than ED_m in cis-trans isomers distinguishing, because the ED_m can not distinguish from cis-trans isomers.

3. RESULTS AND DISCUSSION

3.1 Model Development

In this paper, we considered the interactions of α -, β -, γ -carbon atoms, and used ED_1' , ED_2' and ED_3' to build a multiple linear regression (MLR) model:

$$P = a ED_1' + b ED_2' + c ED_3' + d \quad (12)$$

Where a, b, c and d are constants, P represents the molecular properties. The correlation coefficient (R), the Fischer ratio value (F), and the standard deviation (S) can be used to assess the quality of the models.

3.2 Relationships between the indices ED_m' and properties for alkenes

We regress the boiling point ($b.p.$)^{36,47}, the standard enthalpy of formation ($D_f H_m^0$)⁴⁶, the molar refraction (R_m)⁴⁸ and the molar volume (V_m)⁴⁶ of 44 cis-trans isomers for alkenes with their values of ED_m' , according to Eq. (12), and listed the statistical results of regression in Table 4.

Table 4. Results of regression for alkenes

Property(P) ^{36,46-48}	$P = a ED_1' + b ED_2' + c ED_3' + d$				R	S	F	N
	a	b	c	d				
$b.p.$	70.5347	-15.4246	-2.1993	-89.4146	0.9954	3.6146	1,223	38
$D_f H_m^0$	-39.9037	-5.7244	-5.1517	57.9242	0.9840	4.4194	265	30
R_m	19.4521	-14.6942	-1.4681	0.2419	1.0000	0.0421	72,112	16
V_m	38.2634	1.6749	-11.6817	34.6648	0.9995	0.7469	4,017	16

All the values of R in Table 4 are larger than 0.984, and the adjusted R square (R^2) are above 96.12%. The average absolute relative errors between calculated values and experimental ones for $b.p.$, $D_f H_m^0$, R_m and V_m are 1.84%, 2.55%, 1.39% and 0.57%, respectively. One can observe that the calculations agree well with the experimental values (Figure 5), and most factors describing the property change of cis-trans isomers are opened out. The calculated values and experimental ones for $b.p.$, $D_f H_m^0$, R_m and V_m are listed in Table 5 and Table 6.

Figure 5. The experimental values vs. the calculated values by MLR model.

Table 5. The experimental values, the calculated values and the predicted values of *b.p.* and $D_f H_m^\theta$ for some alkenes.

No.	Compound	<i>b.p.</i> / (°C)			$D_f H_m^\theta$ / (kJ·mol ⁻¹)		
		Exp. ^a	Calc. ^b	Calc. ^c	Exp. ^d	Calc. ^b	Calc. ^c
1	cis-2-butene	3.7	8.7	6.5	-6.99	-9.38	-9.96
2	trans-2-butene	0.8	8.9	6.9	-11.17	-8.91	-9.36
3	cis-2-pentene	36.9	37.2	35.6*	-28.07	-30.63	-31.14*
4	trans-2-pentene	36.3	37.4	36.0	-31.76	-30.21	-30.61
5	cis-2-hexene	68.8	65.5	64.4*	-52.34	-51.88	-52.31
6	trans-2-hexene	67.9	65.7	64.8	-53.89	-51.47	-51.78*
7	cis-3-hexene	66.4	65.2	64.2	-47.61	-51.49	-51.93
8	trans-3-hexene	67.1	65.4	64.5*	-54.43	-51.14	-51.49*
9	cis-4-m-2-pentene	56.3	59.4	58.2	-50.33	-54.73	-54.11
10	trans-4-m-2-pentene	58.6	59.6	58.5	-54.35	-54.34	-53.61
11	cis-2-heptene	98.4	93.2	92.7*	-69.04	-72.83	-73.18*
12	trans-2-heptene	98	93.4	93.1	-73.22	-72.42	-72.65
13	cis-3-heptene	95.8	93.4	93.0*	-69.04	-72.73	-73.08*
14	trans-3-heptene	95.7	93.6	93.3	-73.22	-72.38	-72.64
15	cis-4-m-2-hexene	86.3	87.7	86.4	-73.22	-76.61	-76.46
16	trans-4-m-2-hexene	87.6	87.8	86.7*	-77.4	-76.22	-75.96*
17	cis-5-m-2-hexene	89.5	86.8	86.1	-76.15	-75.48	-74.76
18	trans-5-m-2-hexene	88.1	86.9	86.4*	-80.33	-75.06	-74.23*
19	cis-2-m-3-hexene	86	87.4	86.7	-76.15	-75.58	-74.90
20	trans-2-m-3-hexene	85.9	87.5	87.0	-80.33	-75.26	-74.48
21	cis-3-m-2-hexene	-	90.7	-	-80.75	-79.93	-79.89*
22	trans-3-m-2-hexene	-	90.7	-	-80.75	-79.99	-79.97
23	cis-3-m-3-hexene	-	91.9	-	-80.42	-80.05	-80.04
24	trans-3-m-3-hexene	-	91.9	-	-77.82	-80.11	-80.13*
25	cis-4,4-2m-2-pentene	80.4	76.1	75.1*	-73.64	-81.68	-78.91
26	trans-4,4-2m-2-pentene	76.7	76.3	75.4	-89.79	-81.3	-78.41
27	cis-3,4-2m-2-pentene	-	86.4	-	-87.45	-85.14	-84.70*
28	trans-3,4-2m-2-pentene	-	86.4	-	-87.45	-85.22	-84.80
29	cis-2-octene	125.6	121.4	121.5	-	-94.11	-
30	trans-2-octene	125	121.6	121.8*	-	-93.69	-
31	cis-3-octene	122.9	121.6	121.7	-	-93.98	-
32	trans-3-octene	123.3	121.8	122.0*	-	-93.64	-
33	cis-4-octene	122.5	121.7	121.8	-	-93.97	-
34	trans-4-octene	122.3	121.9	122.1	-	-93.63	-
35	cis-5,5-2m-2-hexene	106.9	103.5	103.0*	-	-102.61	-
36	trans-5,5-2m-2-hexene	104.1	103.7	103.4	-	-102.19	-
37	cis-2,2-2m-3-hexene	105.5	104.5	104.1	-	-102.9	-
38	trans-2,2-2m-3-hexene	100.8	104.6	104.4*	-	-102.59	-
39	cis-2,4-2m-3-hexene	109	113.6	112.6	-	-103.76	-
40	trans-2,4-2m-3-hexene	107.6	113.6	112.6	-	-103.83	-
41	cis-4-m-3-e-2-pentene	116	113.2	110.7*	-	-106.41	-
42	trans-4-m-3-e-2-pentene	114.3	113.2	110.7	-	-106.43	-
43	cis-5-decene	171	178	179.3	-91.09	-102.9	-100.05
44	trans-5-decene	171	178.2	179.6*	-109.45	-102.59	-99.65*

Note: a: from ref. 36 and ref. 47; d: from ref. 46; b: calculated values according to Eq. (12); c: calculated values by RSP method; *: predicted values by RSP method; m = methyl.

Table 6. The experimental values, the calculated values and the predicted values of R_m and V_m for some alkenes.

No.	$R_m / (\text{cm}^3 \cdot \text{mol}^{-1})$			No.	$V_m / (\text{cm}^3 \cdot \text{mol}^{-1})$		
	Exp. ^a	Calc. ^b	Calc. ^c		Exp. ^d	Calc. ^b	Calc. ^c
3	24.9563	24.9154	24.9054	3	106.9707	107.3644	106.8320
4	25.0248	25.0326	25.0288	9	125.7937	124.3375	125.7676
5	29.5522	29.5787	29.5711*	10	125.869	125.2241	127.2047*
6	29.6371	29.6966	29.6954	11	138.6751	139.5453	139.2653
7	29.6715	29.6348	29.6271*	12	140.02	140.4787	140.7782*
8	29.7518	29.7329	29.7304	13	139.6615	139.6932	139.4558
11	34.0614	34.1318	34.1263	14	140.6417	140.4782	140.7282
12	34.2797	34.2491	34.2499*	26	142.52	143.2132	149.0017*
13	34.3036	34.3112	34.3059	29	154.9192	155.6983	155.5162
14	34.4178	34.4099	34.4099*	30	155.8661	156.6317	157.0291*
29	38.7959	38.7582	38.7546	31	156.737	155.8578	155.7286
30	38.8928	38.8755	38.8782	32	156.8904	156.6439	157.0029
31	38.9625	38.9421	38.9388*	33	155.5851	155.9433	155.8438*
32	39.0856	39.0409	39.0429	34	157.1321	156.7318	157.1220
33	38.9471	38.9914	38.9884	43	188.3949	188.2801	188.3979
34	39.0515	39.0905	39.0928*	44	189.5149	189.0698	189.6779*

Note: a: from ref. 48; d: from ref. 46; b: calculated values according to Eq. (13)-(16);
c: calculated values by RSP method; *: predicted values by RSP method.

3.3 Model validation

3.3.1 LOO method

In this method⁴⁰, we take one sample from N samples, do the regression with the experimental values of the remaining $N-1$ samples, and obtain a regression equation for $N-1$ samples, then predict the value of the sample taken-out by the regression equation. Using the same method, we get N prediction values. Finally, N prediction values are regressed with their experimental values, and the regression results, such as the correlation coefficient R_{CV} , the Fischer ratio value F_{CV} , the standard error S_{CV} , are used to test the stability and validity of the model. The results by LOO method are shown in Table 7.

Table 7. The results of LOO method.

Property(P) ^{36,46-48}	Equation of regression	R_{CV}	S_{CV}	F_{CV}	N
$b.p.$	$P_{\text{exp.}} = 0.9970P_{\text{pre.}} + 0.2102$	0.9937	4.1235	2,809	38
$D_f H_m^0$	$P_{\text{exp.}} = 0.9909P_{\text{pre.}} - 0.6386$	0.9753	5.2833	547	30
R_m	$P_{\text{exp.}} = 0.9999P_{\text{pre.}} + 0.0042$	1.0000	0.0520	142,074	16
V_m	$P_{\text{exp.}} = 1.0011P_{\text{pre.}} - 0.3816$	0.9977	1.4796	3,060	16

Figure 6. Comparison between the correlation coefficients R in Table 4 and the correlation coefficients R_{CV} by LOO method.

The values of R_{CV} by LOO method are close to the R in Table 4 (See Figure 6). The multiple correlation coefficients of models about $b.p.$, $D_f H_m^0$, R_m and V_m by LOO method are among 0.9949-0.9958, 0.9796-0.9893, 0.99997-0.99998 and 0.9993-0.9997, respectively. Most of the multiple correlation coefficients fluctuate in a narrow range (See Fig. 7), which demonstrated the high stability and good reliability of the index.

In this method³¹, the M samples are randomly taken from all the N samples as the training set, and the $N-M$ remained samples as the test set. First, build a model with the training set, then predict the $N-M$ values of the test set using the model of the test set, and finally regress $N-M$ prediction values with their experimental ones. The correlation coefficient (R_p), the Fischer ratio value (F_p), and the standard error (S_p) are used to test the predictive ability of the model. The results are shown in Table 8.

Figure 7. The distribution plot of multi-correlation coefficient R in LOO method versus frequency of R for four properties of alkenes.

3.3.2 RSP method

Table 8. The results of the RSP method.

Property(P) ^{36,46-48}	Equation of regression	R_p	S_p	F_p	$N-M$
$b.p.$	$P_{\text{exp.}} = 0.9343P_{\text{pre.}} + 8.1448$	0.9951	3.2618	1,226	14
$D_f H_m^\theta$	$P_{\text{exp.}} = 1.0972P_{\text{pre.}} + 5.7167$	0.9821	4.1824	244	11
R_m	$P_{\text{exp.}} = 0.9977P_{\text{pre.}} + 0.0878$	1.0000	0.0343	74,900	6
V_m	$P_{\text{exp.}} = 1.0277P_{\text{pre.}} - 5.9317$	0.9943	2.5947	345	6

All the values of R for $b.p.$, $D_f H_m^\theta$, R_m , and V_m in Table 8 are larger than 0.98. Moreover, the plots (Fig. 8) of the predicted values and the experimental ones for the test set and the training set can be an evidence of the good quality for the models. All the experimental values, the calculated values and the predicted values mentioned in this paper are listed in Table 5 and Table 6.

Figure 8. The experimental values vs. the predicted values based on the RSP method.

3.4 The contributions of the different ED_m'

We studied the different effects that different ED_m' have on calculated properties for compounds. The relative and fraction contribution of each index were estimated. The relative contribution (Ψ_r) and fraction contribution (Ψ_f) of the relevant indexes are defined as follows⁴⁹:

$$\Psi_r^i = a_i \bar{T}_i \quad (13)$$

$$\Psi_f^i = R^2 \left| \Psi_r^i \right| / \sum_i \left| \Psi_r^i \right| \times 100\% \quad (14)$$

where Ψ_r , Ψ_f and $\bar{T}_i$ are, respectively, the relative contribution, fraction contribution, and the average value of the i th topological index. Square of the correlation coefficient, R^2 , is the coefficient of the determination. The sum is over all the indexes in the model. Results of the contribution analysis are summarized in Table 9.

Table 9. Ψ_r and Ψ_f of QSPR models for the compounds.

Property ^{36,46-48}	$\Psi_r(\Psi_f)$		
	ED_1'	ED_2'	ED_3'
$b.p.$	212.3941 (86.31 %)	-29.2905 (11.90 %)	-2.1363 (0.87 %)
$D_f H_m^0$	-109.8166 (85.41%)	-10.0334 (7.80%)	-4.6626 (3.63%)
R_m	56.2104 (71.19%)	-21.5324 (27.27%)	-1.2080 (1.53%)
V_m	121.8588 (89.59%)	2.9722 (2.19%)	-11.0463 (8.12%)

From Table 9, one can observe that the contributions of each ED_m' to the properties cover a wide range of values. ED_1' plays the most important role in these models. On the contrary, the contributions of ED_2' and ED_3' are less significant. From Fig. 9, it is clear that Ψ_f values of ED_1' for alkenes range between 70% and 90%, which reflects the interaction of α -carbons to be the strongest. Meantime, we can see that the Ψ_f values of ED_2' and ED_3' are in the range of 2%~30% and 1%~9%, respectively.

Figure 9. Comparison of the fraction contribution Ψ_f of different ED_m' .

4 CONCLUSIONS

In this paper, we introduce the bond angle into hidden hydrogen graph of molecules, use the linear distance instead of the sum of bond length between the two atoms, and proposed the new topological descriptor ED_m' . High-quality models of the cis-trans isomers for QSPR study were put forward on the ground of the index ED_m' . By these models, properties of $b.p.$, $D_f H_m^0$, R_m and V_m were calculated and predicted, and the results agreed well with the experimental values. Moreover, the LOO method and the RSP method demonstrate that the models have statistical significance and good stability.

ACKNOWLEDGEMENTS

The project was supported by Science and Technology Projects Foundation of Hunan province (No. 06FJ4104).

REFERENCES

- (1) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, 69, 17-20.
- (2) Balaban, A. T. Highly Discriminating Distance-Based Topological Index. *Chem. Phys. Lett.* **1982**, 89, 399-404.
- (3) Balaban, A. T.; Mills, D.; Basak, S. C. Correlation between Structure and Normal Boiling Points of Acyclic Carbonyl Compounds. *J. Chem. Inf. Comput. Sci.* **1999**, 39, 758-764.
- (4) Balaban, A. T.; Rücker, C. Using Protochirons for Three-Dimensional Coding of Certain Chemical Structures. *J. Chem. Inf. Comput. Sci.* **2001**, 41, 1145-1149.
- (5) Balaban, A. T. Mathematical Chemistry: (3, g)-Cages with Girth g, Topological Indices, and Other Graph-Theoretical Problems. *Fund. Inform.* **2005**, 64, 1-16.

- (6) Randić, M. On characterization of molecular branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609-6615.
- (7) Randić, M.; Zupan, J. On Interpretation of Well-Known Topological Indices. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 550-560.
- (8) Randić, M.; Balaban, A. T.; Basak, S. C. On Structural Interpretation of Several Distance Related Topological Indices. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 593-601.
- (9) Randić, M. Wiener-Hosoya Index - A Novel Graph Theoretical Molecular Descriptor. *J. Chem. Inf. Model.* **2004**, *44*, 373-377.
- (10) Randić, M.; Zupan, J.; Vikić-Topić, D.; Plavšić, D. A Novel Unexpected Use of a Graphical Representation of DNA: Graphical Alignment of DNA Sequences. *Chem. Phys. Lett.* **2006**, *431*, 375-379.
- (11) (a) Gutman, I.; Trinajstić, N. Graph Theory and Molecular Orbitals. *Topics Curr. Chem.* **1973**, *42*, 49-93; (b) Gutman, I.; Ruscic, B.; Trinajstić, N.; Wilcox, Jr. C. F. Graph Theory and Molecular Orbitals. XII. Acyclic Polyenes. *J. Chem. Phys.* **1975**, *62*, 3399-3405. (c) Gutman, I.; Milun, M.; Trinajstić, N. Graph Theory and Molecular Orbitals. XIX. Nonparametric Resonance Energies of Arbitrary Conjugated Systems. *J. Amer. Chem. Soc.* **1977**, *99*, 1692-1704.
- (12) Gutman, I.; Bosanac, S. Topological studies on heteroconjugated molecules. The stability of alternant systems with one heteroatom. *Chem. Phys. Lett.* **1976**, *43*, 371-373.
- (13) Gutman, I.; Rouvray, D. H. New theorem for the wiener molecular branching index of trees with perfect matchings. *Comput. Chem.* **1990**, *14*, 29-32.
- (14) Gutman, I. Selected properties of the Schultz molecular topological index. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1037-1039.
- (15) Trinajstić, N.; Nikolic, S.; Lucic, B.; Amic, D.; Mihalic, Z. The Detour Matrix in Chemistry. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 631-638.
- (16) Trinajstić, N. Chemical Nomenclatures and the Computer. *Comput. Chem.* **1994**, *18*, 435-436.
- (17) Trinajstić, N.; Babic, D.; Nikolic, S.; Plavšić, D.; Amic, D.; Mihalic, Z. The Laplacian matrix in chemistry. *J. Chem. Inf. Comput. Sci.* **1994**, *2*, 368-376.
- (18) Trinajstić, N. *Chemical Graph Theory*; CRC Press: Boca Raton, Florida, 1992.
- (19) Estrada E. Edge Adjacency Relationships and a Novel Topological Index Related to Molecular Volume. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 31-33.
- (20) Estrada, E.; Vilar, S.; Uriarte, E.; Gutierrez, Y. In Silico Studies toward the Discovery of New Anti-HIV Nucleoside Compounds with the Use of TOPS-MODE and 2D/3D Connectivity Indices, I. Pyrimidyl Derivatives. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1194-1203.
- (21) Estrada, E. Application of a novel graph-theoretic folding degree index to the study of steroid-DB3 antibody binding affinity. *Comput. Biol. Chem.* **2003**, *27*, 305-313.
- (22) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure-Activity Analysis*; RSP-Wiley: Chichester, UK, 1986.
- (23) Devillers, J.; Balaban, A. T. Eds. *Topological Indices and Related Descriptors in QSAR and Drug Design*; Gordon&Breach: Amsterdam, The Netherlands, 2000.
- (24) Xu, L.; Hu, C. Y. *The Applications of Graph Theory in Chemistry*; Science Press: Beijing, 2000. (in Chinese)
- (25) Willighagen, E. L.; Denissen, H. M. G. W.; Wehrens, R.; Buydens, L. M. C. On the Use of ¹H and ¹³C 1D NMR Spectra as QSPR Descriptors. *J. Chem. Inf. Model.* **2006**, *46*, 487-494.
- (26) Saçan, M. T.; Erdem, S. S.; Özpınar, G. A.; Balcioglu, I. A. QSPR Study on the Bioconcentration Factors of Nonionic Organic Compounds in Fish by Characteristic Root Index and Semiempirical Molecular Descriptors. *J. Chem. Inf. Model.* **2004**, *44*, 985-992.
- (27) Modarresi, H.; Dearden, J. C.; Modarress, H. QSPR Correlation of Melting Point for Drug Compounds Based on Different Sources of Molecular Descriptors. *J. Chem. Inf. Model.* **2006**, *46*, 930-936.
- (28) Kahn, I.; Fara, D.; Karelson, M.; Maran, U.; Andersson, P. L. QSPR Treatment of the Soil Sorption Coefficients of Organic Pollutants. *J. Chem. Inf. Model.* **2005**, *45*, 94-105.
- (29) Lather, V.; Madan, A. K. Topological models for the prediction of HIV-protease inhibitory activity of tetrahydropyrimidin-2-ones. *J. Mol. Graph. Model.* **2005**, *23*, 339-345.
- (30) Narasimhan, B.; Kumari, M.; Jain, N.; Dhake, A.; Sundaravelan, C. Correlation of antibacterial activity of some N-[5-(2-furanyl)-2-methyl-4-oxo-4H-thieno[2,3-d]pyrimidin-3-yl]-carboxamide and 3-substituted-5-(2-furanyl)-2-methyl-3H-thieno[2,3-d]pyrimidin-4-ones with topological indices using Hansch analysis. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4951-4958.
- (31) Nie, C. M.; Wu, Y. X.; Wu, R. Y.; Jiang, S. H.; Zhou, C. Y. Applications Of A New Topological Index ED_m In Some Aliphatic Hydrocarbons. *J. Theor. Comput. Chem.* **2009**, *8*, 19-45.
- (32) Zhou, C. Y.; Nie, C. M. Molecular descriptors of topology and a study on quantitative structure and property relationships. *Bull. Chem. Soc. Jpn.* **2007**, *80*, 1504-1510.
- (33) Zhou, C. Y.; Nie, C. M.; Li, S.; Li, Z. H. A novel semi-empirical topological descriptor Nr and the application to study on QSPR/QSAR. *J. Comput. Chem.* **2007**, *28*, 2413-2423.
- (34) Crombie, K. Cis-trans isomerism was formerly called gemetricar isomerism. *Q. Rev. Chem. Soc.* **1952**, *6*, 101-140.
- (35) Cao, C. Z.; Yuan, H. Study on the Cis/trans-isomerism of Alkenes by Topological Approach. *Acta Phys-Chim. Sin.* **2005**, *21*, 360-366. (in chinese)
- (36) Yuan, H.; Cao, C. Z. Topological indices based on vertex, edge, ring, and distance: Application to various physicochemical properties of diverse hydrocarbons. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 501-512.
- (37) Randic, M. Graph Theoretical Descriptors of Two-Dimensional Chirality with Possible Extension to Three-Dimensional Chirality. *Chem. Inf. Comput. Sci.* **2001**, *41*, 639-649.
- (38) Golbraikh, A.; Bonchev, D.; Tropsha, A. Novel chirality descriptors derived from molecular topology. *Chem. Inf. Comput. Sci.* **2001**, *41*, 147-158.
- (39) Golbraikh, A.; Bonchev, D.; Tropsha, A. Novel ZE-isomerism descriptors derived from molecular topology and their application to QSAR analysis. *Chem. Inf. Comput. Sci.* **2002**, *42*, 769-787.
- (40) Nie, C. M.; Dai, Y. M.; Wen, S. N.; Li, Z. H.; Zhou, C. Y.; Peng, G. W. Topological homologous regularity for additive property of alkanes. *Acta Chim. Sinica.* **2005**, *63*, 1449-1455. (in chinese)
- (41) Zhou, C. Y.; Nie, C. M.; Li, S.; Wen, S. N.; Peng, G. W.; Li, Z. H. Chemical behavior of topology and its application to QSPR of lanthanide and actinide. *Chin. J. Inorg. Chem.* **2007**, *23*, 25-33. (in chinese)
- (42) Nie, C. M.; Peng, G. W.; Xiao, F. Z.; Li, S.; He, X. M.; Li, Z. H.; Zhou, C. Y. Study on Topological Chemistry of Gas Chromatography Retention Index for Sulfides. *Chin. J. Anal. Chem.* **2006**, *34*, 1560-1564. (in chinese)
- (43) Nie, C. M.; Li, Z. H.; Wen, S. N. A Study for the Relationship between ¹³C NMR Chemical Shifts of Alkanes and the Atomic Ionicity index, Polarizability Effect Index. *Chin. J. Org. Chem.* **2002**, *22*, 46-51. (in chinese)
- (44) Nie, C. M.; Wen, S. N. Equilibrium Electronegativity and ¹³C NMR Chemical Shifts of Alkanes. *Chin. J. Magn. Reson.* **2001**, *18*, 45-50. (in chinese)
- (45) Nie, C. M. Group Electronegativity. *J. Wuhan Univ. (Nat. Sci. Ed.)* **2000**, *46*, 176-180. (in Chinese).
- (46) Yao, Y. B.; Xie, T.; Gao, Y. M. *Handbook of Physical Chemistry*; Shanghai Science Technology Press: Shanghai, 1985. (in Chinese)
- (47) David, R.L. *CRC Handbook of chemistry and physics*; 81st ed.; CRC Press: Florida, 2001.
- (48) Schultz, H.P.; Schultz, E. B.; Schultz, T. P. Topological Organic Chemistry, 9. Graph Theory and Molecular Topological Indices of Stereoisomeric Organic Compounds. *Chem. Inf. Comput. Sci.* **1995**, *35*, 864-870.
- (49) Needham, D. E.; Wei, I. C.; Seybold, P. G. Molecular modeling of the physical properties of alkanes. *J. Am. Chem. Soc.* **1988**, *110*, 4186-4194.