QSPR Study of Some Physicochemical Properties of Sulfonamides Using Topological and Quantum Chemical Indices

Fatemeh. Shafiei* and Abolfazl. Saeidifar
Department of Chemistry, Arak Branch, Islamic Azad University, Arak, Iran,
P.O. Box 38135-567, Arak 38361-1-9131 Iran
f-shafiei@iau-arak.ac.ir; shafa38@yahoo.com

(Received on 6th April 2016, accepted in revised form 2nd Februaru 2017)

Summary: QSPR studies on sulfonamides have been made using recently introduced topological methodology. In this study the relationship between the Randic' (^1X), Balaban (J), Szeged (Sz), Harary (H), Wiener (W), Hyper-Wiener (WW), Wiener Polarity (WP) and one other descriptor, namely, the LUMO energy (E_{lumo}) to the thermal energy (E_{th} kJ/mol), heat capacity (C_V J/molK) and entropy (S J/molK) of 41 sulfonamides is represented. Physicochemical properties and the quantum chemical parameter are taken from the quantum mechanics methodology with HF level using the ab initio 6-31G basis sets. The multiple linear regressions (MLR) and Back ward methods (with significant at the 0.05 level) were employed to give the QSPR models. The satisfactory obtained results show that combining the three descriptors (J, E_{lumo}, H) are excellent descriptors for predicted (C_V)and(S), the three descriptors (^1X, J, E_{lumo}) are useful descriptors for predicted (E_{th}) of the 41 sulfonamides.

Keywords: QSPR, Topological indices, Sulfonamides, Graph theory, Multiple linear regressions (MLR).

Introduction

Quantitative structure–property relationships (QSPRs) have provided a valuable approach in research into physico-chemical properties of organic chemicals [1]. Many investigators have used quantum – chemical parameters [2-6]. Among the different approaches employing computational chemistry, those based on chemical graph theory have been useful in establishing QSPR [7].

The basic strategy of QSPR is to find the optimum quantitative relationship which can then be used for the prediction of the properties of molecular structures including those unmeasured or even unknown [8-10]. The premise of QSPR is that physicochemical properties can be correlated with molecular structure characteristics (geometric and electronic) expressed in terms of appropriate molecular descriptors [11].

QSPR have been traditionally developed by selecting, a priori, an analytical model (typically) linear, polynomial or lag- linear to quantity the correlation between selected molecular indices and desired physicochemical properties, followed by regression analysis to determine model parameters [12-15].

Sulfonamides represent an important class of biologically active compounds. With the sulfanilamides as the lead structure, different classes of pharmacological agents have been obtained such as antibacterial sulfanamides, sulfonamides that inhibit the zinc enzyme carbonic anhydrase (CA's). The ring substituted benzene sulfonamides containing -SO2NH2 groups have similar activities.

The hypoglycemic sulfonamides extensively used in the treatment of some forms of diabetes, antithyroid drugs, and others [16].

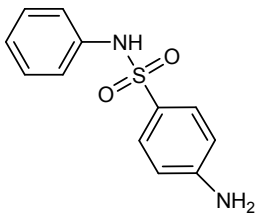
In the present study, the multiple linear regression (MLR) techniques and back ward methods are used for modeling the thermal energy (E_{th} kJ/mol), heat capacity (C_V J/molK) and entropy (S J/molK) of 41 sulfonamides.

This method is useful when there is not any interaction between descriptors and their relation with linear defined activity. Heat capacities are applied in reactions for modification of reactants evaluation. In addition, they are useful for heat - energy balance design calculation. On the other hand, the tests for determining the heat capacity, entropy are expensive and expense much more time. Therefore, we need the models to predict the heat capacity and other physico-chemical properties of molecules.

The proposed QSPR models were based on molecular descriptors (topological indices) that can be calculated for any compound utilizing only the knowledge of its molecular structure (molecular graph).

Experimental

The sulfonamides discussed in this study consist of 41 derivatives with substitution at 2-, 3- and 4- position as well as having some di-substitution on the aromatic nucleus (the N phenyl group). Fig 1 shows the template structure of sulfonamides used in the present study.



Scheme-1: The structural template of sulfonamides.

Quantum Chemistry Calculations

The thermal energy (E_{th}), heat capacity (C_V), entropy (S) and LUMO energy(E_{lumo}) of 41 sulfonamides are taken from the quantum mechanics methodology with Hartree- Fock (HF) level using the ab initio 6-31G basis sets and the standard procedure in GUSSIAN 03. This software is an electronic structure package capable of predicting many properties of atoms, molecules, and reactive systems. The quantum chemistry data of the 41 congeners are listed in Table-1.

Topological Indices

Nowadays, in the literature, hundreds of topological indices, suitable to describe different properties, are reported. The topological indices(Tis) used for the QSPR analysis were Wiener (W)[17], Szeged (Sz) [18], first order molecular connectivity(^{1}X) [19], Balaban(J) [20], Hyper-Wiener(WW) [21], Wiener Polarity(WP)[22] and Harary(H) [23] indices. Moreover, many investigations were carried out with such descriptors [24-26].

All the used topological indices were calculated using all hydrogen suppressed graph by deleting all the carbon hydrogen as well as heteroatomic hydrogen bonds from the structure of the sulfonamides. The descriptors were calculated with chemicalize program [27]. Seven topological indices tested in the present study are recorded in Table-2.

Statistical Analysis

Structure- Property models (MLR models) are generated using the multi linear regression procedure of SPSS version 16. The thermal energy(E_{th} kJ/mol), heat capacity (C_V J/molK) and entropy(S J/molK) are used as the dependent variable and ^{1}X, J, Sz, H, WP, WW and LUMO energy(E_{lumo}) indices as the independent variables. The models are assessed with R value (correlation coefficient), the R^2 (coefficient of determination), the R^2- adjusted, the RMSE value (root of the mean square of errors), the F value (Fischer statistic), the DW value (Durbin-Watson) and the Sig (Significant).

Table-1: Structural details and their thermal energy (E_{th}), heat capacity (C_V) and entropy (S) for the sulfonamides used in present study

Substituents	No.	E_{th} kJ/mol	Cv J/molK	S J/molK	Substituents	No.	E_{th} kJ/mol	Cv J/molK	S J/molK
4-NMe_2	1	886.871	292.805	609.210	4-CN	22	670.995	256.397	555.246
2-OMe	2	768.982	268.306	571.920	4-$COCH_3$	23	784.933	280.779	589.062
2- OC_2H_5	3	852.131	287.137	601.845	4-NO_2	24	683.455	265.656	570.980
4-OMe	4	767.603	269.238	572.359	4-SO_2CH_3	25	783.198	310.896	631.460
4- OC_2H_5	5	851.809	286.539	623.510	2,3-di-Me	26	833.442	278.835	576.853
4-Me	6	751.108	256.021	566.244	2-Me,5-Cl	27	727.709	272.298	578.838
2-Me	7	751.501	255.875	548.320	2-Me,6-Cl	28	728.265	270.275	569.053
3-Me	8	751.087	256.213	567.088	3,4-di-Cl	29	621.900	263.904	580.309
H	9	669.561	231.802	521.463	3,5-di-Cl	30	621.708	264.657	582.654
3- OC_2H_5	10	850.630	288.625	607.647	2-Cl,4-OMe	31	744.220	284.470	596.954
3- OMe	11	767.427	269.564	577.212	2-OMe, 4-Cl	32	745.357	284.612	600.102
4-Cl	12	645.873	247.974	550.289	2-Cl,4-NO_2	33	660.068	280.929	596.231
4-Br	13	645.839	248.969	561.081	2-Me,4-NO_2	34	764.555	289.653	600.679
2-Br	14	647.344	248.359	542.882	4-Me,2-NO_2	35	765.500	288.178	603.295
4-SO_2NH_2	15	750.640	310.770	636.986	2-Br,4-NO_2	36	660.181	281.628	605.695
3-SO_2NH_2	16	749.491	311.410	641.542	4-NO2,2-CF_3	37	706.048	316.388	638.432
3-Cl	17	645.693	248.233	552.270	2-Cl	38	646.700	247.711	549.620
3-Br	18	645.622	249.241	562.381	2-SO_2NH_2	39	751.986	305.094	599.667
3-NO_2	19	682.214	266.354	577.517	4-Ethyle	40	835.160	274.045	584.565
3-CF_3	20	692.154	283.366	606.719	3-Ethyle	41	835.085	274.250	586.132
4-CF_3	21	693.052	283.015	605.348					

Table-2: Descriptor values used in present study.

Comp. No.	$^1\chi$	J	H	W	WW	WP	Sz	E_{lumo}/ev	Comp.No.	$^1\chi$	J	H	W	WW	WP	Sz	E_{lumo}/ev
1	9.381	1.800	64.46	880	3188	29	1324	2.965	22	9.009	1.801	59.81	758	2633	27	1148	2.285
2	9.026	1.912	60.80	714	2285	28	1060	2.938	23	9.382	1.800	64.46	880	3188	29	1324	2.203
3	9.526	1.912	65.16	831	2769	29	1209	2.965	24	9.382	1.800	64.46	880	3188	29	1324	0.980
4	9.009	1.801	59.81	758	2633	27	1148	2.910	25	9.682	1.816	69.60	1004	3746	31	1502	2.312
5	9.509	1.771	63.85	897	3324	28	1341	2.910	26	8.899	1.921	61.19	710	2269	29	1068	2.856
6	8.471	1.825	55.66	638	2081	25	974	2.883	27	8.882	1.916	61.05	712	2278	28	1072	2.557
7	8.488	1.890	56.24	616	1918	26	930	2.802	28	8.899	1.950	61.45	700	2197	29	1048	2.448
8	8.471	1.855	55.9	627	1994	25	952	2.856	29	8.882	1.861	60.61	732	2432	28	1112	2.312
9	8.077	1.837	51.29	536	1649	23	818	2.829	30	8.865	1.887	60.76	722	2349	27	1092	2.258
10	9.509	1.836	64.41	864	3030	28	1275	2.802	31	9.420	1.875	65.21	847	2927	30	1273	2.611
11	9.009	1.853	60.22	736	2448	27	1140	2.774	32	9.420	1.924	65.63	825	2742	30	1229	2.747
12	8.471	1.825	55.66	638	2081	25	974	2.557	33	9.793	1.880	70.06	974	3497	32	1456	0.680
13	8.471	1.825	55.66	638	2081	25	974	2.560	34	9.793	1.880	70.06	974	3497	32	1456	0.680
14	8.488	1.890	56.24	616	1918	26	930	2.693	35	9.793	1.880	70.89	974	3127	32	1456	0.762
15	9.680	1.730	69.60	1004	3746	31	1502	2.258	36	9.793	1.880	70.06	974	3497	32	1456	0.734
16	9.680	1.810	70.37	960	3365	31	1414	1.958	37	11.004	2.052	86.85	1325	4778	38	1921	0.435
17	8.471	1.855	55.90	627	1994	25	952	2.557	38	8.488	1.900	56.24	616	1918	26	930	2.638
18	8.471	1.855	55.90	627	1994	25	952	2.584	39	9.700	1.910	71.43	916	3028	32	1326	1.958
19	9.382	1.869	65.05	847	2905	29	1258	0.843	40	9.010	1.620	59.81	758	2633	27	1148	2.883
20	9.682	1.900	70.37	960	3365	31	1414	2.285	41	9.010	1.670	60.22	736	2448	27	1104	2.856
21	9.682	1.816	69.60	1004	3746	31	1502	2.421									

Results and Discussion

Several linear QSPR models involving three-eight descriptors are established and strongest multivariable correlations are identified by the back ward method are significant at the 0.05 level and regression analysis of the SPSS program.

In the first of this study we drown scattering plots of C_V, S and E_{th} versus the seven topological indices (^{1}X, J, W, Sz, WW, WP, H) and, E_{lumo}. Some of these plots are given in Fig. (1-3), respectively.

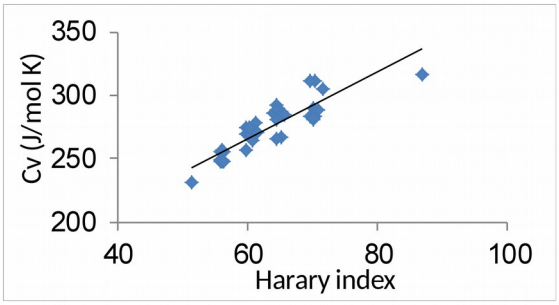


Fig. 1: Plots of the Harary index (H) versus heat of capacity (C_v) of 41sulfonamides.

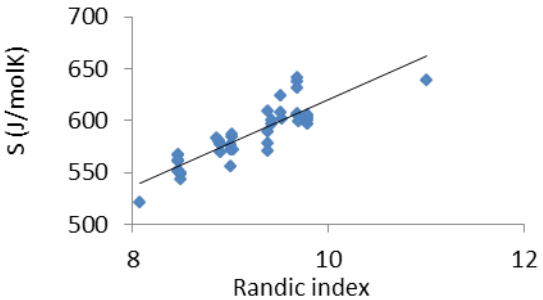


Fig. 2: Plots of the Randic index (^{1}X) versus entropy(S) of 41 sulfonamides.

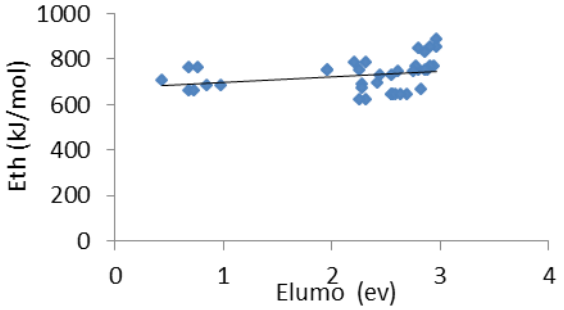


Fig. 3: Plots of the LUMO energy (E_{lumo}) versus thermal energy (E_{th}) of 41 sulfonamides.

Distribution of the dependent variable against the independent variable for 41chemicals employed in developing quantitative structure-Properties relationship. For obtaining appropriate QSPR model we have used maximum R^2 method and followed Back ward regression analysis. The predictive ability of the model is discussed on the basis of predictive correlation coefficient.

QSPR Models for Heat Capacity (C_V)

Initial regression analysis indicated that combination of seven topological indices and E_{lumo} plays a dominating role in modeling the heat capacity. In Table-3 are given the regression parameters and quality of correlation of the proposed models for heat capacity of 41 sulfonamides.

It turns out that the heat capacity (C_V) has a highly correlation with all descriptors as well as with a combination of the four parameters, namely, the LUMO energy (E_{lumo}), Balaban (J), Harary (H) and Wiener Polarity (WP) indices. Fig 4 shows the linear correlation between the observed and the predicted heat capacity values obtained using Eq. (1).

Table-3: Regression parameters and quality of correlation of the proposed models for the heat capacity.

Model	independent variables	R	R^2	R^2_{adj}	RMSE	F
1	E_{lumo}, J, WP, ^{1}X, H ,Sz ,W, WW	0.954	0.911	0.888	6.718	40.804
2	E_{lumo}, J, WP, ^{1}X, H ,Sz ,W	0.954	0.911	0.892	6.620	48.015
3	E_{lumo}, J, WP, ^{1}X, H,W	0.954	0.910	0.895	6.529	57.586
4	E_{lumo}, J, WP, H,W	0.954	0.910	0.897	6.451	70.736
5	E_{lumo}, J, WP,H	0.953	0.908	0.898	6.426	88.938

Model 3.1.5

C_V =28.092– 3.759J+ 0.489H+ 0.918 WP +2.605 E_{lumo} (1)

N=41, R=0.953, R^2=0.908, R^2_{adj}=0.898
RMSE=6.426, F=88.938, Sig=0.000, DW=1.429

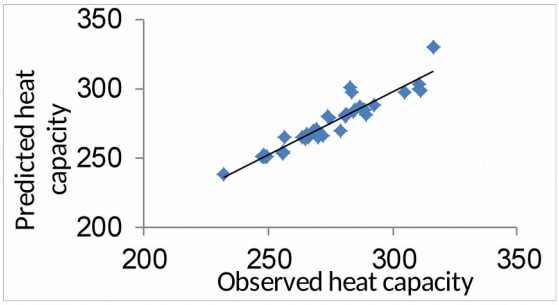


Fig. 4: Comparison between the predicted and observed values of heat capacity by MLR (Eq. 1).

QSPR Models for Thermal Energy (E_{th})

In Table-4 are given the regression parameters and quality of correlation of the proposed models for the thermal energy of 41 sulfonamides.

Table-4: Regression parameters and quality of correlation of the proposed models for the thermal energy.

Model	independent variables	R	R^2	R^2_{adj}	F	RMSE
1	E_{lumo}, J, WW, WP, ^{1}X, H, Sz, W	0.823	0.678	0.597	8.414	45.704
2	E_{lumo}, J, WW, WP, ^{1}X, H, W	0.815	0.665	0.593	9.339	45.921
3	E_{lumo}, J, WW, WP, ^{1}X, H	0.804	0.647	0.584	10.370	46.431
4	E_{lumo}, J, WW, ^{1}X, H	0.797	0.635	0.582	12.156	46.539
5	E_{lumo}, J, WW, ^{1}X	0.782	0.611	0.568	14.130	47.350

Statistically significant models are obtained when four descriptors are used and that the quality of the model goes on improving with higher parametric modeling (Table-4), the tetra parametric model containing four descriptors (E_{lumo}, J, WW, ^{1}X) is found as below:

Model 3.2.5

E_{th} = -129.111 + 59.458^{1}X - 102.257 J– 0.032 WW +418.521 E_{lumo} (2)

N= 41, R= 0.782, R^2 =0.611, R^2_{adj} =0.568, S=47.351
F= 14.130, Sig = 0.000, DW=0.947

Fig 5 shows the linear correlation between the observed and the predicted thermal energy values obtained using Eq. (2).

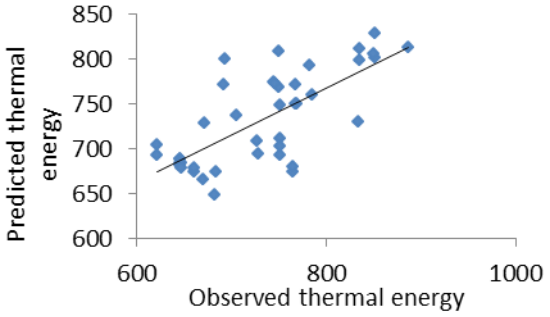


Fig. 5: Comparison between the predicted and observed values of thermal energy by MLR (Eq.2).

In Table-5 are given the regression parameters and quality of correlation of the proposed models for the entropy of 41 sulfonamides.

Table-5: Regression parameters and quality of correlation of the proposed models for the entropy.

Model	independent variables	R	R^2	R^2_{adj}	RMSE	F	Sig
1	E_{lumo}, J, WP, ^{1}X, H, Sz, W, WW	0.937	0.879	0.848	10.795	28.988	0.000
2	E_{lumo}, J, WP, ^{1}X H, Sz, W	0.937	0.879	0.852	10.651	34.015	0.000
3	E_{lumo}, J, WP, H, Sz, W	0.936	0.877	0.855	10.566	40.250	0.000
4	E_{lumo}, J, H, Sz, W	0.935	0.874	0.856	10.521	48.573	0.000
5	E_{lumo}, J, H, Sz	0.934	0.873	0.859	10.426	61.736	0.000
6	E_{lumo}, J, H	0.932	0.869	0.858	10.441	81.700	0.000

It turns out that the entropy(S) has a good correlation with all descriptors as well as with a combination of the three parameters, namely, the LUMO energy (E_{lumo}), Balaban (J) and Harary (H) indices. Fig 6 shows the linear correlation between the observed and the predicted heat capacity values obtained using Eq. (3).

Model 3.3.6

S = 102.931 -23.221J+1.155H+3.110E_{lumo} (3)

N=41, R= 0.932, R^2 =0.869, R^2_{adj} =0.858,
RMSE=10.441, F= 81.700, DW=1.761, Sig = 0.000

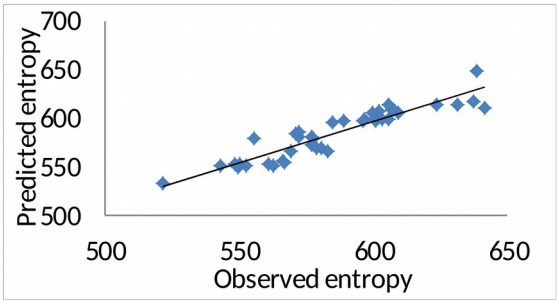


Fig. 6: Comparison between the predicted and observed values of entropy(S) by MLR (Eq.3).

We studied the relationship between topological indices and the LUMO energy (E_{lumo}) to the thermal energy (E_{th}), heat capacity (C_v) and entropy (S) of 41 sulfonamides**.**

In this study, to find the best model for predict the properties mentioned, we will use the following:

Verification and Validity of models

In this section for verification and validity of the regression models, we will focus on the Durbin-Watson statistic and unstandardized predicted and residual values and collinearity.

Test for Autocorrelation by Using the Durbin-Watson statistic

The Durbin-Watson statistic ranges in value from 0 to 4. A value near 2 indicates non-autocorrelation; a value toward 0 indicates positive autocorrelation; a value toward 4 indicates negative autocorrelation. Therefore the value of Durbin-Watson statistic is close to 2 if the errors are uncorrelated. In our all models, the value of Durbin-Watson statistic is close to 2(See eq.1-3) and hence the errors are uncorrelated.

Multicollinearity

Multicollinearity in regression is a condition that occurs when some predictor variables in the model are correlated with other predictor variables. Severe multicollinearity is problematic because it can increase the variance of the regression coefficients, making them unstable. Multicollinearity does not affect the goodness of fit and the goodness of prediction. The coefficients (linear discriminant function) cannot be interpreted reliably, but the fitted (classified) values are not affected. Multicollinearity has the same effect in discriminant analysis as in regression.

In addition, multicollinearity test done to avoid habits in the decision making process regarding the partial effect of independent variables on the dependent variable. Good regression model should not happen correlation between the independent variables or not happen multicollinearity. Test multicollinearity as a basis the variance inflation factor (VIF) value of multicollinearity test results using SPSS. If the VIF value lies between1-10, then there is no multicollinearity, and if the VIF<1 or >10, then there is multicollinearity

In all our final models, the Multicollinearity has not existed, because the values of correlations between independent variables are not near to one and VIFs value lies between 1-10. Furthermore we have computed Q^2 (Eq.4) by 50% of data, randomly, that are positive and less than one.

$$Q^2 = 1 - \frac{\sum(Y_i - \hat{Y}_{i|i})^2}{\sum(Y_i - \bar{Y})^2} \qquad Q^2 \leq 1 \qquad (4)$$

where the notation i|i indicates that the response is predicted by a model estimated when the i-th sample was left out from the training set.

We studied the validation of linearity between the molecular descriptors in the models 3.1.5, 3.2.5 and 3.3.6. We obtained by SPSS the pearson coefficient correlation and collinearity statistics as follows Tables (6-8).

For model 3.1.5 the pearson correlation (H, WP) is near one, and VIF(H),VIF(WP)>10, therefore there is a linearity between H and WP. After removed WP from this model, we corrected model 3.1.5 as follows:

$Cv=111.995-47.897J+3.595H+10.431E_{lumo}$ (5)

N=41, R=0.949, R^2=0.900, R^2_{adj}=0.892, DW=1.252
RMSE =6.617, F=110.800, Sig=0.000, Q^2=0.956

Table-6: Correlation between the molecular descriptors (model 3.1.5).

Pearson correlations(model 3.1.5)					Collinearity statistical		Corrected model
	J	H	WP	E_{lumo}	Tolerance	VIF	VIF
J	1	0.293	0.350	-0.297	0.790	1.266	1.117
H		1	0.984	-0.664	0.30	33.808	1.821
WP			1	-0.678	0.028	35.931	-
E_{lumo}				1	0.537	1.863	1.826

Similarity model 3.1.5 we obtained corrected model 3.2.5 as follows:

Table-7: Correlation between the molecular descriptors (model 3.2.5).

Pearson correlations(model 3.2.5)					Collinearity Statistical		Corrected model
	^1X	J	WW	E_{lumo}	Tolerance	VIF	VIF
^1X	1	197	0.959	-0.628	0.055	18.048	1.650
J		1	0.048	-0.297	0.641	1.559	1.097
WW			1	-0.604	0.056	17.787	-
E_{lumo}				1	0.561	1.783	1.740

Table-8: Correlation between the molecular descriptors (model 3.3.6)

Pearson correlations				Collinearity statistical	
	J	H	E_{lumo}	Tolerance	VIF
J	1	0.293	-0.297	0.895	1.117
H		1	0.059	0.549	1.821
E_{lumo}			1	0.548	1.826

$$E_{th}=94.004-230.189J+70.118E_{lumo}+98.635^1X \quad (6)$$

N=41, R=0.723, R^2=0.523, R^2_{adj}=0.485, DW=1.081
RMSE =51.694, F=13.544, Sig=0.000, Q^2=0.999

In model 3.3.6 there is no multicollinearity between J, H, E_{lumo} (Table-8).

Regular Residuals

The residual that is the difference between the observed value of the dependent variable (y) and the predicted value (ŷ). Comparison between predicted and observed values of thermal energy, heat capacity and entropy of respect sulfonamides show in Table-9. Plot the residuals, and use other diagnostic statistics, to determine whether our model is adequate and the assumptions of regression are met. The residuals can also identify how much a model explains the variation in the observed data.

The residuals values of heat capacity, thermal energy and entropy expressed by Eqs.(3,5, 6) show in Fig (7-9). The residual plot shows a fairly random pattern. This random pattern indicates that a linear model provides a decent fit to the data.

Table-9: Comparison between predicted and observed values of thermal energy, heat capacity and entropy of respect sulfonamides

Comp. No.	Observed (Cv)	Predicted (Cv)	Residual	Observed (Eth)	Predicted (Eth)	Residual	Observed (S)	Predicted (S)	Residual
1	292.805	288.431	4.374	886.871	812.862	74.009	609.210	606.716	2.494
2	268.306	269.628	-1.322	768.982	750.172	18.810	571.920	581.507	-9.587
3	287.137	285.583	1.554	852.131	801.382	50.749	601.845	608.061	-6.216
4	269.238	271.094	-1.856	767.603	772.083	-4.480	572.359	586.518	-14.159
5	286.539	287.054	-0.515	851.809	828.306	23.503	623.510	614.454	9.056
6	256.021	254.744	1.277	751.108	711.599	39.509	566.244	556.625	9.619
7	255.875	252.871	3.004	751.501	692.634	58.867	548.320	552.986	-4.666
8	256.213	253.888	2.325	751.087	702.800	48.287	567.088	554.648	12.440
9	231.802	237.896	-6.094	669.561	666.188	3.373	521.463	534.668	-13.205
10	288.625	284.827	3.798	850.630	805.771	44.859	607.647	609.624	-1.977
11	269.564	268.658	0.906	767.427	750.577	16.850	577.212	582.107	-4.895
12	247.974	251.343	-3.369	645.873	688.740	-42.867	550.289	553.029	-2.740
13	248.969	251.375	-2.406	645.839	688.951	-43.112	561.081	553.062	8.019
14	248.359	251.734	-3.375	647.344	684.991	-37.647	542.882	551.783	-8.901
15	310.77	302.887	7.883	750.640	808.893	-58.253	636.986	618.535	18.451
16	311.41	298.694	12.716	749.491	769.443	-19.952	641.542	610.747	30.795
17	248.233	250.769	-2.536	645.693	681.835	-36.142	552.270	551.349	0.921
18	249.241	251.051	-1.810	645.622	683.728	-38.106	562.381	551.647	10.734
19	266.354	265.113	1.241	682.214	648.286	33.928	577.517	579.494	-1.977
20	283.366	297.794	-14.428	692.154	771.851	-79.697	606.719	609.422	-2.703
21	283.015	300.468	-17.453	693.052	800.723	-107.671	605.348	615.625	-10.277
22	256.397	264.574	-8.177	670.995	728.259	-57.264	555.246	579.623	-24.377
23	280.779	280.483	0.296	784.933	759.530	25.403	589.062	598.361	-9.299
24	265.656	267.726	-2.070	683.455	673.775	9.680	570.980	584.868	-13.888
25	310.896	299.331	11.565	783.198	793.080	-9.882	631.460	614.422	17.038
26	278.835	269.744	9.091	833.442	729.824	103.618	576.853	573.429	3.424
27	272.298	266.361	5.937	727.709	708.332	19.377	578.838	569.517	9.321
28	270.275	265.033	5.242	728.265	694.540	33.725	569.053	567.304	1.749
29	263.904	264.858	-0.954	621.900	703.814	-81.914	580.309	569.893	10.416
30	264.657	263.589	1.068	621.708	692.366	-70.658	582.654	566.949	15.705
31	284.47	283.843	0.627	744.220	774.622	-30.402	596.954	600.660	-3.706
32	284.612	284.424	0.188	745.357	772.879	-27.522	600.102	599.418	0.684
33	280.929	280.896	0.033	660.068	674.864	-14.796	596.231	598.663	-2.432
34	289.653	280.896	8.757	764.555	674.864	89.691	600.679	598.663	2.016
35	288.178	284.735	3.443	765.500	680.614	84.886	603.295	599.568	3.727
36	281.628	281.460	0.168	660.181	678.650	-18.469	605.695	599.259	6.436
37	316.388	330.460	-14.072	706.048	737.540	-31.492	638.432	649.924	-11.492
38	247.711	250.681	-2.970	646.700	678.833	-32.133	549.620	550.617	-0.997

39	**305.094**	**297.715**	**7.379**	**751.986**	**748.396**	**3.590**	**599.667**	**606.200**	**-6.533**
40	**274.045**	**279.481**	**-5.436**	**835.160**	**811.952**	**23.208**	**584.565**	**596.405**	**-11.840**
41	**274.25**	**278.279**	**-4.029**	**835.085**	**798.550**	**36.535**	**586.132**	**593.308**	**-7.176**

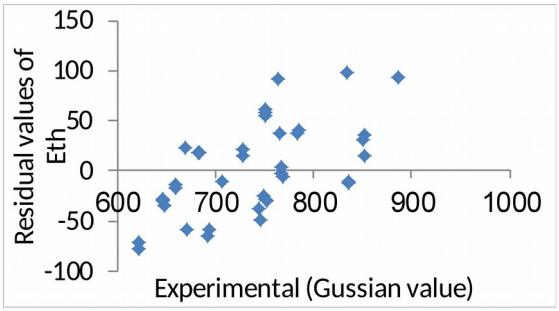


Fig. 7: Plot of residuals against experimental value with Eq. (6) for the thermal energy of 41 sulfonamides.

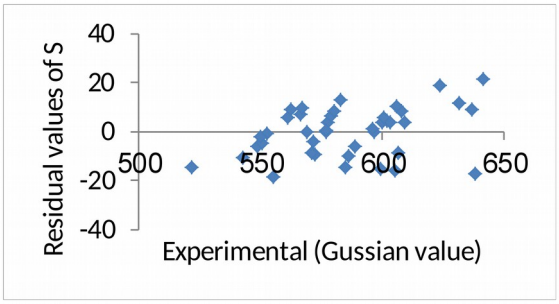


Fig. 8: Plot of residuals against experimental value with Eq. (3) for the entropy of 41 sulfonamides.

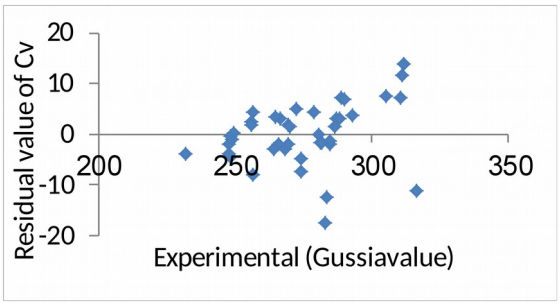


Fig. 9: Plot of residuals against experimental value with Eq. (5) for the heat capacity of 41 sulfonamides.

Conclusions

Graph theory has provided the chemist with a variety of very useful tools. Tis contain valuable structural information as evidenced by the success of their widespread applications in QSAR/QSPR. In this work, the relationship between topological indices (J, W, WW, WP, H, ^1X, Sz), E_{lumo} and the heat capacity (C_V), entropy (S), thermal energy (E_{th}) of 41 sulfonamides were studied.

The aforementioned results and discussion lead us to conclude that combining the three descriptors (E_{lumo}, J, H) could be used successfully for modeling and predicting the heat capacity (C_V) and the entropy (S). Three descriptors (E_{lumo}, J, ^1X) could be predict the thermal energy of compounds. The training set models established by MLR method have good correlation of physico-chemical properties, which means QSPR models, could be used for prediction of the heat capacity (C_V), entropy (S), thermal energy (E_{th}) for a set of 41sulfonamides.

Acknowledgments

The author is thankful to Islamic Azad University Arak, Iran for their financial support.

References

1. M. Randic', Quantitative Structure–Property Relationship: Boiling Points of Planar Benzenoids, *New J. Chem*., **20**, 1001 (1996).
2. S. E. Rita, H. I. Sadigm and A. S. Bahjat, Quantum Chemical QSAR Study of 1– phenyl–X–benzimidazoles as Inhibitors of PDGFRT Tyrosin Kinase, *Int. J. Pharm. Tech Res*., **3**, 2183 (2011).
3. A. S. Bahjat, S. E. Rita, H. I. Sadigm and A. H. Kawkab, Theoretically Predicted Descriptors Based Quantitative Structure–Activity Relationship Study of the Activity of Acridines Against B–16 Melanoma, *Am. J. Sci*., **8**, 773 (2011).
4. A. H. Kawkab, A. H. Wisam and H. I. Sadigm, QSAR Study and Improving it of Some Schiff- Base Ligands as Anticancer for Prostate Cancer, *J. Chem. Pharm Res*., **4**, 1702 (2012).
5. M. H. I. Sadiq, A. H. Kawkab and A. S. A. M. Hasanian, QSPR Study of the Intrinsic Viscosity[η] of Polyisobutylene Solution, *Der pharmacia Lettre*.,**4**, 1826 (2012).
6. A. A. Toropov, O. M. Nabiev and F.Torrens, QSPR Modeling of Hydrocarbon Dipole Moments by Means of Correlation Weighting of Local Graph Invariants, *J. Theor. Comput. Chem*., **2**, 139 (2003).
7. H. Redmond, J.E. Thompson, Evaluation of a quantitative structure–property relationship (QSPR) for predicting mid-visible refractive

index of secondary organic aerosol (SOA), *Phys. Chem. Chem. Phys*., **13**, 6872 (2011).

8. M. Uzma, R. Sitra and V.Wolfgang , 3D-QSPR Method of Computational Technique Applied on Red Reactive Dyes by Using CoMFA Strategy, *Int. J. Mol. Sci*., **12**, 8862 (2011).
9. A. Antreas, M. Georgia and M. Kalliopi, Prediction of high weight polymers glass transition temperature using RBF neural networks, *J. Mol. Struct. (Theochem)*., **1**,716 (2005).
10. R. Irana, K. Victor, O. Olga, and A. Sergei, QSPR analysis of peroxidase substrates reactivity of the new phenolic substrates with satisfactory reliability, *Ch & Ch .T*., **3**,255 (2009).
11. A. Sabljic, D. Horvatic, GRAPH III: A computer program for calculating molecular connectivity indices on microcomputers, *J. Chem. Inf. Comput. Sci*., **33**,292 (1993).
12. L. Poglina, Molecular modeling by linear combination of connectivity indexes, *J. Phy. Chem*., **99**,925 (1995).
13. S. Chaterjee, A. Hadi and B. Price, *Regression Analysis by Examples*, 3nd. Ed. Wiley VCH: New York (2000).
14. H. Waterbeemd, Vande (Ed) *Chemometric Method in Molecular Design, in Methods and Principles in Medicinal Chemistry*, Vol. 12, VCH, Weinleium (Ger) (1995).
15. J. Devillers, W. Karcher, (Eds) *Applied Multivariate Analysis in SAR and Environmental Statistics* Vol. 2*, in Chemical and Environmental Sciences*, Klawer Academic Pub. Dorderch (Netherland) (1991).
16. J. Y.Winum, A. Scozzafava, J. L. Montero, C. T.Supuran, In Vivo Evaluation of Selective Carbonic Anhydrase Inhibitors as Potential Anticonvulsant Agents, *Med. Res. Rev*., **25**,186 (2005).
17. H. Wiener, Structural Determination of Paraffin Boiling Points, *J. Am. Chem. Soc*., **69**, 17 (1947).
18. P. V. Kadikar, N. V. Deshpande and G. Domotor, The Szeged Index and an Analogy with the Wiener Index, *J. Chem. Inf. Compt. Sci*., **35**, 547 (1995).
19. M. Randic', On Characterization of Molecular Branching, *J. Am. Chem. Soc.*, **97**, 6609 (1975).
20. A. T. Balaban, Highly Discriminating Distance Based Topological Indices, *Chem. Phys. Lett*., **89**, 399 (1982).
21. I. Gutman, A New Hyper−Wiener index, *Croat. Chem. Acta*., **77**, 61 (2004).
22. A. Behmaram, H. Yousefi- Azari and A. R. Ashrafi, Wiener Polarity Index of Fullerenes and Hexagonal System, *Appl. Math.Lett*., **25**, 1510 (2012).
23. C. K. Das, B. Zhou and N. Trinajstic, Bounds on Harary index, *J. Math. Chem*., **46**, 1369 (2009).
24. F. Shafiei, H. Hosseini, Quantitative Structure Property Relationship Models for the Prediction of Gas Heat Capacity of Benzene Derivatives UsingTopological Indices, MATCH *Commun. Math. Comput. Chem*.,**75**, 583 (2016).
25. J. Singh, M. Lakhwani, P. V. Khadikar, A. T. Balaban and C. T. Supuran, QSAR Study on the Inhibition of the Human Carbonic Anhydrase Cytosolic Isozyme VII, Roumaine de Chimie., **51**, 691 (2006).
26. G. Melagraki, A. Afantitis, H. Sarimveis, O. Igglessi-Markopoulou and C. T. Supuran QSAR Study on Para-Substituted Aromatic Sulfonamides as Carbonic Anhydrase II Inhibitors Using Topological Information Indices, *Bio & Med. Chem*., **14**, 1108 (2006).
27. Web Search Engine Developed by ChemAxon; software available at http:// *WWW. Chemicalize. Org*.