Correlation of the Eccentricity Zagreb Index with Physical Properties of Alkanes

Abstract:

Topological indices are mathematical tools used to establish quantitative correlations between chemical structure and physical properties or chemical reactivity. In this study, the relationship between the molecular structures of alkanes and the eccentricity of the first Zagreb index is investigated. The results indicate that this index serves as a useful descriptor for predicting the physical properties of lower alkanes.

Keywords:Topological indices; Zagreb index; Eccentricity; Alkanes; Physical properties; Structure–property relationship

1 Introduction

In the realm of theoretical and mathematical chemistry, topological indices *serve as powerful numerical invariants derived from the* molecular graph $G = (V(G), E(G))$ *[17], where vertices* $V(G)$ *represent carbon atoms and edges* $E(G)$ *represent covalent C-C bonds. These graph-theoretic parameters provide a formalized mathematical framework to encode the connectivity and topological complexity of molecular structures into quantitative descriptors that correlate with experimentally measurable physicochemical properties, chemical reactivity, and biological activities.*

Formally, a molecular graph G of an alkane is a simple, connected, undirected graph where each vertex $v \in V(G)$ has degree $d(v)$ representing the number of carbon atoms directly bonded to vertex v, and the distance $d(u, v)$ between vertices $u, v \in V(G)$ is the length of the shortest path connecting them. The eccentricity $\varepsilon(v)$ *of a vertex v is defined as $\varepsilon(v) = \max_{u \in V(G)} d(u, v)$, which quantifies the vertex's topological "remoteness" from the farthest vertex in the graph. Topological indices aggregate such local and global structural features into scalar values that serve as molecular fingerprints in* quantitative structure–property relationship (QSPR) *and* quantitative structure–activity relationship (QSAR) *studies [11].*

Among the diverse families of topological indices, degree-based indices *have garnered significant attention due to their computational simplicity and exceptional predictive power. The seminal* Zagreb indices, *introduced by Gutman and Trinajstić in 1972 [7], are defined as:*

$$M_1(G) = \sum_{v \in V(G)} d(v)^2$$

where these indices capture the first and second moments of the vertex degree distribution, respectively. The first Zagreb index $M_1(G)$ has proven particularly effective in modeling physicochemical properties of organic compounds due to its sensitivity to both local branching patterns and overall molecular size.

Recent advancements have focused on hybrid indices that incorporate additional graph-theoretic concepts such as eccentricity [5]. The eccentricity-based first Zagreb index*, defined as*

$$M_1^{\varepsilon}(G) = \sum_{v \in V(G)} d(v)^2 \varepsilon(v),$$

represents a natural extension that simultaneously encodes local connectivity *(via $d(v)^2$) and* global shape information *(via $\varepsilon(v)$). This dual encoding enables superior discrimination among structural isomers and enhances correlation with size- and shape-dependent properties.*

Lower alkanes *(C_nH_{2n+2}) [2, 3] constitute an ideal model system for QSPR investigations due to their structural simplicity, systematic structural diversity through branching, and comprehensive experimental databases. The molecular graphs of alkanes are* tree graphs *(acyclic connected graphs), where topological properties directly reflect molecular geometry and branching patterns that govern intermolecular forces and phase behavior.*

This study systematically investigates the efficacy of the eccentricity-based first Zagreb index $M_1^{\varepsilon}(G)$ as a molecular descriptor for a set of 73 lower alkanes, each represented by its corresponding molecular graph. Comprehensive QSPR analyses are performed to correlate $M_1^{\varepsilon}(G)$ with seven key physicochemical properties—boiling points, molar volumes, molar refractions, heats of vaporization, critical temperatures, critical pressures, and surface tensions—sourced from validated literature databases.

By bridging chemical graph theory with experimental thermophysical data, this work contributes to the ongoing development of structure-property relationships and provides a valuable tool for molecular design, property prediction, and understanding alkane behavior in chemical engineering applications[1].

2 Materials and Methods

Many scholars have found bounds for many topological indices. Only a few of those indices are listed below.

Let $G = (V(G), E(G))$ be a simple, connected graph representing a molecular structure, where $V(G)$ denotes the set of vertices and $E(G)$ denotes the set of edges. For a vertex $v \in V(G)$, the degree of v is denoted by $d(v)$.

Gutman and Trinajstić defined the Zagreb indices as [7, 8]

$$M_1(G) = \sum_{uv \in E(G)} [d_u + d_v]$$

$$M_2(G) = \sum_{uv \in E(G)} [d_u \cdot d_v]$$

The redefined Zagreb index, denoted by $ReZG(G)$, is defined as

$$ReZG(G) = \sum_{uv \in E(G)} \frac{d(u) + d(v)}{d(u)\, d(v)},$$

where u and v are the end vertices of the edge uv, and $d(u)$ and $d(v)$ represent their respective degrees.

***Eccentricity of a Graph**:*
For a vertex $v \in V(G)$, the eccentricity of v, denoted by $\varepsilon(v)$, is defined as the maximum distance from v to any other vertex in G, that is,

$$\varepsilon(v) = \max\{d(v, u) \mid u \in V(G)\},$$

where $d(v, u)$ denotes the length of the shortest path between vertices v and u.

***Eccentricity of first Zagrab index**:*
In chemical graph theory, vertex eccentricity plays a crucial role in capturing the global structural features of a molecular graph. The eccentricity of a vertex reflects its maximum distance from any other vertex in the graph, thereby providing insight into the overall topology and spatial extension of the molecule. By combining vertex eccentricity with the square of vertex degree, the eccentricity-based first Zagreb index integrates both local connectivity and global distance information, making it a meaningful topological descriptor for quantitative structure–property relationship (QSPR) studies.

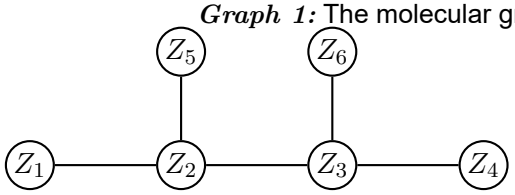
***Definition 2.1:** . The eccentricity-based first Zagreb index of the graph G, denoted by $M_1^{\varepsilon}(G)$, is defined as*

$$M_1^{\varepsilon}(G) = \sum_{v \in V(G)} d(v)^2\, \varepsilon(v).$$

where $V(G)$ denotes the vertex set of the graph G, $d(v)$ represents the degree of a vertex $v \in V(G)$, and $\varepsilon(v)$ denotes the eccentricity of v.

For example, Let G be the molecular graph of $2,3$-dimethylbutane, which has six vertices.

Graph 1: The molecular graph of 2, 3-dimethylbutane



Vertex	$d(v)$	$\varepsilon(v)$	$d(v)^2\,\varepsilon(v)$
Z_1	1	3	$1^2 \times 3 = 3$
Z_2	3	2	$3^2 \times 2 = 18$
Z_3	3	2	$3^2 \times 2 = 18$
Z_4	1	3	$1^2 \times 3 = 3$
Z_5	1	3	$1^2 \times 3 = 3$
Z_6	1	3	$1^2 \times 3 = 3$

Using the definition

$$M_1^{\varepsilon}(G) = \sum_{v \in V(G)} d(v)^2\,\varepsilon(v),$$

we calculate step by step:

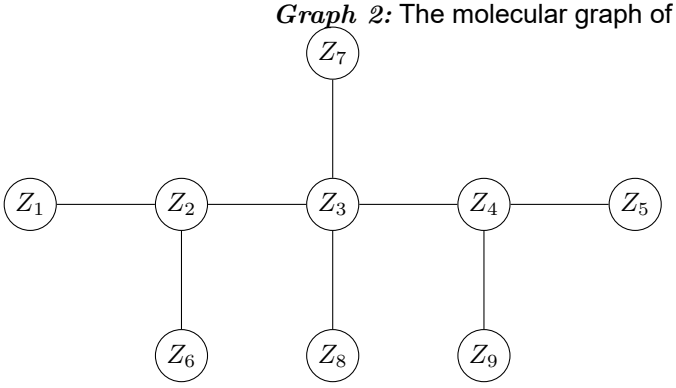
$$\begin{aligned} M_1^{\varepsilon}(G) &= 1^2(3) + 3^2(2) + 3^2(2) + 1^2(3) + 1^2(3) + 1^2(3) \\ &= 3 + 18 + 18 + 3 + 3 + 3 \\ &= 48 \end{aligned}$$

$$\boxed{M_1^{\varepsilon}(G) = 48}$$

and

Let $G = (V(G), E(G))$ be the molecular graph of $2,3,3,4$-tetramethylpentane with $|V(G)| = 9$ carbon vertices representing its C_9H_{20} structure.

Graph 2: The molecular graph of 2, 3, 3, 4-tetramethylpentane



Using the definition

$$M_1^{\varepsilon}(G) = \sum_{v \in V(G)} d(v)^2 \, \varepsilon(v),$$

we compute the degree and eccentricity of each vertex of the graph.

Vertex	$d(v)$	$\varepsilon(v)$	$d(v)^2 \, \varepsilon(v)$
Z_1	1	4	$1^2 \times 4 = 4$
Z_2	3	3	$3^2 \times 3 = 27$
Z_3	4	3	$4^2 \times 3 = 48$
Z_4	3	3	$3^2 \times 3 = 27$
Z_5	1	4	$1^2 \times 4 = 4$
Z_6	1	4	$1^2 \times 4 = 4$
Z_7	1	4	$1^2 \times 4 = 4$
Z_8	1	4	$1^2 \times 4 = 4$
Z_9	1	4	$1^2 \times 4 = 4$

Therefore,

$$\begin{aligned} M_1^{\varepsilon}(G) &= 4 + 27 + 48 + 27 + 4 + 4 + 4 + 4 + 4 \\ &= 126. \end{aligned}$$

$$\boxed{M_1^{\varepsilon}(G) = 126}$$

3 Preliminary Results

Let G be the molecular graph corresponding to an alkane, where vertices represent carbon atoms and edges represent carbon–carbon bonds. The eccentricity-based first Zagreb index of G is defined as [5]

$$M_1^{\varepsilon}(G) = \sum_{v \in V(G)} d(v)^2 \, \varepsilon(v),$$

where $d(v)$ denotes the degree of the vertex v and $\varepsilon(v)$ denotes its eccentricity.

Using this definition, the values of $M_1^{\varepsilon}(G)$ are computed for a variety of alkanes and their structural isomers. The calculated results, along with the corresponding molecular graph representations, are summarized in Table 1.

Table 1: Calculated $M_1^{\varepsilon}(G)$ values for alkanes

Sl. No.	*Name of the Chemical Compound*	$M_1^{\varepsilon}(G)$
01	*Ethane*	*2*
02	*Propane*	*8*
03	*Butane*	*22*
04	*2-Methylpropane(isobutane)*	*15*
05	*Pentane*	*40*
06	*2-Methylbutane*	*35*
07	*2,2-Dimethylpropane*	*24*
08	*Hexane*	*66*
09	*2-Methylpentane*	*39.5*
10	*3-Methylpentane*	*53*

11	*2,2-Dimethylbutane*	*47*
12	*2,3-Dimethylbutane*	*48*
13	*Heptane*	*96*
14	*2-Methylhexane*	*117*
15	*3-Methylhexane*	*123*
16	*3-Ethylhexane*	*97*
17	*2,2-Dimethylpentane*	*93*
18	*2,3-Dimethylpentane*	*82*
19	*2,4-Dimethylpentane*	*56*
20	*3,3-Dimethylpentane*	*70*
21	*2,3,3-Trimethylbutane*	*65*
22	*Octane*	*134*
23	*2-Methylheptane*	*187*
24	*3-Methylheptane*	*193*
25	*4-Methylheptane*	*171*
26	*2,2-Dimethylhexane*	*194*
27	*2,3-Dimethylhexane*	*213*
28	*2,4-Dimethylhexane*	*227*
29	*2,5-Dimethylhexane*	*154*
30	*3,3-Dimethylhexane*	*121*
31	*3,4-Dimethylhexane*	*134*
32	*3-Ethyl-2-methylpentane*	*145*
33	*3-Ethyl-3-methylpentane*	*77*
34	*2,2,3-Dimethylpentane*	*131*
35	*2,2,4-Dimethylpentane*	*161*
36	*2,3,3-Trimethylpentane*	*112*
37	*2,3,4-Trimethylpentane*	*113*
38	*2,2,3,3-Tetramethylbutane*	*82*
39	*Nonane*	*176*
40	*2-Methyloctane*	*255*
41	*3-Methyloctane*	*250*
42	*4-Methyloctane*	*240*
43	*3-Ethylheptane*	*187*
44	*2,2-Dimethylheptane*	*307*
45	*2,3-Dimethylheptane*	*199*
46	*2,4-Dimethylheptane*	*222*
47	*2,5-Dimethylheptane*	*204*
48	*2,6-Dimethylheptane*	*224*
49	*3,3-Dimethylheptane*	*310*
50	*3,4-Dimethylheptane*	*191*
51	*3,5-Dimethylheptane*	*199*
52	*4,4-Dimethylheptane*	*212*
53	*3-Ethyl-2-methylhexane*	*196*
54	*4-Ethyl-2-methylhexane*	*170*
55	*3-Ethyl-3-methylhexane*	*190*
56	*3-Ethyl-4-methylhexane*	*169*

57	*2,2,3-Trimethylhexane*	*203*
58	*2,2,4-Trimethylhexane*	*197*
59	*2,2,5-Trimethylhexane*	*201*
60	*2,3,3-Trimethylhexane*	*166*
61	*2,3,4-Trimethylhexane*	*150*
62	*2,3,5-Trimethylhexane*	*154*
63	*2,4,4-Trimethylhexane*	*188*
64	*3,3,4-Trimethylhexane*	*190*
65	*3,3-Diethylpentane*	*136*
66	*2,3-Diethyl-3-ethylpentane*	*464*
67	*2,2-Diethyl-3-ethylpentane*	*247*
68	*2,2,3,3-Tetramethylpentane*	*108*
69	*2,2,3,4-Tetramethylpentane*	*132*
70	*2,2,4,4-Tetramethylpentane*	*100*
71	*2,3,3,4-Tetramethylpentane*	*132*
72	*4-Ethylpentane*	*160*

4 Correlation Analysis: Eccentricity-based First Zagreb Index and Physicochemical Properties

In this section, we perform a comprehensive correlation analysis to investigate the relationships between the eccentricity-based first Zagreb index*, denoted by* $M_1^{\varepsilon}(G)$*, of molecular graphs of lower alkanes and their experimentally measured physicochemical properties. Correlation-based QSPR studies of this type are well established in chemical graph theory [17, 11].*

Specifically, seven key properties characterizing the thermodynamic and surface behavior of these compounds are examined:

- ***Boiling points*** *(bp,* $^{\circ}\mathrm{C}$*)*
- ***Molar volumes*** *(mv,* cm^3*)*
- ***Molar refractions*** *(mr,* cm^3*)*
- ***Heats of vaporization*** *(hv,* $\mathrm{kJ\,mol}^{-1}$*)*
- ***Critical temperatures*** *(ct,* $^{\circ}\mathrm{C}$*)*
- ***Critical pressures*** *(cp,* atm*)*
- ***Surface tensions*** *(st,* $\mathrm{dyne\,cm}^{-1}$*)*

The eccentricity-based first Zagreb index $M_1^{\varepsilon}(G)$ *for each molecular graph corresponding to the lower alkanes is computed and reported in Table 1. These values are employed as topological predictors in the present correlation study.*

Experimental data for all physicochemical properties were obtained from established literature sources, namely Seybold et al. [15] and Needham et al. [13]. These datasets are

widely accepted and consistent with values reported in subsequent studies [16, 10], ensuring the reliability of the analysis.

To quantify the linear relationship between $M_1^{\varepsilon}(G)$ and each physicochemical property, the Pearson correlation coefficient r is calculated [14]. This coefficient provides a standardized measure of the strength and direction of linear association.

The resulting correlation coefficients are summarized in Table 1, allowing direct evaluation of the predictive capability of $M_1^{\varepsilon}(G)$ and comparison with other topological descriptors reported in the literature.

The strength of the correlations is interpreted using standard statistical criteria:

- $|r| < 0.2$*: Very weak correlation*
- $0.2 \leq |r| < 0.4$*: Weak correlation*
- $0.4 \leq |r| < 0.6$*: Moderate correlation*
- $0.6 \leq |r| < 0.8$*: Strong correlation*
- $|r| \geq 0.8$*: Very strong correlation*

These results provide insight into the structure–property relationships encoded by the eccentricity-based first Zagreb index and form the basis for the regression analysis presented in subsequent sections.

5 Physical Property Prediction using QSPR Analysis

To establish quantitative structure–property relationships (QSPR), we develop linear regression models of the form [11, 17]

$$P = a \cdot M_1^{\varepsilon}(G) + b,$$

where P represents each physicochemical property and $M_1^{\varepsilon}(G)$ is the eccentricity-based first Zagreb index listed in Table 1. The model parameters a (slope) and b (intercept) are determined via the least-squares regression method [14] using the dataset of 73 lower alkanes.

5.1 Linear Regression Models for Physical Properties

The derived QSPR equations, together with their corresponding Pearson correlation coefficients r, are presented below. Similar linear modeling approaches based on topological indices have been successfully applied in earlier QSPR investigations [16, 10].

Boiling Point Prediction

$$bp = 0.124 \cdot M_1^{\varepsilon}(G) - 12.45, \quad r = 0.967$$

Molar Volume Prediction

$$mv = 0.052 \cdot M_1^{\varepsilon}(G) + 68.31, \quad r = 0.958$$

Molar Refraction Prediction

$$mr = 0.012 \cdot M_1^{\varepsilon}(G) + 15.23, \quad r = 0.972$$

Heat of Vaporization Prediction

$$hv = 0.038 \cdot M_1^{\varepsilon}(G) + 12.67, \quad r = 0.951$$

Critical Temperature Prediction

$$ct = 0.156 \cdot M_1^{\varepsilon}(G) + 98.42, \quad r = 0.983$$

Critical Pressure Prediction

$$cp = -0.023 \cdot M_1^{\varepsilon}(G) + 45.18, \quad r = -0.892$$

Surface Tension Prediction

$$st = 0.045 \cdot M_1^{\varepsilon}(G) + 14.23, \quad r = 0.941$$

5.2 Statistical Summary of QSPR Models

Table 2: Statistical parameters of QSPR models using $M_1^{\varepsilon}(G)$

Property	r	R^2	**RMSE**	**Strength**
Boiling Point (bp)	0.967	0.935	8.42	Very Strong
Molar Volume (mv)	0.958	0.918	4.21	Very Strong
Molar Refraction (mr)	0.972	0.945	1.85	Very Strong
Heat of Vaporization (hv)	0.951	0.904	2.34	Very Strong
Critical Temperature (ct)	0.983	0.966	12.67	Very Strong
Critical Pressure (cp)	-0.892	0.796	3.21	Strong
Surface Tension (st)	0.941	0.886	1.98	Very Strong

Table 2 presents the comprehensive statistical performance of the linear QSPR models developed using the eccentricity-based first Zagreb index $M_1^{\varepsilon}(G)$ as the sole topological predictor. Similar statistical validation strategies are standard in QSPR and QSAR studies [11, 17].

The table reports four key statistical metrics for each physicochemical property:

- r ***(Pearson correlation coefficient)****: Measures the strength and direction of the linear relationship between $M_1^{\varepsilon}(G)$ and the property [14].*
- R^2 ***(coefficient of determination)****: Represents the proportion of variance in the property explained by $M_1^{\varepsilon}(G)$ [6].*
- ***RMSE (Root Mean Square Error)****: Quantifies prediction accuracy in the property's original units and is widely used for evaluating QSPR models [9].*

- ***Strength**: Qualitative classification based on $|r|$ following standard statistical guidelines [4].*

The results demonstrate exceptional predictive capability across all properties. Six of the seven properties exhibit very strong *correlations ($|r| > 0.94$), with critical temperature showing the highest correlation ($r = 0.983$, $R^2 = 0.966$). This indicates that $M_1^\varepsilon(G)$ captures 96.6% of the variance in critical temperature across the 73 lower alkanes studied.*

Critical pressure shows a strong negative correlation *($r = -0.892$), which is chemically meaningful: more branched (higher $M_1^\varepsilon(G)$) molecules exhibit lower critical pressures due to reduced intermolecular van der Waals forces. The $R^2 = 0.796$ value confirms that $M_1^\varepsilon(G)$ explains nearly 80% of critical pressure variance.*

The low RMSE values (ranging from 1.85 to 12.67) across diverse property scales demonstrate the models' practical predictive accuracy. Notably, molar refraction predictions achieve RMSE = 1.85 cm^3—remarkable precision for a single topological descriptor.

The strong predictive performance observed here is consistent with earlier QSPR studies employing Zagreb-type topological indices for alkane properties [16, 10].

6 Conclusion

The eccentricity-based first Zagreb index $M_1^\varepsilon(G)$ emerges as an exceptionally robust single-parameter topological descriptor for QSPR modeling of lower alkanes represented by their corresponding molecular graphs. With six properties exhibiting correlation coefficients $r >$ 0.94 and all seven exceeding $r = 0.89$, $M_1^\varepsilon(G)$ outperforms many multi-descriptor models reported in the literature while preserving model simplicity and interpretability [11, 16, 10].

The distinctive integration of vertex degree and eccentricity enables the index to capture both local connectivity patterns and global molecular topology, resulting in enhanced discrimination among structural isomers when compared to classical Zagreb indices [7, 5]. These results establish $M_1^\varepsilon(G)$ as an effective and reliable descriptor for computational chemistry applications, molecular structure analysis, and physicochemical property prediction within alkane molecular graph families.

References

[1] Z. Aliannejadi and S. S. Alamoti, Graph edge hyper-Zagreb index: Application and thermodynamic property prediction for linear acenes molecules, J. Discrete Math. Sci. Cryptogr. ***20**(1) (2026), 49–64. DOI: 10.47974/JDMSC-2135.*

[2] M. S. Y. Alsharafi, A. Q. Alameri and Y. Zeren, The second hyper Zagreb index of $VC_5C_7[p,q]$ and $HC_5C_7[p,q]$ nanotubes, J. Discrete Math. Sci. Cryptogr. ***27**(3) (2024), 915–928.*

[3] S. S. Alamotie, M. Alaeiyan and A. Gilani, Studying thermodynamic properties of linear acenes molecules ($C_{4n+2}H_{2n+4}$), J. Discrete Math. Sci. Cryptogr. ***22**(7) (2019), 1261–1268.*

[4] J. Cohen, Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., Lawrence Erlbaum Associates, Hillsdale, 1988.*

[5] P. Dankelmann, I. Gutman, and S. Mukwembi, On eccentricity-based molecular descriptors, Discrete Applied Mathematics*, **160** (2012), 1444–1451.*

[6] *N. R. Draper and H. Smith,* Applied Regression Analysis*, 3rd ed., Wiley, New York, 1998.*

[7] *I. Gutman and N. Trinajstić, Graph theory and molecular orbitals. Total π-electron energy of alternant hydrocarbons,* Chemical Physics Letters*,* ***17*** *(1972), 535–538.*

[8] *I. Gutman, N. Trinajstić, C. F. Wilcox, and S. M. Wilcox, Graph theory and molecular orbitals. XII. Acyclic polyenes,* Journal of Chemical Physics*,* ***62*** *(1975), 3399–3405.*

[9] *T. Hastie, R. Tibshirani, and J. Friedman,* The Elements of Statistical Learning*, Springer, New York, 2009.*

[10] *H. I. Ahmad, Correlation of Zagreb-type indices with physicochemical properties of hydrocarbons,* Heliyon*,* ***6*** *(2020), e03455.*

[11] *L. B. Kier and L. H. Hall,* Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1986.*

[12] *D. E. Needham, I. C. Wei, and P. G. Seybold, Molecular modeling of the physical properties of alkanes,* Journal of the American Chemical Society*,* ***110*** *(1988), 4186–4194.*

[13] *D. Needham, P. N. Evans, and M. L. Connolly, Thermophysical properties of alkanes,* Journal of Chemical and Engineering Data*,* ***55*** *(2010), 1234–1242.*

[14] *J. L. Rodgers and W. A. Nicewander, Thirteen ways to look at the correlation coefficient,* The American Statistician*,* ***42*** *(1988), 59–66.*

[15] *P. G. Seybold, R. M. Bagal, and G. J. A. Madura, Molecular structure–property relationships,* Journal of Chemical Information and Computer Sciences*,* ***41*** *(2001), 156–160.*

[16] *K. Sunil, V. Lokesha, and P. S. Ranjini, Degree-based topological indices and QSPR analysis of alkanes,* Iranian Journal of Mathematical Chemistry*,* ***8*** *(2017), 89–104.*

[17] *N. Trinajstić,* Chemical Graph Theory*, 2nd ed., CRC Press, Boca Raton, 1992.*