Journal of Mathematical Control Science and Applications (JMCSA)
Vol. 3 No. 1 (January-June, 2017), ISSN : 0974-0570

QSPR ANALYSIS OF ALDEHYDES USING THE FIRST MULTIPLICATIVE HYPER-ZAGREB INDEX

MANJULA K. M.

Abstract. Quantitative Structure–Property Relationship (QSPR) analysis is an effective computational approach for predicting the physicochemical properties of chemical compounds using graph-theoretical descriptors. In this study, the first multiplicative hyper-Zagreb index is employed to investigate the thermodynamic properties of twenty-four straight-chain aldehydes. The computed descriptor values are correlated with thermal energy (E_{th}), heat capacity (C_V), and entropy (S) using a logarithmic regression model. The predictive performance of the model is evaluated using the correlation coefficient (r) and the coefficient of determination (R^2). The results demonstrate that the first multiplicative hyper-Zagreb index is an effective molecular descriptor for modeling the thermodynamic properties of aldehydes.

1. Introduction

Aldehydes are a significant class of organic compounds containing a terminal carbonyl functional group ((−CHO)). Their high chemical reactivity makes them important intermediates in a wide range of synthetic and industrial applications. These compounds readily participate in oxidation, reduction, condensation, and polymerization reactions, enabling their use in the production of pharmaceuticals, agrochemicals, polymers, dyes, perfumes, solvents, synthetic resins, and specialty chemicals. Lower aliphatic aldehydes, such as ethanal and propanal, are commonly employed as industrial feedstocks, whereas higher homologues are primarily utilized in the manufacture of surfactants, lubricants, and flavoring agents. Owing to these diverse applications, aldehydes have been extensively investigated for their structural and physicochemical characteristics.

The physicochemical behavior of organic compounds is strongly influenced by their molecular structures. Thermal energy, heat capacity, and entropy are important thermodynamic parameters that provide insight into molecular stability and energy distribution. Thermal energy represents the total internal energy associated with translational, rotational, and vibrational molecular motions. Heat capacity indicates the amount of heat required to raise the temperature of a substance by one degree and reflects the accessibility of molecular energy states. Entropy quantifies the degree of molecular disorder and the distribution of energy among microscopic configurations. These thermodynamic quantities are widely used to

2000 *Mathematics Subject Classification.* 05C90, 05C92, 05C07, 05C38.
Key words and phrases. Degree of a vertex, Topological indices, QSPR analysis.

analyze molecular stability, reaction mechanisms, phase transitions, and chemical kinetics.

Straight-chain aliphatic aldehydes constitute an ideal homologous series for studying the influence of molecular size on thermodynamic properties. As the number of carbon atoms increases from ethanal to pentacosanal, molecular mass and conformational flexibility increase, resulting in systematic variations in thermal energy, heat capacity, and entropy. Such predictable trends provide a suitable basis for establishing quantitative relationships between molecular structure and thermodynamic behavior. Consequently, homologous aldehydes have frequently been employed as model compounds in quantitative structureproperty relationship (QSPR) studies [6].

Graph theory has become an effective mathematical approach for representing molecular structures and analyzing their properties. In this framework, atoms and chemical bonds are represented by vertices and edges, respectively, and the resulting molecular graphs are characterized using numerical descriptors known as topological indices. These descriptors efficiently encode structural information and have been successfully applied to predict a wide variety of physicochemical, biological, and pharmacological properties. As a result, graph-theoretical descriptors provide a computationally efficient alternative to extensive experimental investigations.

Reliable thermodynamic data are essential for developing and validating QSPR models. Experimental or computational values of thermal energy ((E_{th})), heat capacity ((C_V)), and entropy ((S)) serve as suitable response variables for assessing the predictive performance of molecular descriptors. The systematic variation in the thermodynamic properties of homologous aldehydes makes this series particularly suitable for investigating structureproperty relationships [8].

In this study, a dataset comprising twenty-four straight-chain aldehydes ranging from ethanal to pentacosanal is considered. The molecular formula together with the corresponding values of thermal energy ((E_{th})), heat capacity ((C_V)), and entropy ((S)) for each compound are presented in Table 1. These thermodynamic properties are employed as dependent variables in the development of QSPR models based on graph-theoretical descriptors. The selected dataset enables a comprehensive evaluation of the predictive capability of the proposed descriptors and regression models.

This investigation aims to establish quantitative relationships between graph-theoretical descriptors and the thermodynamic properties of straight-chain aldehydes. The developed QSPR models are expected to demonstrate the effectiveness of graph-based descriptors in predicting important physicochemical properties with high accuracy. The findings contribute to the advancement of computational molecular modeling and support the development of efficient tools for molecular property prediction and rational molecular design.

2. QSPR Analysis of Aldehydes

Quantitative Structure–Property Relationship (QSPR) analysis is an established computational approach for investigating the relationship between the molecular

structure of chemical compounds and their physicochemical properties. By utilizing graph-theoretical molecular descriptors, QSPR models enable the prediction of experimentally measurable properties directly from molecular structures, thereby minimizing the need for expensive and time-intensive laboratory experiments. Due to their reliability, computational efficiency, and predictive capability, QSPR techniques have become valuable tools in chemistry, pharmaceutical research, materials science, and chemical engineering.

In the present study, a homologous series of twenty-four straight-chain aldehydes, extending from ethanal to pentacosanal, is selected for QSPR investigation. Aldehydes are an important class of organic compounds characterized by the presence of a terminal aldehyde functional group ($-$CHO). They are widely employed in the manufacture of pharmaceuticals, polymers, fragrances, solvents, and numerous specialty chemicals. The systematic increase in carbon chain length within this homologous series provides an ideal framework for examining the influence of molecular structure on thermodynamic behavior. The molecular formulas and the corresponding values of thermal energy (E_{th}), heat capacity (C_V), and entropy (S) are summarized in Table 2. These thermodynamic parameters are considered as the response variables for the present QSPR analysis [2, 3].

Thermal energy (E_{th}) represents the total internal energy arising from the translational, rotational, and vibrational motions of molecules. Heat capacity (C_V) describes the quantity of heat required to raise the temperature of a substance by one unit under constant-volume conditions, whereas entropy (S) quantifies the degree of molecular randomness and the distribution of energy among accessible microscopic states. Since these thermodynamic properties are closely associated with molecular size, connectivity, and structural complexity, they constitute suitable target properties for QSPR modeling [1].

In this investigation, the first multiplicative hyper-Zagreb index, denoted by $HII_1(G)$, is employed as the molecular descriptor. The descriptor values are calculated from the molecular graphs of the selected aldehydes and subsequently correlated with the thermodynamic properties using linear, quadratic, cubic, logarithmic, power, and exponential regression models. The predictive performance of each regression model is assessed using the correlation coefficient (r) and the coefficient of determination (R^2). The computed values of $HII_1(G)$ for all the considered aldehydes are listed in Table 1, whereas the corresponding regression analyses and discussion are presented in the subsequent sections [5].

The first multiplicative hyper-Zagreb index, denoted by $HII_1(G)$, was introduced by Kulli [4] and is defined as follows:

$$HII_1(G) = \prod_{uv \in E(G)} \left(d_G(u) + d_G(v)\right)^2,$$

here $d_u(G)$ and $d_v(G)$ denote the degrees of the vertices u and v in the graph G, respectively, while $E(G)$ represents the edge set of G.

TABLE 1. First Multiplicative Hyper-Zagreb Index of the considered aldehydes

Compound	$HII_1(G)$
Ethanal	81
Propanal	1296
Butanal	20736
Pentanal	331776
Hexanal	5308416
Heptanal	84934656
Octanal	1358954496
Nonanal	21743271936
Decanal	3.47892×10^{11}
Undecanal	5.56628×10^{12}
Dodecanal	8.90604×10^{13}
Tridecanal	1.42497×10^{15}
Tetradecanal	2.27995×10^{16}
Pentadecanal	3.64792×10^{17}
Hexadecanal	5.83667×10^{18}
Heptadecanal	9.33866×10^{19}
Octadecanal	1.49419×10^{21}
Nonadecanal	2.39070×10^{22}
Eicosanal	3.82512×10^{23}
Heneicosanal	6.12019×10^{24}
Docosanal	9.79230×10^{25}
Tricosanal	1.56677×10^{27}
Tetracosanal	2.50683×10^{28}
Pentacosanal	4.01093×10^{29}

TABLE 2. Structural details and their thermal energy (E_{th}), heat capacity (C_V) and entropy (S) for the aldehydes used in the present study

Compound	**Formula**	**No.**	$\mathbf{E}_{\text{th}}$ (kcal/mol)	$\mathbf{C_V}$ (cal/molK)	**S** (cal/molK)
Ethanal	C_2H_4O	1	40.074	10.472	62.483
Propanal	C_3H_6O	2	60.162	14.719	69.255
Butanal	C_4H_8O	3	80.131	19.195	76.521
Pentanal	$C_5H_{10}O$	4	100.175	23.569	83.110
Hexanal	$C_6H_{12}O$	5	132.672	26.877	90.594
Heptanal	$C_7H_{14}O$	6	139.940	32.586	97.626
Octanal	$C_8H_{16}O$	7	159.879	37.083	104.813
Nonanal	$C_9H_{18}O$	8	179.812	41.597	112.474
Decanal	$C_{10}H_{20}O$	9	199.764	46.095	119.946
Undecanal	$C_{11}H_{22}O$	10	219.717	50.587	127.356
Dodecanal	$C_{12}H_{24}O$	11	239.673	55.061	134.061
Tridecanal	$C_{13}H_{26}O$	12	259.626	59.553	141.197
Tetradecanal	$C_{14}H_{28}O$	13	279.545	64.080	149.267
Pentadecanal	$C_{15}H_{30}O$	14	299.493	68.575	156.600
Hexadecanal	$C_{16}H_{32}O$	15	319.434	73.070	163.611
Heptadecanal	$C_{17}H_{34}O$	16	339.444	77.480	171.456
Octadecanal	$C_{18}H_{36}O$	17	360.034	81.825	178.205
Nonadecanal	$C_{19}H_{38}O$	18	380.088	86.245	185.815
Eicosanal	$C_{20}H_{40}O$	19	400.070	90.739	193.172
Heneicosanal	$C_{21}H_{42}O$	20	420.055	95.226	200.520
Docosanal	$C_{22}H_{44}O$	21	440.039	99.715	207.846
Tricosanal	$C_{23}H_{46}O$	22	460.024	104.206	215.418
Tetracosanal	$C_{24}H_{48}O$	23	480.006	108.692	222.462
Pentacosanal	$C_{25}H_{50}O$	24	499.992	113.185	230.223

TABLE 3. Statistical parameters for the logarithmic QSPR model for the first multiplicative hyper-Zagreb index $HII_1(G)$ and the physicochemical properties of aldehydes

Physicochemical properties	R	R^2	Adjusted R^2	F	Sig.
$\mathbf{E}_{\text{th}}$ (kcal/mol)	1.000	1.000	1.000	67012.756	0.000
$\mathbf{C_V}$ (cal/molK)	1.000	1.000	1.000	293587.418	0.000
$\mathbf{S}$ (cal/molK)	1.000	1.000	1.000	383621.119	0.000

Logarithmic Regression Model: By utilizing the data presented in Tables 1 and 2, logarithmic regression models are developed for the selected physicochemical properties of aldehydes. The logarithmic regression model is expressed as

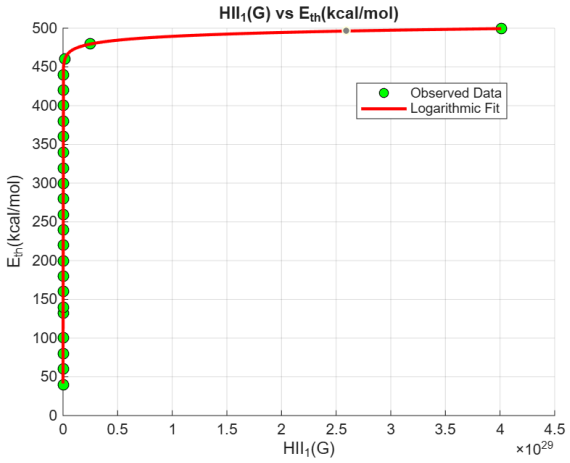
$$y = a + b\log(x),$$

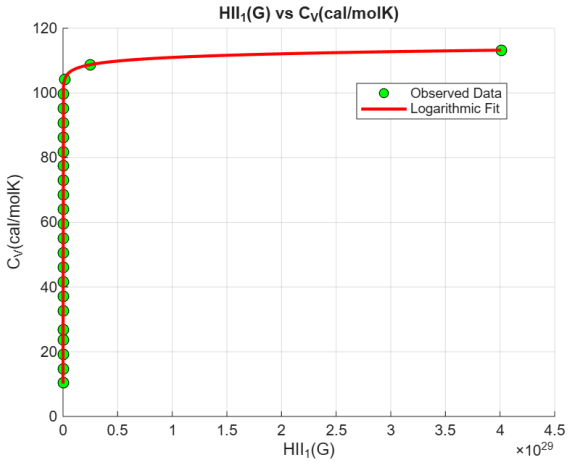
where y denotes the physicochemical property, a is the regression coefficient (intercept), b is the regression constant (slope), and x represents the molecular descriptor. The logarithmic regression models corresponding to the first multiplicative hyper-Zagreb index are given below:

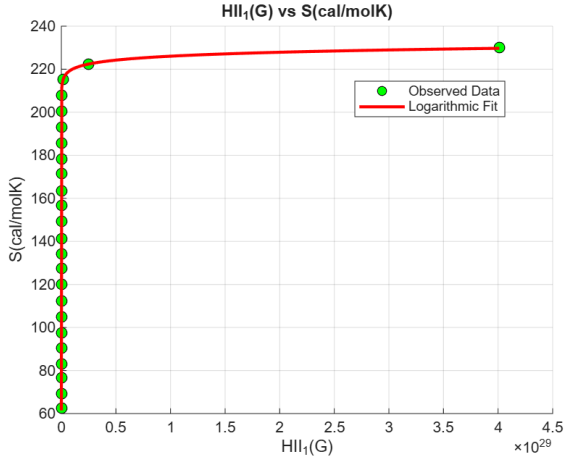
The logarithmic regression models for the first multiplicative hyper-Zagreb index $HII_1(G)$ are given by:

$$\begin{aligned}
\mathbf{E}_{\text{th}}\,(\text{kcal/mol}) &= 51.130 + 5.491\,\log(HII_1),\\
\mathbf{C_V}\,(\text{cal/molK}) &= 2.815 + 0.334\,\log(HII_1),\\
\mathbf{S}\,(\text{cal/molK}) &= 36.396 + 2.846\,\log(HII_1).
\end{aligned}$$

The correlations of the first multiplicative hyper-Zagreb index, $HII_1(G)$, with the selected physicochemical properties of aldehydes obtained using the logarithmic regression model are depicted in the figures below.







3. Conclusion

In this study, the first multiplicative hyper-Zagreb index was employed as a molecular descriptor to investigate the thermodynamic properties of twenty-four straight-chain aldehydes through QSPR analysis. The descriptor values were correlated with thermal energy (E_{th}), heat capacity (C_V), and entropy (S) using a logarithmic regression model. The obtained values of the correlation coefficient (R) and the coefficient of determination (R^2) demonstrate a strong correlation between the proposed descriptor and the selected physicochemical properties. Therefore, the first multiplicative hyper-Zagreb index can be regarded as an effective molecular descriptor for predicting the thermodynamic properties of aldehydes and may serve as a useful tool in future QSPR investigations of related organic compounds.

References

[1] A. Alaghebandi and F. Shafiei, *QSPR modeling of heat capacity, thermal energy and entropy of aliphatic Aldehydes by using Topological Indices and MLR method*, Iranian Journal of Mathematical Chemistry, **7**(2) (2016), 235–251.

[2] D. Bonchev, *Over all connectivities/topological complexities: A new powerful tool for QSPR/QSAR*, J. Chem. Inf. Comput. Sci., **40** (2000), 934–941.

[3] A. A. Gakh, E. G. Gakh, B. G. Sumpter, D. W. Noid, *Neural network graph theory approach to the prediction of the physical properties of organic compounds*, J. Chem. Inf. Comput. Sci., **34**(4) (1994), 832–839.

[4] V. R. Kulli, *Multiplicative hyper-Zagreb indices and coindices of graphs*, International Journal of Pure Algebra, **6**(7) (2016), 342–347.

[5] Z. Lin, J. Xu, X. Zheng, Z. Li, *Study on quantitative structure- property relationship of chain hydrocarbons, aldehydes and alkanones by molecular distance-edge vector*, Acta Phys. Chem. Sin., **16** (2000), 153–161.

[6] V. Sharma, R. Goswami, A. K. Madan, *A novel highly discriminating topological descriptor for structureproperty and structureactivity studies*, J. Chem. Inf. Comput. Sci., **37** (1997), 273–282.

[7] G. H. Shirdel, H. Rezapour and A. M. Sayadi, *The Hyper–Zagreb Index of Graph Operations*, Iranian Journal of Mathematical Chemistry, **4**(2) (2013), 213–220.

[8] N. Trinajstić, *Chemical Graph Theory*, 2nd Edition, CRC Press, Boca Raton, (1992).

Manjula K. M.: Department of Mathematics, Government First Grade College for Women, Hassan-573 202, India.

E-mail address: manjula.km.gowda@gmail.com