Investigating the Connection between Atom Connectivity, Connectivity Energy, and the Physico-Chemical Properties of Amino Acids

M. Kamal Kumar[1], Shajidmon Kolamban[1], Bushra Hibras Al Sulaimi[1], Salah Salim Rashid Khanayafar Al Hajri[2], Hour Juma Nasser Al Harbi[3]

[1, 3,] Department of Information Technology, College of Computing and Information Sciences, University of Technology and Applied Sciences, Muscat, Oman
[2] Department of Applied Sciences, College of Applied Sciences and Pharmacy, University of Technology and Applied Sciences, Muscat, Oman

E-mail: kamalmvz@gmail.com*, shajidmon.kolamban@utas.edu.om, bushra.alsulaimi@utas.edu.om, 34j2145@utas.edu.om, 16s2013@utas.edu.om.

ABSTRACT

This research explores the use of molecular graph theory to predict the physical and chemical properties of twenty amino acids. Specifically, it examines how the energy of a graph, derived from the eigenvalues of an adjacency matrix, correlates with the chemical reactivity of amino acids. In molecular graph theory, a chemical compound is represented as a graph, where atoms are treated as nodes and bonds as edges. The energy of the graph is calculated by summing the absolute values of the eigenvalues of the adjacency matrix.
However, the basic adjacency matrix treats all atoms and bonds as equivalent, which can limit the accuracy of predictions. To address this, the research introduces atom-connectivity matrices and connectivity matrices to better differentiate between different types of atoms and bond connectivity. The key proposition of this research is that by incorporating atom-connectivity matrices, predictions of amino acid properties are more accurate than those based solely on the adjacency matrix.

AMS Classification (2000): 15A45, 05C50, 05C69, 05C90.

Key Words: Adjacency matrix; Molecular graph; Energy of Graph; Atom Connectivity; Connectivity Energy; Regression Models.

1. INTRODUCTION

In industrial chemistry the concept of energy of graph was introduced. It was noticed that the total electron energy was related to certain numerical values. The generation of heat of hydrocarbon is one of the examples of the numerical quantity. This numerical quantity could be evaluated as the energy of the molecular graph associated with the chemical compound. In the molecular graph the vertices, the position of the atoms and the edges are the bonds between the two atoms in the molecule. Linear transformation and the related geometry could be predicted using the Eigen values and Eigen vectors.

In the year 1940 Coulson [3] studied total *pi*-electron energy (E) and general characteristics with respect to mathematical sciences. This study was completely with respect to the Huckel Molecular Orbital model. Because of Coulson efforts and interest in this field of sciences, many other mathematicians and scientists showed their interest in this field. Many authors [2-6, 8-17] obtained exhaustive results on total *pi*-electron energy (E) and its dependence on the general molecular structure and other parameters of the molecule.

There were significant results obtained using the relation between the physical and chemical properties of the molecules and the energy of graphs. Hence this field of study was chosen by a lot of researchers for their study and enormous research papers were obtained in the field of graph energies. One serious drawback for the energy of the graph is that the adjacent matrix would represent the different types of atoms as a single node without differentiating between them. Also, the types of bonds like single, double and triple were represented as a single edge. Hence the importance of these different types of bonds and atoms couldn't be figured out in the energy of the graph. Hence to differentiate the different types of atoms and bond connectivity atom-connectivity and connectivity matrices are introduced in the study of energy of graph. The chemical applications of atom connectivity energy and connectivity energy and various types of energies were studied [1, 25, 28-30].

Generally, determining the chemical and physical properties of a compound can range from very easy to very complex, depending on factors such as the nature of the compound, the type of property, the available data and the methods used. For simple compounds, these properties are often readily available through experimental data or standard reference materials. For more complex compounds, especially novel ones, it may require advanced computational or experimental techniques to obtain accurate results. Hence, we are exploring the prediction of the physical and chemical properties of the compound through a much easier method by using the special type of matrices. This new innovative approach is to analyze the different types of energy of graph and their significance with respect to the different type of atoms and bonds in the molecules. Many mathematicians explore different energy concepts in graphs without connecting them to molecular structures. Hence the energy of graph was represented as theoretical chemistry and not practical chemistry. By the introduction of atom connectivity matrix (ACM) and in general connectivity matrix (CM) the energy of the graph can be used for practical chemistry too.

This study explores how these matrices (atom-connectivity and connectivity) can be used to predict the physical and chemical properties of compounds, specifically amino acids. Amino acids are the building blocks of proteins, and their chemical properties, such as solubility, reactivity, and stability, can be influenced by the molecular structure, which can be studied through Energy of the molecular graph.

This paper is a unique approach to find the physico-chemical properties of Amino Acids. The study could be extended to other molecules and find the unique properties of any arbitrary molecule using the concept of mathematical graph energies.

2. LITERATURE REVIEW

In the last three decades the study on graph energies has grown exponentially. There are thousands of papers published all around the world related to this field of mathematics. Ivan Gutman who introduced the mathematical concept of energy of graph in 1978 had done a survey on the energy of graphs and found that there are more than 70 different types of energy of graph introduced by various research across the globe [8]. It is surprising to note that to date there is no unique way to represent any molecule in the computer. The representation of the molecule in the computer helps the scientists and researchers to do huge computations easily. Among the many ways to represent a molecule in the computer, one way is the atom connectivity and connectivity matrix. This matrix gives the information of different types of atoms and the different types of bonds of the molecules. [1, 28-30]. Different mathematicians had introduced various types of energies of graph and had given exhaustive results in this field of mathematics. Due to simplicity of this concept extensive results are obtained in this field energy of graph. [19-25].

Among the various types of energies of graphs few are very common in usage in literature. Few to mention like Energy of graph, Distance Energy of graph, Laplace Energy Graph, Sign less Energy of Graph, Sign less Laplace Energy of Graph also for all the above energies of the graph the principal diagonal elements are zero hence we introduced the domination set in the principal diagonal elements which led to the field of domination energy of the respective energies of the graph.

To do the computation on molecules in the computer precise and unique information of the molecules need to be stored in the system. Usually, this precise and unique information of the molecules is stored in the form of matrices. Hence

in 1988, S. C. Basak and V. R. Magnuson introduced Atom connectivity [1] for computer representation of molecules, Spialter [28-30] introduced the concept of connectivity matrix. Mathavi Manisekar and S. Lalitha have also obtained few results in Dissociation Energy for Amino Acids [26].

2.1. DEFINITIONS

All the equations has been chosen so as to be fully analogous to the definition of Energy of Graph: $E = E(G) = \sum_{i=1}^{n} |\lambda_i|$ (1)

Where $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq, \ldots\ldots\ldots, \geq \lambda_n$ are the ordinary graph eigenvalues, that is, the eigenvalues of the adjacency matrix of $A(G)$ of graph G. Similar to the above-mentioned definition, the energy obtained from the Atom connectivity matrix is called Atom connectivity energy and the energy obtained from connectivity matrix is called connectivity energy of the molecular graph of the molecules.

The atom-connectivity matrix, denoted by ACM, and defined as the matrix whose elements as given as $[ACM]_{ij} = \begin{cases} b_{ij} & if\ vertices\ i\ and\ j\ are\ adjacent \\ s_i & if\ i = j \\ 0 & otherwise \end{cases}$ (2)

Where b_{ij} is the bond order (1, 1.5, 2, 3) single, double and triple bound between atoms i *and* j, and s_i stands for the chemical symbol of the atom i, for matrix calculation s_i is considered to be zero.

Hydrogen-suppressed atom-connectivity matrix denoted by HS-ACM [28-30] is the matrix obtained for the molecule without considering the hydrogen atoms for calculation. This was basically done to save the system processing time.

The connectivity matrix, denoted by CM, and defined as the matrix whose elements as given as $[CM]_{ij} = \begin{cases} b_{ij} & if\ vertices\ i\ and\ j\ are\ adjacent \\ n_i & if\ i = j \\ 0 & otherwise \end{cases}$ (3)

Where b_{ij} is the bond order between atoms i and j, and n_i stands for the atomic number of the atom i.

3. RESULTS AND DISCUSSION

Peptide and Proteins are formed by combining organic compounds of amino acids. Amino acids are called the foundation blocks for Proteins and Peptide. Basic amino groups and carboxyl groups are the basic compounds containing the amino acids as given below in the figure 1. Amino groups are represented by

(−NH2) and carboxyl groups are represented by (−COOH). In the formation of proteins 20 amino acids are involved.

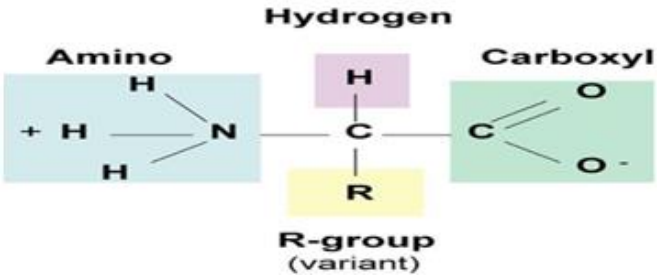


Figure 1: Basic Amino Acid Structure.

Hellberg [18], developed three principal properties for 20 amino acids and variation in amino acid sequence within sets of peptides is described by three principal properties hydrophilicity (z1), bulk (z2) and electronic properties (z3). These are derived from principal components analysis (PCA) of a matrix of 29 physicochemical variables for 20 amino acids. Apart from this, two descriptors are derived from the side chain surface area and atomic charge. One descriptor is isotropic surface area (ISA) is the portion of the solute molecule, which is accessible to nonspecific interactions with the solvent, water. The other one is electronic charge index (ECI), ECI is the sum of absolute value of CNDO/2 charges of the side chain atoms. ISA and ECI were derived and QSAR methodologies were used to correlate structural variation with variation in activity for the sets of peptides. One physical property of amino acid is molecular weight (MW), molecular weight is a measure of the sum of the atomic weight values of the atoms in a molecule. Molecular weight is used in chemistry to determine stoichiometry in chemical reactions and equations. Molecular weight is commonly abbreviated by M.W. or MW. Molecular weight is either unit less or expressed in terms of atomic mass units (amu) or Daltons (Da). Mass is the average mass of amino acid.

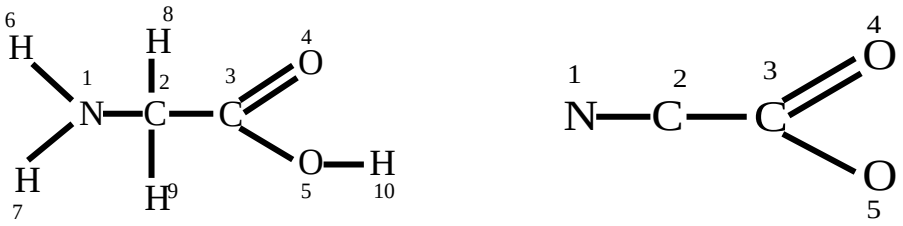


Figure 2: The structure of Glycine labeled with hydrogen and without hydrogen.

Like the above amino acid glycine all the twenty amino acids were labeled, and various energies were found and listed in the tabulated in the table 2 below.

	1	2	3	4	5	6	7	8	9	10
1	0	1	0	0	0	1	1	0	0	0
2	1	0	1	0	0	0	0	1	1	0
3	0	1	0	1	1	0	0	0	0	0
4	0	0	1	0	0	0	0	0	0	0
5	0	0	1	0	0	0	0	0	0	1
6	1	0	0	0	0	0	0	0	0	0
7	1	0	0	0	0	0	0	0	0	0
8	0	1	0	0	0	0	0	0	0	0
9	0	1	0	0	0	0	0	0	0	0
10	0	0	0	0	1	0	0	0	0	0

	1	2	3	4	5
1	0	1	0	0	0
2	1	0	1	0	0
3	0	1	0	1	1
4	0	0	1	0	0
5	0	0	1	0	0

	1	2	3	4	5	6	7	8	9	10
1	N	1	0	0	0	1	1	0	0	0
2	1	C	1	0	0	0	0	1	1	0
3	0	1	C	2	1	0	0	0	0	0
4	0	0	2	O	0	0	0	0	0	0
5	0	0	1	0	O	0	0	0	0	1
6	1	0	0	0	0	H	0	0	0	0
7	1	0	0	0	0	0	H	0	0	0
8	0	1	0	0	0	0	0	H	0	0
9	0	1	0	0	0	0	0	0	H	0
10	0	0	0	0	1	0	0	0	0	H

	1	2	3	4	5
1	N	1	0	0	0
2	1	C	1	0	0
3	0	1	C	2	1
4	0	0	2	O	0
5	0	0	1	0	O

	1	2	3	4	5	6	7	8	9	10
1	7	1	0	0	0	1	1	0	0	0
2	1	6	1	0	0	0	0	1	1	0
3	0	1	6	2	1	0	0	0	0	0
4	0	0	2	8	0	0	0	0	0	0
5	0	0	1	0	8	0	0	0	0	1
6	1	0	0	0	0	1	0	0	0	0
7	1	0	0	0	0	0	1	0	0	0
8	0	1	0	0	0	0	0	1	0	0
9	0	1	0	0	0	0	0	0	1	0
10	0	0	0	0	1	0	0	0	0	1

	1	2	3	4	5
1	7	1	0	0	0
2	1	6	1	0	0
3	0	1	6	2	1
4	0	0	2	8	0
5	0	0	1	0	8

The above matrices are the adjacency matrices, Atom Connectivity matrices and Connectivity matrices with and without hydrogen respectively. These matrices are used for calculating Energy of Graph, Atom Connectivity Energy of Graph and Connectivity Energy of Graph respectively. For calculations, the atoms names are replaced with 0 in Atom Connectivity matrices.

S.No	Amino acid	SY	z_1	z_2	z_3	*ISA*	ECI	MW	M
1	Alanine	A	0.07	-1.7	0.09	62.9	0.1	89	71.0788
2	Valine	V	-2.7	-2.5	-1.3	120.9	0.1	117	99.1326
3	Leucine	L	-4.2	-1	-1	154.4	0	131	113.1594
4	Isoleucine	I	-4.4	-1.7	-1	149.8	0.1	131	113.1594
5	Proline	P	-1.2	0.88	2.23	122.4	0.2	115	97.1167
6	Phenylalanine	F	-4.9	1.3	0.45	189.4	0.1	165	147.1766
7	Tryptophan	W	-4.8	3.65	0.85	179.2	1.1	204	186.2132
8	Methionine	M	-2.5	-0.3	-0.4	132.2	0.3	149	131.1986
9	Lysine	K	2.84	1.41	-3.1	102.8	0.5	147	128.1741
10	Arginine	R	2.88	2.52	-3.4	52.98	1.7	175	156.1875
11	Histidine	H	2.41	1.74	1.11	87.38	0.6	155	137.1411
12	Glycine	G	2.23	-5.4	0.3	19.93	0	75	57.0519
13	Serine	S	1.96	-1.6	0.57	19.75	0.6	105	87.0782
14	Threonine	T	0.92	-2.1	-1.4	59.44	0.7	119	101.1051
15	Cysteine	C	0.71	-1	4.13	78.51	0.2	121	103.1388
16	Tyrosine	Y	-1.4	2.32	0.01	132.2	0.7	181	163.176
17	Asparagine	N	3.22	1.45	0.84	17.87	1.3	132	114.1039
18	Glutamine	Q	2.18	0.53	-1.1	19.53	1.4	146	128.1307
19	Aspartic	D	3.64	1.13	2.36	18.46	1.3	132	115.0886
20	Glutamic	E	3.08	0.39	-0.1	30.19	1.3	146	129.1155

Table 1: Physico-chemical properties of Amino acids [7, 27, 33].

S. No	Amino acid	S Y	MCD EFG	EWH	EW OH	ACE WH	ACEW OH	CEW H	CEW OH
1	Alanine	A	2.97	13.85	6	15.62	7.57	47.93	41
2	Valine	V	5.12	13.87	6	15.62	7.57	48	41
3	Leucine	L	5.92	20.57	9.34	22.20	10.90	62.42	59.00
4	Isoleucine	I	6.28	20.57	10.07	22.59	11.65	62.43	58.99
5	Proline	P	5.21	19.88	9.68	21.64	11.23	63.00	53

6	Phenylalanine	F	8.25	26.82	14.57	29.16	17.58	87.99	76.99
7	Tryptophan	W	10.67	32.32	19.07	37.53	24.87	107.9	95.99
8	Methionine	M	5.83	20.39	10.18	22.15	11.74	78.99	69.00
9	Lysine	K	6.35	26.49	11.24	28.24	12.80	81.00	66.00
10	Arginine	R	7.77	28.64	13.44	31.73	16.53	94.00	80
11	Histidine	H	8.23	23.43	13.23	26.83	16.05	82	72.99
12	Glycine	G	3.34	10.81	5.23	12.59	6.77	40	35
13	Serine	S	4.36	15.64	7.66	17.40	9.22	56.00	49.00
14	Threonine	T	4.94	18.78	8.42	20.53	9.98	63.99	55
15	Cysteine	C	4.29	15.64	7.66	17.40	9.22	64.00	57.07
16	Tyrosine	Y	9.40	27.91	15.69	31.99	20.25	96.00	84.99
17	Asparagine	N	5.11	18.84	9.34	22.40	12.39	70.00	62.00
18	Glutamine	Q	6.77	21.99	10.92	25.55	14.01	78	68.00
19	Aspartic	D	5.04	16.42	9.34	19.75	12.39	69.00	63
20	Glutamic	E	6.56	19.57	10.92	22.90	14.01	77.0	69

Table 2: Various energies of the graphs [7, 27, 33].

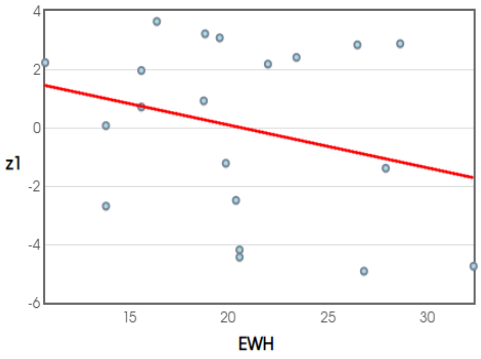
EWH: Energy of graph with hydrogen
EWOH: Energy of graph without hydrogen
ACEWH: Atom Connectivity Energy of graph with hydrogen
ACEWOH: Atom Connectivity Energy of graph without hydrogen
CEWH: Connectivity Energy of graph with hydrogen
CEWOH: Connectivity Energy of graph without hydrogen
MCDEFG: Minimum Connected Dominating Energy of fuzzy graphs [27]
M: Average mass of amino acid
$z_{1:}$ Hydrophilicity, $z_{2:}$ Bulk, z_3 : Electronic Properties, ISA: Isotropic Surface Area, ECI : Electronic Charge Index, MW : Molecular Weight

The above tables are not merely computer-based algorithms or direct calculations. In turn it's the new type of parameter defined in literature to find the property of any molecule using a very simple concept of mathematics. In order words we are in the process of finding the property of molecule not in labs but in computer-based calculations using matrices and their eigen values as a predictable mathematical-Modelling.

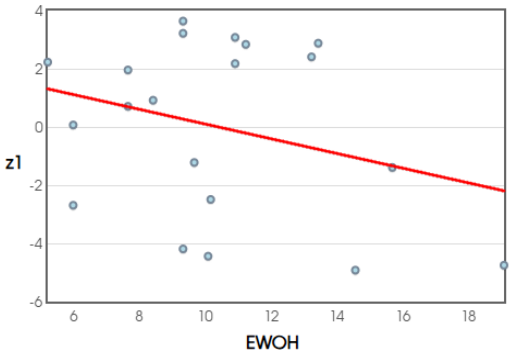
4. REGRESSION MODELS

In this section, we consider the following very simple regression models: Linear regression model: $Y = ax^2 + b$ and Quadratic regression model: $Y = ax^2 + bx + c$. For this study, the seven physico-chemical properties of amino acids (which are listed in Table 1) and the minimum connected dominating energy of the molecular fuzzy graphs (MCDEFG) of amino acids [27] and atom connectivity and connectivity energy of the graph are considered for the regression analysis. The various energies of the graphs are compared with all these physico-chemical properties of amino acids, and we have obtained the following results. Considering

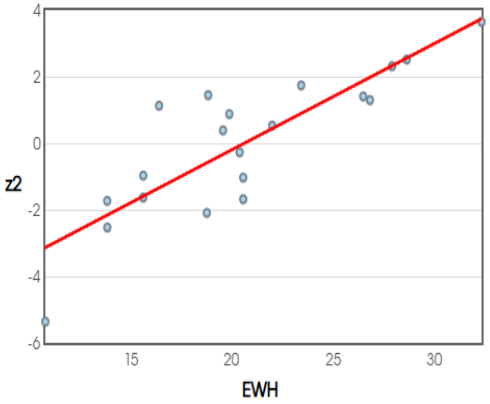
Null Hypothesis $H_0 : \beta = 0$, Alternative Hypothesis, $H_\alpha : \beta \neq 0$, $\alpha = 0.05$



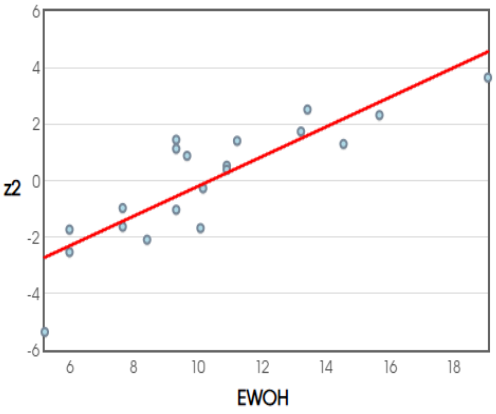
Regression Line: $Z1 = -0.1471 \cdot EWH + 3.0354$
Correlation: $r = -0.2754$
R-squared: $r^2 = 0.0758$



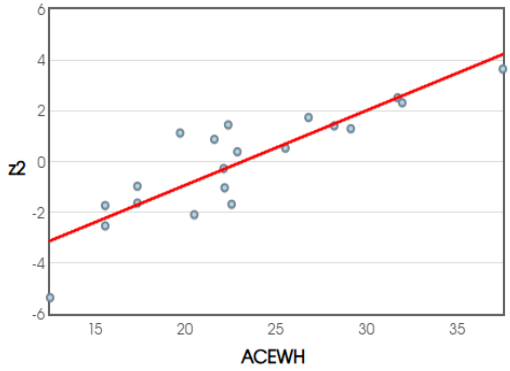
Regression Line: $Z1 = -0.253 \cdot EWOH + 2.6335$
Correlation: $r = -0.2918$
R-squared: $r^2 = 0.0851$



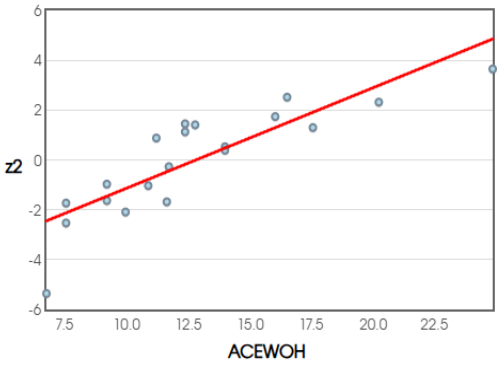
Regression Line: $Z2 = 0.3196 \cdot EWH - 6.5901$
Correlation: $r = 0.8422$
R-squared: $r^2 = 0.7092$



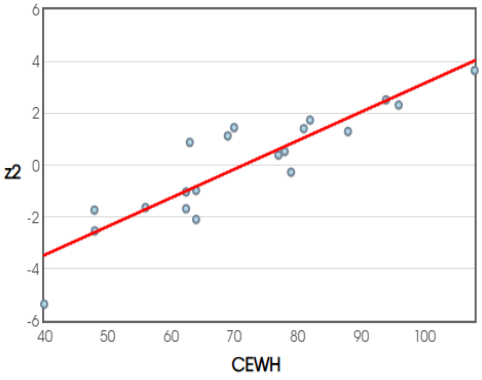
Regression Line: $Z2 = 0.5256 \cdot EWOH - 5.4647$
Correlation: $r = 0.853$
R-squared: $r^2 = 0.7277$



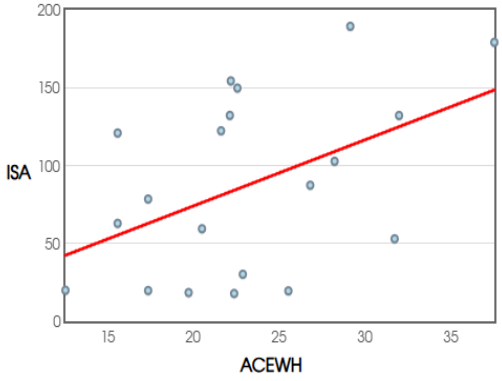
Regression Line: $Z2 = 0.2945 \cdot ACEWH - 6.8277$
Correlation: $r = 0.8704$
R-squared: $r^2 = 0.7576$



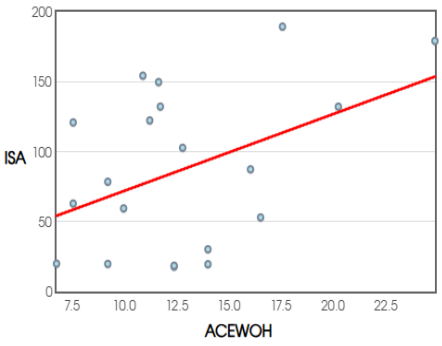
Regression Line: $Z2 = 0.4035 \cdot ACEWOH - 5.1788$
Correlation: $r = 0.8514$
R-squared: $r^2 = 0.725$



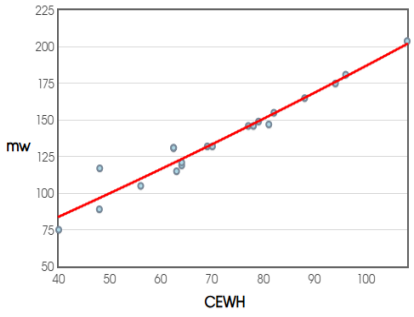
Regression Line: $Z2 = 0.1103 \cdot CEWH - 7.8824$
Correlation: $r = 0.8988$
R-squared: $r^2 = 0.8078$



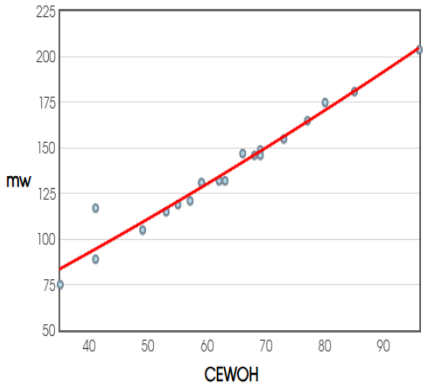
Regression Line: $ISA = 4.2543 \cdot ACEWH - 11.1572$
Correlation: $r = 0.4655$
R-squared: $r^2 = 0.2167$

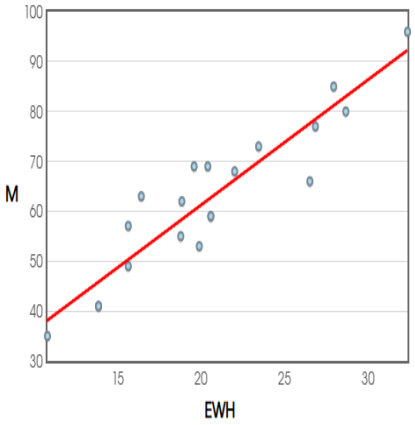


Regression Line: $ISA = 5.4979 \cdot ACEWOH + 16.9251$
Correlation: $r = 0.4295$
R-squared: $r^2 = 0.1844$



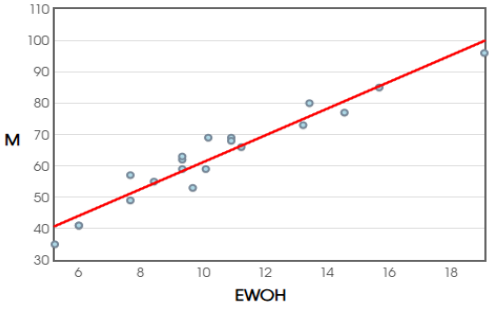
Regression Polynomial: $MW = 0.0021 \cdot CEWH^2 + 1.4187 \cdot CEWH + 23.7324$
R-squared: $r^2 = 0.9505$
Adjusted R-squared: $r^2_{adj} = 0.9477$
Residual Standard Error: 7.2697 on 17 degrees of freedom

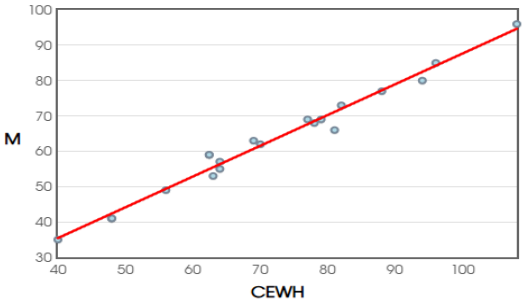




Regression Polynomial: $\mathrm{MW} = 0.0034 \cdot \mathrm{CEWOH}^2 + 1.544 \cdot \mathrm{CEWOH} + 25.5546$
R-squared: $r^2 = 0.9594$
Adjusted R-squared: $r^2_{\text{adj}} = 0.9571$
Residual Standard Error: 6.5819 on 17 degrees of freedom

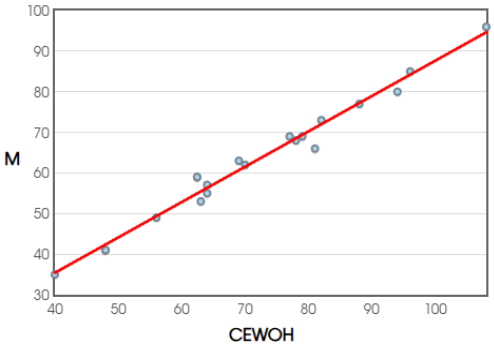
Regression Line: $\mathrm{M} = 2.5171 \cdot \mathrm{EWH} + 10.9423$
Correlation: $r = 0.9248$
R-squared: $r^2 = 0.8552$

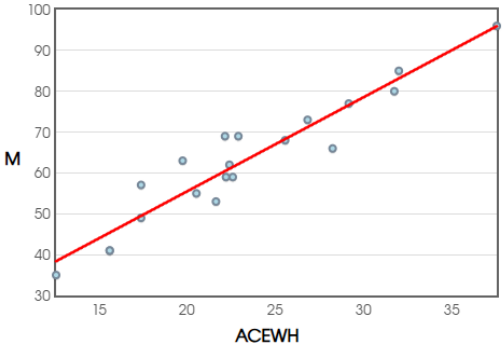




Regression Line: $\mathbf{M = 4.2645 \cdot EWOH + 18.5017}$
Correlation: $r = 0.9651$
R-squared: $r^2 = 0.9313$

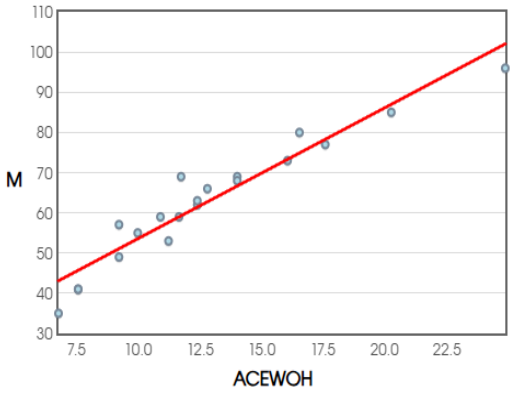
Regression Line: $\mathbf{M = 0.8711 \cdot CEWH + 0.5769}$
Correlation: $r = 0.9899$
R-squared: $r^2 = 0.98$



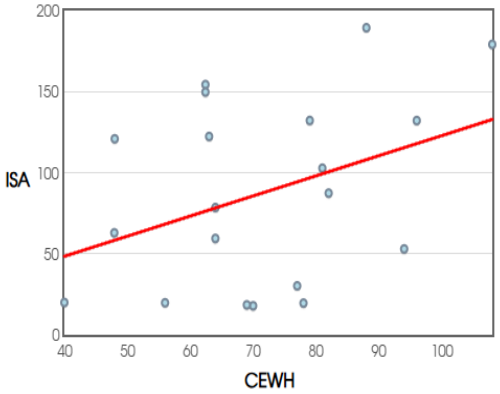


Regression Line: $\mathbf{M = 0.8711 \cdot CEWOH + 0.5769}$
Correlation: $r = 0.9899$
R-squared: $r^2 = 0.98$

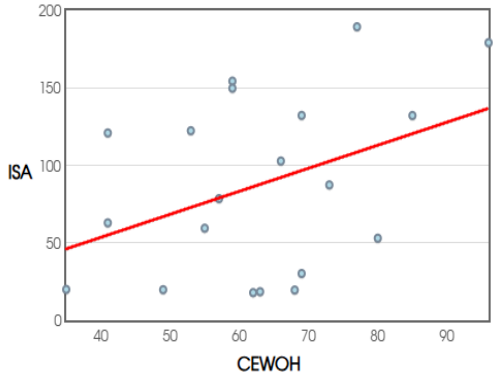
Regression Line: $\mathbf{M = 2.306 \cdot ACEWH + 9.3752}$
Correlation: $r = 0.9504$
R-squared: $r^2 = 0.9032$



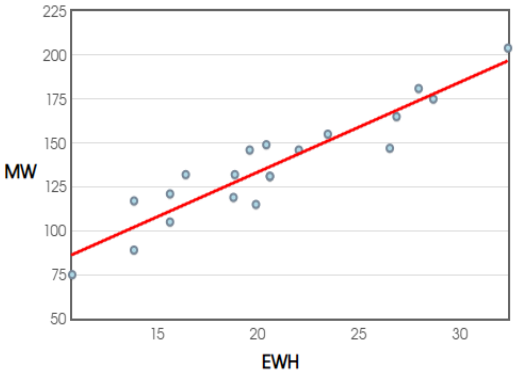
Regression Line: $\mathrm{M} = 3.2573 \cdot \mathrm{ACEWOH} + 21.0383$
Correlation: $r = 0.9583$
R-squared: $r^2 = 0.9183$



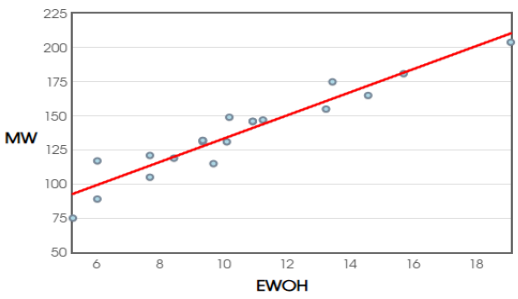
Regression Line: $\mathrm{ISA} = 1.2418 \cdot \mathrm{CEWH} - 1.2723$
Correlation: $r = 0.3747$
R-squared: $r^2 = 0.1404$



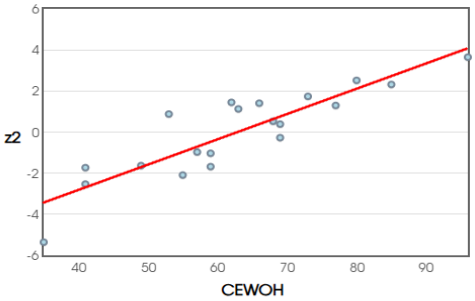
Regression Line: $\mathrm{ISA} = 1.4865 \cdot \mathrm{CEWOH} - 5.928$
Correlation: $r = 0.3947$
R-squared: $r^2 = 0.1558$



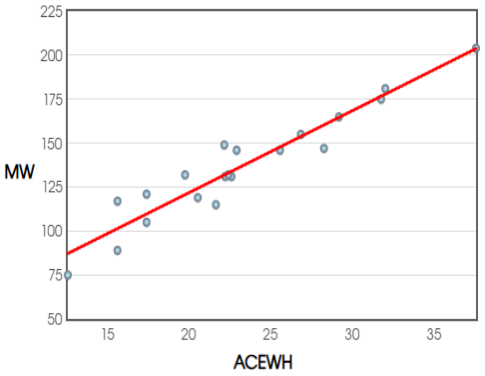
Regression Line: $\mathrm{MW} = 5.1325 \cdot \mathrm{EWH} + 30.8998$
Correlation: $r = 0.9333$
R-squared: $r^2 = 0.8711$



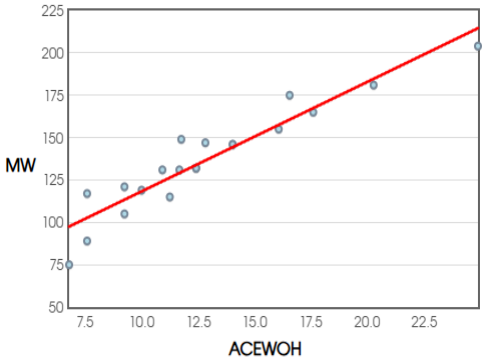
Regression Line: $\mathrm{MW} = 8.498 \cdot \mathrm{EWOH} + 48.3689$
Correlation: $r = 0.9518$
R-squared: $r^2 = 0.906$



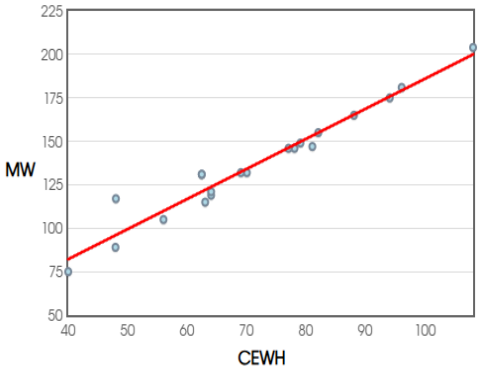
Regression Line: $\mathrm{Z2} = 0.1231 \cdot \mathrm{CEWOH} - 7.7385$
Correlation: $r = 0.8832$
R-squared: $r^2 = 0.78$



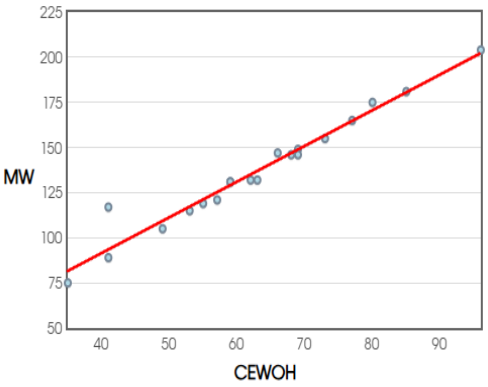
Regression Line: $\mathrm{MW} = 4.6705 \cdot \mathrm{ACEWH} + 28.4364$
Correlation: $r = 0.9527$
R-squared: $r^2 = 0.9077$



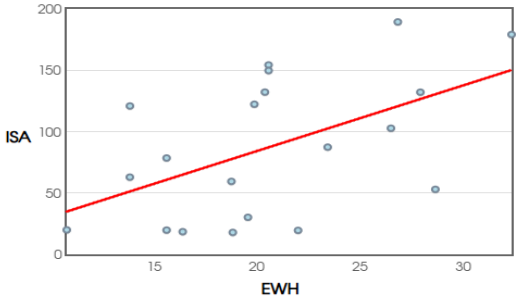
Regression Line: $\mathrm{MW} = 6.4683 \cdot \mathrm{ACEWOH} + 53.7139$
Correlation: $r = 0.9419$
R-squared: $r^2 = 0.8871$



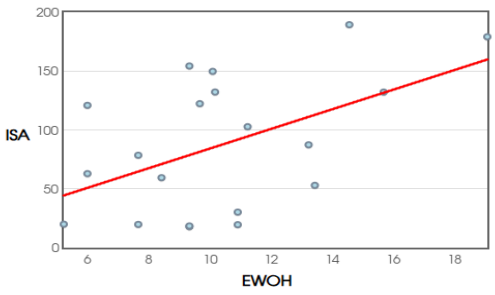
Regression Line: $\mathrm{MW} = 1.7328 \cdot \mathrm{CEWH} + 12.8763$
Correlation: $r = 0.9746$
R-squared: $r^2 = 0.9498$



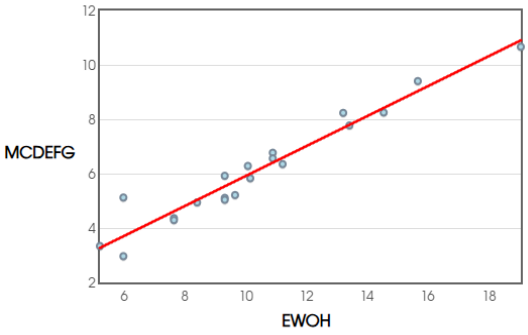
Regression Line: $\mathrm{MW} = 1.9779 \cdot \mathrm{CEWOH} + 12.4338$
Correlation: $r = 0.9789$
R-squared: $r^2 = 0.9583$



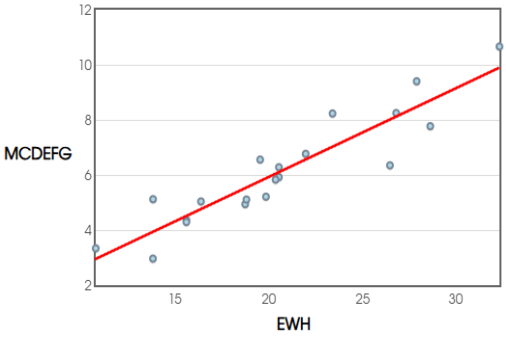
Regression Line: $\mathrm{ISA} = 5.3642 \cdot \mathrm{EWH} - 23.1259$
Correlation: $r = 0.5233$
R-squared: $r^2 = 0.2738$



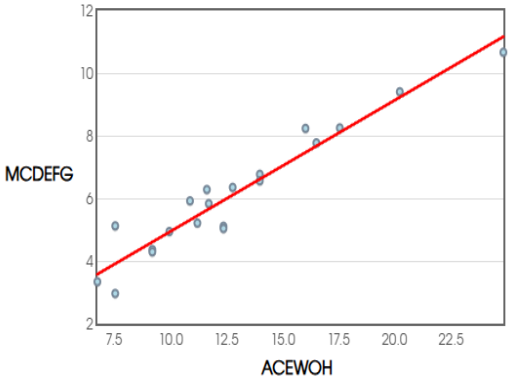
Regression Line: $\mathrm{ISA} = 8.3386 \cdot \mathrm{EWOH} + 0.779$
Correlation: $r = 0.501$
R-squared: $r^2 = 0.251$



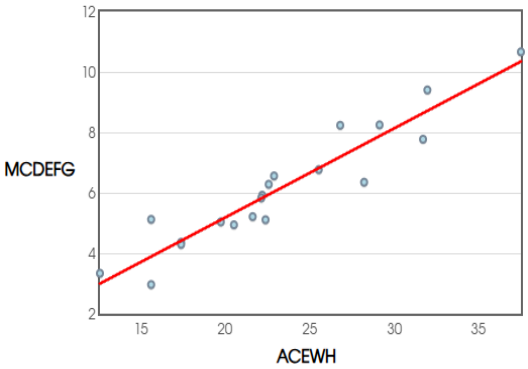
Regression Line: $\mathrm{MCDEFG} = 0.551 \cdot \mathrm{EWOH} + 0.3914$
Correlation: $r = 0.9693$
R-squared: $r^2 = 0.9395$



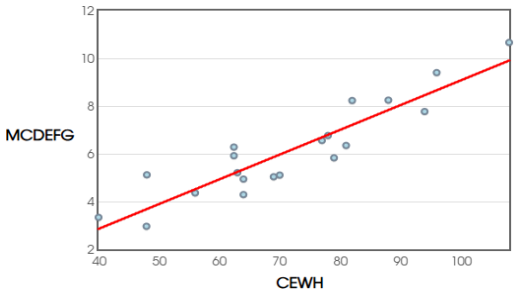
Regression Line: $\mathrm{MCDEFG} = 0.323 \cdot \mathrm{EWH} - 0.5401$
Correlation: $r = 0.9226$
R-squared: $r^2 = 0.8511$



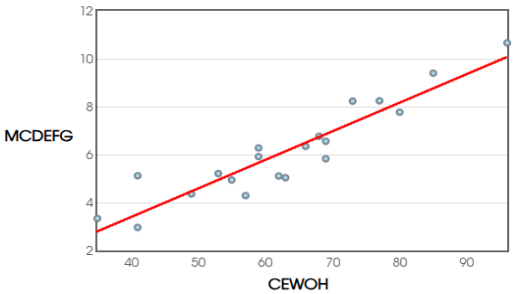
Regression Line: $\mathrm{MCDEFG} = 0.4192 \cdot \mathrm{ACEWOH} + 0.7399$
Correlation: $r = 0.9588$
R-squared: $r^2 = 0.9192$



Regression Line: $\mathrm{MCDEFG} = 0.2946 \cdot \mathrm{ACEWH} - 0.7111$
Correlation: $r = 0.9439$
R-squared: $r^2 = 0.891$



Regression Line: $\mathrm{MCDEFG} = 0.1037 \cdot \mathrm{CEWH} - 1.2921$
Correlation: $r = 0.9161$
R-squared: $r^2 = 0.8393$



Regression Line: $\mathrm{MCDEFG} = 0.1194 \cdot \mathrm{CEWOH} - 1.383$
Correlation: $r = 0.9282$
R-squared: $r^2 = 0.8615$

Figure 3: Regression models with various comparison of the data.

5. CONCLUSION

By using very simple regression models with the physical property and the energies of the graph the following observations are made.

- The properties Z_1, Z_3 and ECI cannot be predicted using any of the graph-derived energy metrics, as the obtained energy values are significantly larger than the values of the given properties.
- Z_2, MW, and M can be accurately predicted using various energy metrics derived from the graph. The regression model predicts Z_2 with up to 70% accuracy when compared to EWH and EWOH, and with up to 75% accuracy when compared to ACEWH and ACEWOH.
- The ISA can be reasonably predicted using EWH and EWOH, with predictions achieving up to 30% accuracy based on the regression model.
- This demonstrates that the Minimum Connected Dominating Energy of fuzzy graphs can be predicted with up to 95% accuracy using metrics such as energy, atom connectivity, or connectivity energy of the graph.
- The properties such as M and MW can be predicted with up to 95% accuracy using a regression model.

6. FUTURE WORK

- Various other types of energies like Domination, Distance, and Laplacian energy can be analyzed.
- Analyze whether the results obtained could be generalized of any molecule.
- Apply advanced regression models for better prediction of physical and chemical properties of the molecules.
- Analyze various other physical and chemical properties of the molecules.

7. ACKNOWLEDGEMENT

Funding: The research leading to these results has received Research Project Funding from The Research Council (TRC) of the Sultanate of Oman, under Commissioned Research Program, Agreement No.: MoHERI/BFP/UTAS/2023, Project No. BFP/URG/EBR/23/061, Project Title: Domination Energies of various Chemical Graphs: A detailed Study.

8. This research article is Compliance with Ethical Standards.
On behalf of all authors, the corresponding author states that there is no conflict of interest.
"No datasets were generated or analyzed during the current study"

9. REFERENCES

[1] Basak. S. C and V.R. 1988. Magnuson V. R, Determining structural similarity of chemicals using graph-theoretic indices, Discrete Applied Mathematics 19,17-44.

[2] B. Zhou. 2007. On the largest eigenvalues of the distance matrix of a tree, MATCH Commun. Math. Comput. Chem. 58, 657-662.

[3] C. A. Coulson. 1940 On the calculation of the energy in unsaturated hydrocarbon molecules, Proc. Cambridge Phil. Soc., 36, 201 – 203.

[4] Danail Bonchev and Douglas J. Klein. 2002. On the Wiener Number of Thorn Trees, Stars, Rings, and Rods, CROATICA CHEMICA ACTA, CCACAA 75 (2) 613-620.

[5] D.M.Cvetkovic, M.Doob, and H. Sachs. 1980. Spectra of graphs theory and applications, Academic press, New York.

[6] E. Sampath Kumar. 2010. Graphs and Matrices, Proceedings of National Workshop on Graph theory Applied to chemistry, chapter-2, Centre for Mathematical Sciences, Pala Campus, Kerala, 1-3 Feb.

[7] Elizabeth R, Collantes, William J. Dunn III. 1995. Amino Acid Side Chain Descriptors for Quantitative Structure-Activity Relationship Studies of Peptide Analogues, J. Med. Chem. 38, 2705-2713.

[8] Gutman and Boris Furtula. 2017, Survey of Graph Energies, Mathematics Interdisciplinary Research 2, 85 – 129.

[9] Gutman. 1978. The energy of a graph, Ber. Math. Statist. Sekt. Forschungszenturm Graz., 103, 1-22.

[10] Gutman, O. E. Polansky. 1986. Mathematical Concepts in Organic Chemistry, Springer.

[11] Gutman. 2000. The energy of a graph: old and new results, Algebraic Combinatorics and Applications, Springer, 196-211.

[12] Gutman. 2005. Topology and stability of conjugated hydrocarbons. The dependence of total ¼-electron energy on molecular topology, J. Serb. Chem. Soc., 70 441-456.

[13] G. Indulal, I.Gutman, A.Vijaya Kumar. 2008. On Distance Energy of a graph, MATCH Commun. Math. Comput. Chem. 60,461-472.

[14] Gutman. 2007. Sekt. Forchungsz, J.A. de la Pena and J. Rada, On the energy of regualr graphs, MATCH Commun. Math. Comput. Chem., 57, 351-361.

[15] H.B. Walikar. 2007. The energy of a graph: Bounds, Graph theory Lecture notes, Department of Computer Science, Karnatak University, Dharwad.

[16] H.S. Ramane, D.S. Revankar, Ivan Gutman, Siddani Bhaskara rao, B.D. Acharya, H.B. Walikar. 2008. Estimating the distance energy of graphs, Graph theory notes of New York LV.

[17] H.S. Ramane, H.B. Walikar, S.B Rao, B.D. Acharya, P.R. Hampiholi, S.R. Jog and Gutman. 2005. Spectra and energies of iterated line graphs of regular graphs, Appl. Math. Lett., 18, 679-682.

[18] Hellberg. S, Sjostrom.M, Skagerberg.B, Wold.S. 1987. Peptide Quantitative StructureActivity Relationships, a Multivariate Approach, J. Med. Chem. 30, 1126- 1135.

[19] Kamal Kumar. M. 2012. Domination energy of graphs, IJAM, Vol. 3, No. 1, 417-424.

[20] Kamal Kumar. M. 2011. Relation between domination number, energy of graph and rank. International Journal of Mathematics and Scientific Computing, 1(1), 58-61.

[21] Kamal Kumar. M. 2014. Characteristic polynomial and domination energy of some special class of graphs. International Journal of Mathematical and Combinatorics, 1(1), 37-48.

[22] Kamal Kumar. M, Johnson J. J, Winson. R. 2018. Different Domination energies in graphs, International Journal of Mathematical and Combinatorics, Vol.4, 103-145.

[23] Kamal Kumar. M, Shajidmon Kolamban. 2018. Various Domination energies in graphs, International Journal of Mathematical and Combinatorics, Vol. 3, No.12, 108-12.

[24] Kamal Kumar.M, Shajidmon Kolamban, Bushra Hibras Al Sulaimi. 2022. The Energy of Graph “A New Perspective”, Proceedings of International Conference(Virtual) on Interdisciplinary Research on Applied Science, Management and Technology (ICIRASMT), India, May 26-27.

[25] Kamal Kumar. M, Shajidmon Kolamban, Bushra Hibras Al Sulaimi. 2022. THE ENERGY OF GRAPH: A NEW PERSPECTIVE, Global Journal of Modeling and Intelligent Computing (GJMIC), Vol. 2, No. 2.

[26] Mathavi Manisekar, S. Lalitha A. 2015. Comparison between the Adjacency Matrix Energy and Dissociation Energy for Amino Acids, International Journal of Engineering Technology, Management and Applied Sciences, August, Volume 3, Issue 8.

[27] Neeta M. Chougule, D.B.Unde. 2024. The Minimum connected dominating energy of fuzzy graphs and its application in structural characterization of amino acids,DCDIS A,Mathematical Analysis, Preprint.

[28] Spialter. L. 1963. The atom connectivity matrix (ACM) and its characteristic polynomial (ACMCP): A new computer-oriented chemical nomenclature, J. Am. Chem. Soc. 85 2012-2013.

[29] Spialte. L. 1964. The Atom Connectivity Matrix (ACM) and Its Characteristic Polynomial (ACMCP), J. Chem. Doc. 4, pp. 261-269.

[30] Spialter. L. 1964. The Atom Connectivity Matrix Characteristic Polynomial (ACMCP) and Its Physico-Geometeric (Topological) Significance, J. Chem. Doc. 4, pp. 269-274.

[31] Tamara Anthony Carter, Richard A. Tapia, Anne Papakonstantinou. 1995. An Introduction to Linear Algebra for Pre-Calculus Students, CEEE.

[32] V. Nikiforov. 2007. The energy of graphs and matrices. J.Math. Anal. Appl., 326, 1472-75.

[33] Zhao Li , Jijun Tang , Fei Guo. 2016. Identification of 14-3-3 Proteins Phosphopeptide-Binding Specificity Using an Affinity-Based Computational Approach. PLoS One. Feb 1;11(2): e0147467. doi: 10.1371/journal.pone.0147467.