Iraqi National Journal of Chemistry

Journal homepage:
http://iqnjc.com/Default.aspx

Dielectric Constant (ε) for Pentane substituted, Using Quantitative Structure-Property Relationship (QSPR) Techniques.

[1]Kawkab A. Hussain, [2]Wisam A. Radhi, [1]Sadiq M-H. Ismael,

[1]Department of Chemistry – College of Education Pure Science

[2]Department of Chemistry, Polymer Research Center, University of Basrah

University of Basrah – Iraq

Abstract

Quantitative Structure–Property Relationship (QSPR) models based on molecular descriptors derived from molecular structures have been used for the prediction for computed the dielectric constant (ε) of pentane substituted. QSPR model includes some Molecular descriptors, regression quality indicates that these descriptors provide valuable information and have significant role in the assessment of the dielectric constant of compounds understudy. four QSPR equation for the prediction of dielectric constant have been drawn up by using the multiple regression technique. (Eqs 1-4) with the values of R^2 range from 0.862-0.969, Ra^2 range from 0.702-0.923, Q^2 range from 0.86-0.93 and the values of S range from 0.989-0.488, while the values of F range from 75.127-172.796. The results show excellent model by Eq 3. with high of R^2, Ra^2, F and minimum S by using five parameters [D.M and T.E], was found and indicate that these parameters have important role in determining the properties of dielectric constant . this result encourages the application of QSPR to a wider selection of compounds properties and to other classes of compounds, including industrial, biopolymers.

Keywords. Pentane substituted, predicted dielectric constant (ε), (QSPR) Model.

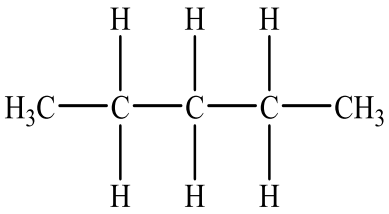
Introduction

Dielectric constant **(ε)**, which is the ratio of the actual electric displacement to the electric field strength when an external field is applied to the substance [1]. The dielectric constant **(ε)** is the most important it has a significant effect on the properties and processing characteristics of the compounds, the experiments are costly and time consuming, nowadays much interest is devoted to the prediction of physicochemical properties of molecules, such as their biological activity, chemical property, their retention on chromatographic systems, or electrochemical property,etc[2-6]. The quantitative structure-activity/property relationship (QSAR/QSPR) is a successful strategy for prediction of surfactant properties based on modeling between calculated descriptors from molecular structures of the surfactants and chemical or physical properties of the solution. QSPR has also become a well-established and proven technique to correlate diverse physicochemical properties of compounds, ranging from simple to complex, with molecular structure, through a variety of descriptors of the chemical structures. Most QSAR/QSPR treatments utilize a program to calculate descriptors and then try to select a small number of descriptors in a purely empirical fashion to form an equation. This is derived from a so-called “training set” of compounds for which a property of interest has been measured. QSPR methodology has been aided by new software tools, which allow chemists to elucidate and to understand how molecular structure influences properties [7-14]. In this work we demonstrate the usefulness some of the parameters in deriving predictive QSPR models. The relation between the dielectric constant **(ε)** and quantum chemical calculation parameters to know the physicochemical behavior of pentane substituted compounds, and to find out the effect of various the structural, chemical, physical and other properties of these compounds understudy on experimental dielectric constant **(ε)** .

Modeling and Geometry Optimization

The quantum chemical calculations were performed for 14 compounds understudy with the Gaussain [15]. Geometrical optimizations were carried out using Density Functional Theory (DFT) method at B3LYB/6-31G(2d) level of theory[16]. The experimental dielectric constant (**ε**) data of 14 pentane substituted compounds under study has been taken from reference[1]. Structures of 14 pentane substituted compounds shown in Figure.1 .

No	R
1	Pentane
2	2-Methylpentane
3	3-Methylpentane
4	1-Bromo-2-Ethoxypentane
5	2-Bromo-3-Ethoxypentane
6	3-Bromo-2-Ethoxypentane
7	3-Bromopentane
8	2,2-Dimethylpentane
9	2,3-Dimetheylepentane
10	2,4-Dimethylpentane
11	3,3-Dimethylpentane
12	1-Ethoxypentane
13	1-Fluoropentane
14	1-Bromopentane



Pentane

Figure 1. Molecular structure of compounds used in the present study

Results and Discussion

The relationship between dielectric constant (**ε**) and various descriptors (Physiochemical and alignment-independent) were established by sequential multiple regression analysis (MLR) in order to obtain QSPR models. The best multilinear regression (BMLR) procedure was used to find the best correlation models from the selected noncollinear descriptors. The descriptors Table 1., which were significant for experimental data, were selected by QSPR-contingency module. To establish the statistical correlation, the physicochemical parameters were taken as independent variables and dielectric constant (ε) as dependent variable. The best model was selected on the basis of statistical parameters viz observed with high coefficient of multiple (R^2), adjusted coefficient of multiple determination (Ra^2), cross-validation correlation coefficient (Q^2), sequential Fischer test (F) and low standard error of estimate (S) . were employed to judge the validity of regression equation and evaluate the obtained QSPR models[17&18].

In the first the prediction model of QSPR study has been make up with assist of the next descriptors HOMO energy, LUMO energy, Dipole moment(μ), Mass, Polarizability, Total Energy, Hydration Energy, refractivety,Table 2 , can be directly related with experimental data of dielectric constant(ε). The 1 and 8- descriptor correlations of the dielectric constant (ε) were given in eqs (1-4) respectively and

the resulting parametric models are depicted in figures. 2-5, along with statistical

No	LUMO	HOMO	D.M	T.E	Mass	pol	Ref	H.E
1	2.4256	-8.5362	0.0835	-197.778	72.15	9.95	24.81	3.16
2	2.3595	-8.3019	0.1071	-237.081	86.18	11.78	29.36	3.36
3	2.1774	-8.3365	0.0953	-237.091	86.18	11.78	29.36	3.31
4	0.0291	-6.8875	2.1031	-2922.78	195.1	16.88	43.63	1.63
5	-0.1001	-6.8646	1.2524	-2922.78	195.1	16.88	43.45	1.49
6	0.0829	-6.9999	3.0544	-2922.78	195.1	16.88	43.45	1.54
7	-0.0076	-7.3052	2.2904	-2768.94	151.05	12.58	32.51	2.66
8	1.9301	-8.2502	0.0619	-276.406	100.2	13.62	33.83	3.51
9	2.7668	-8.2086	0.0859	-276.403	100.2	13.62	33.9	3.47
10	2.052	-8.198	0.1877	-276.403	100.2	13.62	33.9	3.44
11	1.9088	-8.2121	0.0913	-276.403	100.2	13.62	33.83	3.41
12	2.3265	-6.8881	0.9726	-351.621	116.2	14.26	36.23	1.72
13	2.1586	-8.6573	2.0476	-297.016	90.14	9.86	25.04	2.87
14	-1.5799	-6.2121	2.0051	-2763.79	149.03	12.38	32.5	1.18

parameters of the regression.

Table 2. Descriptors as the independent variables used for QSPR analysis of compounds.

Definition *of Descriptors Used in This Study*.

ΔE= Energy. GAP =Different between HOMO and LUMO is energy gaps in eV, LUMO= The energy of Lowest Unoccupied Molecular Orbital in eV, HOMO= The energy of Highest Occupied Molecular Orbital in eV, D.M= Dipole moment in debyes., H.E= Hydration Energy in Kcal/mol, Pol=Polarizability, T.E=Total Energy in a.u., (**ε**)= dielectric constant.

The prediction set, consisted of 14 molecules, was used to evaluate the generated model. It is clear that many MLR models will result using stepwise multiple regression procedure; among them we have to choose the best one. It is common to consider four statistical parameters for this purpose. These parameters are the number of descriptors, coefficient of multiple (R^2) for training and prediction sets, standard error (SE) for training and prediction sets, (F)statistic, adjusted coefficient of multiple determination (Ra^2) and The cross-validation correlation coefficient (Q^2). A reliable MLR model is one that has high R^2, Ra^2, $Q^2$and F values, low SE and least number of descriptors. In addition to these, the model should have a high predictive ability. Establish the statistical correlation, the physicochemical parameters were taken as independent variables and dielectric constant (ε) as dependent variable. The solutions of the above multiple linear regressions (MLR) are given by equations 1 to 4, calculated to generated equations 1 to 4.

The first model when depend on only one parameter [D.M] gave model with coefficient of the multiple R^2 values for this model of 0.862, as equation 1. The suggest that the dielectric constant (ε) increases with increase values of this descriptor.

ε = 2.219D.M+1.754……..Eq1

Statistical characteristics of the obtained equation :

R^2 = 0.862 F = 75.127 S = 0.989 Ra^2= 0.702 Q^2=0.86

The relationship between the experimental data and predicted dielectric constant for pentane substituted as shown in Fig.2**.**

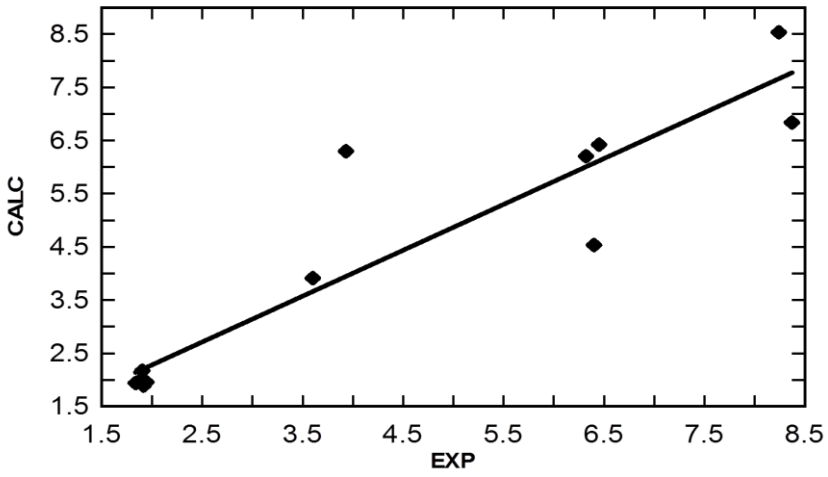


Figure 2. Plot of (ε) prediction versus (ε) experimental using Eq 1.

Second model. of the dielectric constant (ε) of pentane compounds depends on T.E value. Second model depends on only 1 parameter gave good model with change in the coefficient of multiple R^2 values to 0.888.

ε = -1.871E-03T.E+1.806….Eq2

Statistical characteristics of the obtained equation :

R^2= 0.888 F =95.724 S = 0.889 Ra^2= 0.879 Q^2=0.88

Fig 3. Show the relationship between the experimental data and predicted dielectric constant **(ε)**.

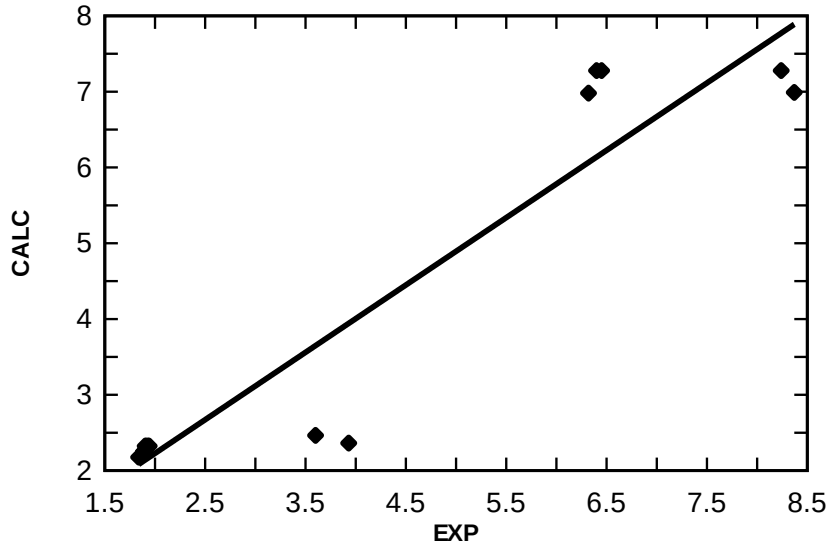


Figure 3. Plot of (ε) prediction versus (ε) experimental using Eq 2

.

Two- parameter correlations of the pentane compounds were given in eq 3. In this equation it could be seen that increases of the coefficient of multiple when depends on 2 parameters [T.E and D.M] by collection the eq1 and eq2 to generate the eq3.

ε = 1.542-1.005E-03T.E+1.150 D.M……Eq3

Statistical characteristics of the obtained equation :

R^2 = 0.969 F = 172.796 S = 0.488 Ra^2= 0.963 Q^2=0.968

From eq.3 the coefficient 1.542 is intercept, for every unit increase in T.E , we expect a 1.7826 E-04 unit decrease in the dielectric constant every unit increase in D.M , a 1.15 unit increase in dielectric constant is predicted, staying all other variable constant. The large value of F-statistic interpretation large part from data in model 3(eq 3) and this means random differences is few, Also the difference between the R^2 and R^2 adj value is less than 0.3 indicates that the number of descriptors involved in the QSAR model is acceptable. The number of descriptors is not acceptable if the difference is more than 0.3(19)**.**

The relationship between the experimental data and predicted dielectric constant (**ε**), are given in Fig.4**.**

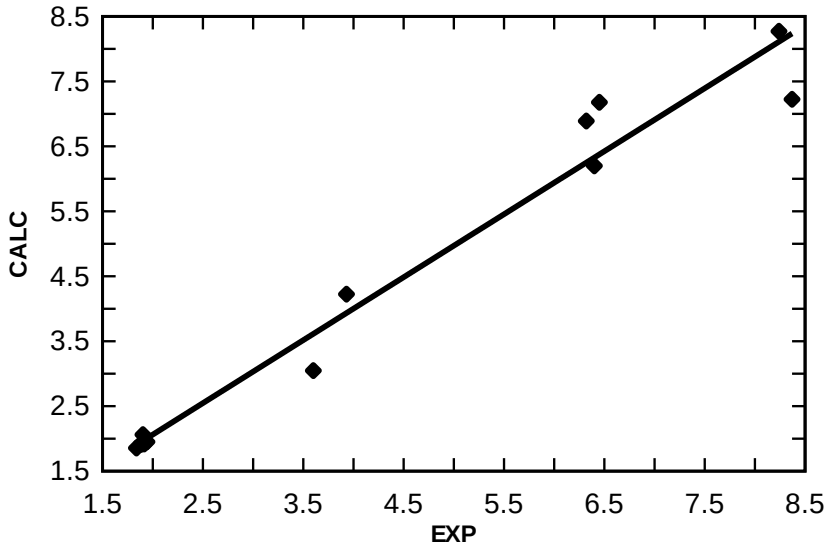


Figure 4. Plot of (ε) prediction versus (ε) experimental using Eq 3.

When replacement of a parameter [T.E] in eq 3.by the parameters [Mass], gave model predicted in this study Eq 4. This model equation depends on the 2 parameters[D.M, Mass], The resulting decreasing coefficient of multiple & cross-validation correlation coefficient, *F*-test statistic and increase standard error which means the mass parameter have weak statistic affect on dielectric constant compared with T.E in eq 4.

ε = -0.434+2.386E-02Mass+1.471D.M……Eq4

Statistical characteristics of the obtained equation

R^2 = 0.934 F = 79.013 S = 0.710 Ra^2=0.923 Q^2=0.93

From eq 4 the coefficient -0.434 is intercept, for every 2.386 E-02 unit increase in mass , we expect a 2.386 E-02 unit increase in the dielectric constant and every unit increase in D.M , a 1.15 unit increase in dielectric constant is predicted, staying all other variable constant.

Fig 5. Represents the relationship between the experimental data and predicted dielectric constant (**ε**)**.**

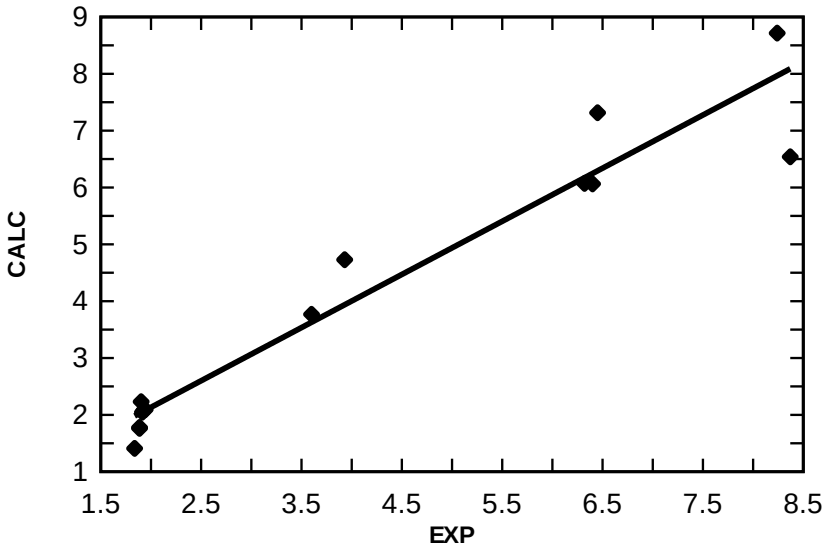


Figure 5. Plot of (ε) prediction versus (ε) experimental using Eq 4.

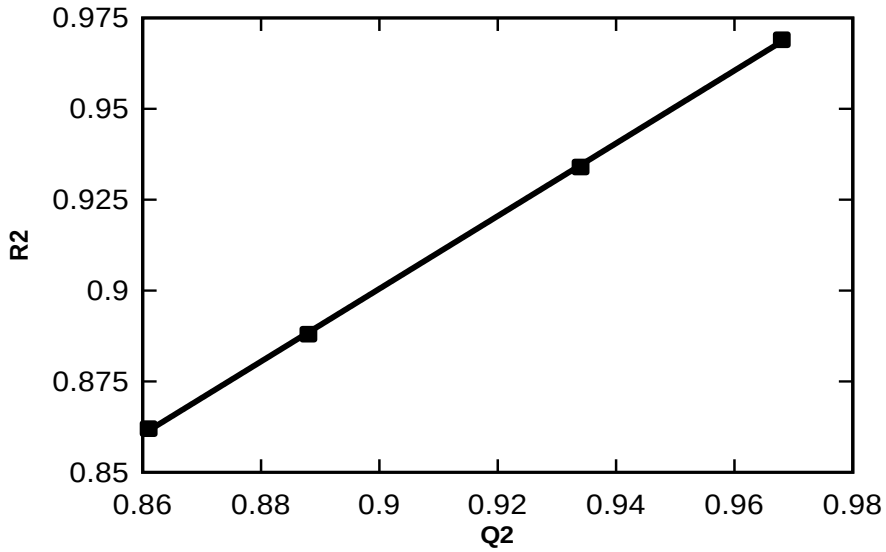


Fig 6. The Correlation between Q^2 vs R^2Obtained by Eq 1-4.

The quality of models can be evaluate by correlation coefficient squared (R^2), coefficient of determination, or in the best, by external validation. Cross validation (Q^2) is used to judge the productivity of the model, if no data remain for external validation. The correlation coefficient of regression between experimental and data estimated by cross validation is cross- validated correlation coefficient Q^2. The QSPR model is "good" if $Q^2 > 0.5$, excellent if $Q^2 > 0.9$. In this study the final model results from training on the entire dataset using the signature height and descriptor count assessed to have the highest predictive accuracy as measured by the (Q^2). A representative plot showing Q^2 as a function of the correlation coefficient R^2 count as illustrated in figure 6, for the predicted dielectric constant (**ε**) dataset.The fact can be well established from the figure 6 showing a comparative plot of the values of Q^2, R^2 pred for the 4 different equations models that all the validations for the real MLR

model confirm the self-consistency, robustness and good prediction power of the model, its stability to resampling's and the absence of chance correlation. The real MLR model shows excellent performance in predictive of predicted dielectric constant (**ε**), based on the high values for Q^2 and R^2 which are confident that the QSPR model gives good predictions predicted dielectric constant (**ε**) of compounds that may be used to understanding better for this type of compounds(20-21) .

From the residual values (Figs. **7** and 8), it can be clearly seen that the lower residual values show that there is a minimal difference between the experimental value and the predicted value of the dielectric constant (**ε**) of this test set, also the behavior of the residuals in terms of the predictions follows a normal distribution. No molecule in the set exhibits a residual larger than 2S that can be considered as an outlier(22-23).

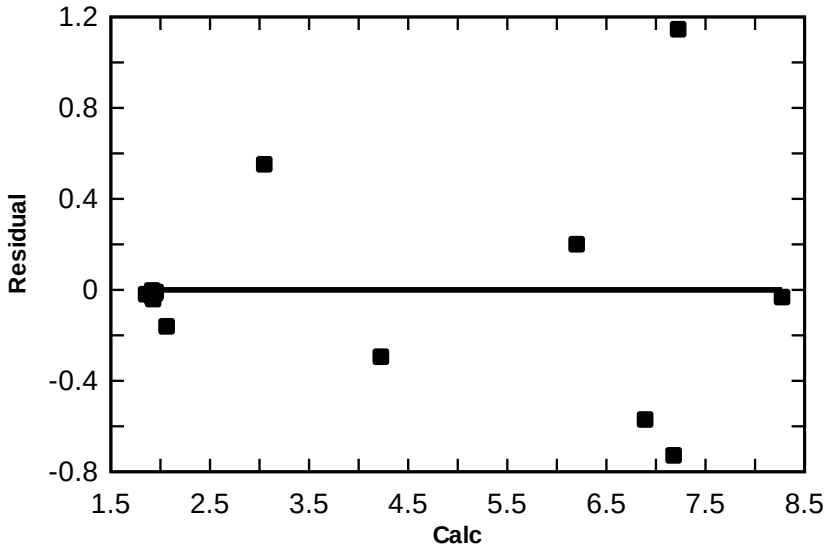


Fig. (7). Histogram of residual values obtained from QSPR model for equation 3, depends on D.M and T.E, R^2= 0.969.

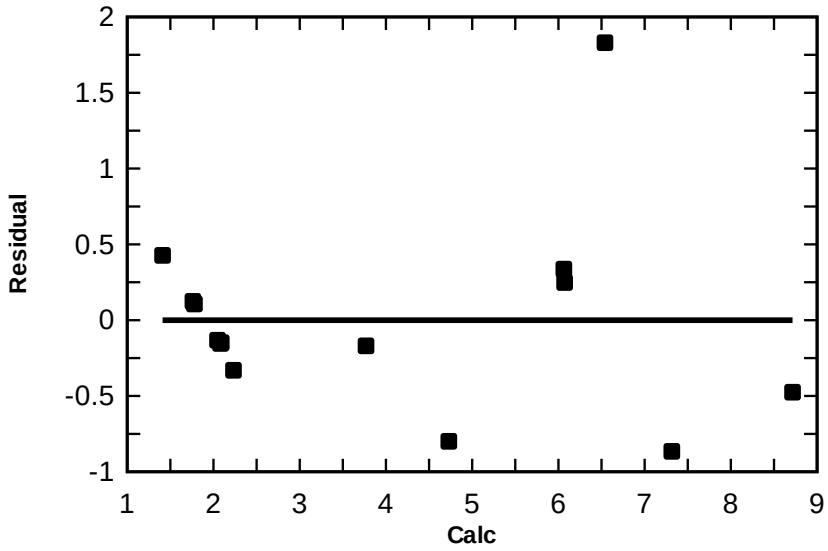


Fig. (8). Histogram of residual values obtained from QSPR model for equation 4depends on Mass and D.M, R^2 = 0.934.

Consequently, among different models, the best model was chosen, whose specifications are presented in Table 3. It is obvious that as the number of descriptors increase the R^2 will increase. Also shows the effect of increasing the number of descriptors on R^2 values. It can be seen from this table that increasing the number of parameters only up to 2 has a large influence on improving correlation 1-4- parameter models for each of compound understudy, which it has high correlation coefficient & adjusted coefficient of multiple determination (Ra^2), Cross validation (Q^2), less standard error (SE) and high F values.

Table 3. Statistical parameters of the lineal regressions models obtained for the 6 kinds of descriptors

No	Descriptors	R^2	F	S	Ra^2	Q^2
1	D.M	0.862	75.127	0.989	0.702	0.861
2	T.E	0.888	95.724	0.889	0.879	0.888
3	D.M + T.E	0.969	172.796	0.488	0.963	0.968
4	D.M + Mass	0.934	79.013	0.710	0.923	0.934

It could be seen from Table 4. The predicted of the dielectric constant **(ε)** values obtain from Eq. 1-4 in this study and comparable with the experimental values in the Reference [1]. It is obvious from this table 4. that the relations between descriptors which calculations in this study and experimental the dielectric constant **(ε)**values are excellent.

Table 4. Predicated Experimental data depends on Eq 1. & Eq 4.

No	(ε) Exp	Calc By Eq1	Calc By Eq2	Calc By Eq3	Calc By Eq4
		R^2=0.88	R^2=0.86	R^2=0.934	R^2=0.969
1	1.837	2.176	1.939	1.409	1.856
2	1.886	2.25	1.991	1.779	1.926
3	1.886	2.25	1.965	1.761	1.912
4	6.45	7.278	6.422	7.315	7.177
5	6.4	7.278	4.534	6.063	6.199
6	8.24	7.278	8.534	8.715	8.271
7	8.37	6.99	6.838	6.539	7.223
8	1.915	2.324	1.891	2.047	1.917
9	1.929	2.324	1.944	2.082	1.945
10	1.902	2.324	2.170	2.232	2.062
11	1.942	2.324	1.956	2.090	1.951
12	3.6	2.464	3.913	3.769	3.047
13	3.93	2.362	6.299	4.729	4.224
14	6.32	6.980	6.204	6.071	6.890

Conclusion

Quantum chemical calculated parameters can be successfully used for the derived a designer QSPR capable of predicting the dielectric constant (ε) values. The study indicated that predicted dielectric constant (ε) values for pentane compounds can be modeled. The values of R^2 for the QSPR models Eqs. 1-4 range from 0.862-0.969, the values of Ra^2 range from 0.963-0.702, Q^2 range from 0.86-0.968 and the values

of S in the Eqs. 1-4 range from 0.969-0.488, while the values of F range from 172.127-79.013which are statistically significant at the 99% level. The values of R^2, Ra^2 ,F Q^2,and S suggest that the QSPR models Eqs. 1-4 are predictive and validate. From all the results the Eq 3 that contain T.E and D.M descriptors have smaller the value of S and the larger the value of F ,R^2,Ra^2 and Q^2, which indicate predictive power for eq 3. On the other hand the T.E and D.M descriptors play an important role in effect on dielectric constant (ε) properties of compounds, which allow chemists to elucidate and to understand how molecular structure influences properties.

References:

[1] John A. D., ***LANGE'S HANDBOOK OF CHEMISTRY***., Fifteenth Edition,. Copyright 1999,by McGraw-Hill, Inc.

[2]Sameer D, Gagan D A., ***International Journal of Engineering and Advanced Technology (IJEAT).*** Volume-3, Issue-5, June 2014.

[3] Inaan M. A. , Harith I. J., Huda K. A., and Sewench N. R., ***The Iraqi Journal For Mechanical And Material Engineering,*** Vol.12, No.1, 2012.

[4] Farid. M. T., Ahmad. I, Aman. S., Kanwal. M, Murtazai. G., Ali, I. Ahmad, and M. Ishfaq., ***Journal of Ovonic Research*** .,Vol. 11, No. 1, January - February, 2015.

[5]Rattikorn Y., Supon A. and Pitak L., *Songklanakarin* ***J. Sci. Technol****.* Vol. 26 No. 4 Jul.- Aug.2004.

[6]Bhupesh N and Pothaji K., ***International Journal of Physical and Mathematical Sciences*** ,Vol 5,No 1 ,2015.

[7] Alan R. K, Dimitar A. D and Mati K, ***A Journal of Chemical Science***. Z. Naturforsch,61b, 373 –384.2006. http://znaturforsch.com.

[8] Ashrafi F, Rostami A. A and Mahdavipour N.***African Journal of Pure andApplied Chemistry*** ,Vol. 2 (1), pp. 006-009,2008*.*

[9] Kawkab A. H Al-A, Methem N. A, Sadiq M-H. I, ***Journal of Basrah Researches(Sciences),*** *37. 4A* / 111-115.2011

[10]Bahjat A. S, Rita S. E,Sadiq M-H.I and Kawkab A.H. ***American Journal of Applied Sciences***, 8 (8), 773-776.2011.

[11]Kawkab A. H, Sadiq M-H. I and Wisam A-H. R, ***J. Chem. Pharm. Res***.4(3):1702-1707.2012,

[12] Vladyslav K , Jack R. S, Doyle K, Sascha A, Joachim K, William J. W,. ***Polymer*** **,**45, 7367– 7379,2004.

[13]Aihong L, Xueye W, Ling W, Hanlu W, Hengliang W., ***European Polymer Journal***. 43,989– 995,2007.

[14] Sadiq M-H I, Kawkab A. H and Abbas D. M, ***Chem Sci Trans.,*** 4(2), pp 355-360

DOI:10.7598/cst2015, 2015.

[15]Frisch M. J., Trucks ,G. W., Schlegel , et al.,. *Gaussian Inc., Waling ford, CT,* . 2004. Available on line . www .**gaussian**.com.

[16] Lee.C., Yang., W and Parr ., R. G., ***Phys. Rev. Lett***., B37, 785-789, 1998.

[17] Rudolf Kiralj and Márcia M. C. Ferreira., ***J. Braz. Chem. Soc***., vol. 20, No. 4, 770-787,

2009.

[18] S. P. Cardoso, J. A. C. P. Gomes, L. E. P. Borges and E. Hollauer., ***Brazilian Journal of Chemical Engineering.,*** Vol. 24, No. 04, pp. 547 - 559, October - December, 2007.

[19] Ravichandran V., Harish R., Abhishek J., Shalini S., Christapher P. V., and Ram K. A., ***Interantional Journal of Drug Design and Discovery***., 2., 3.,2011.

[20] Jord.i P-F., María. G-L., Jorge. G., Ramón. G-D., ***Int. J. Mol. Sci.*** 2011,12, 12811292; doi:10.3390/ijms12021281.

[21] Rudolf. K., Márcia. M. C. F.,***J. Braz. Chem. Soc***.2009,20, 4,770-787.

[22] Liza T. B., and Junie B. B., ***Journal of Chemical and pharmaceutical Research***, 6(8),2014.

[23] Andrew G. M., Pablo R. D., Francisco M. F., Eduardo A. C., Ezequiel W., ***Chemical Physics Letters***,462 , 2008.

Iraqi National Journal of Chemistry 2015; 15 (3)

التنبأ بخواص ثابت العزل الكهربائي للبنتان المعوض بأستخدام تقنية العلاقة التركيب – الخاصية الكمية

كوكب علي حسين ، وسام عبد الحسن راضي، صادق محمد حسن اسماعيل ،

جامعة البصرة –كلية التربية للعلوم الصرفة- قسم الكيمياء

الخلاصة:

تقدم الدراسة موديلات العلاقة التركيبية الخصائصية الكمية QSPR اعتمادا على حساب الموصوفات الجزيئية المشتقة من التركيب الجزيئي للبنتان ومعوضاته للتنبؤ وحساب ثابت العزل الكهربائي للبنتان ومعوضاته .تضمنت تقنيات QSPR بعض الموصوفات الجزيئية والتي اثبتت بان هذه الموصوفات تعطي معلومات وقيم مهمة ولها دور كبير في تحديد خواص ثابت العزل الكهربائي للمركبات المدروسة.بأستخدام اربعة موديلات QSPR (eqs 1-4) وكانت الدوال الكمية المستحصلة من المعادلات ضمن المدى

R^2 =0.969, Q^2 =0.86-0.93, Ra^2 = 0.963, S = 0.488, F = 172.796.

اظهرت الدراسة بأن افضل موديل ممتاز كان باستخدام معادلة eq 3. اعتمادا على القيم العالية ل R^2 ,Q^2, Ra^2, F واقل القيم ل S والذي اعتمد على المتغيرات الجزيئية التالية

[D.M, T.E], والتي اظهرت بأن لها تأثير ودور كبير في تحديد خواص ثابت العزل الكهربائي للمركبات المدروسة. ومن هذا الموديل تبين الدراسة التأثير المميز لهذه الموصوفات على ثابت العزل الكهربائي لمركبات البنتان المدروسة وتشجع هذه النتائج لدراسة وتطبيق تقنيات QSPR لمدى اوسع من المركبات