Original Research Article

Comparative study of CBS-Q calculated and experimental Pka values for fluoro-acetoxy derivative

Abstract

The pKa values were calculated for some acetoxy group molecules using CBS-Q method which is one of the Complete Basis Set methods to find accurate energies. The acetoxy group molecules were also planned by Quantitative Structure Activity Relationship (QSAR) to study their effect on paraoxonase1 activity.

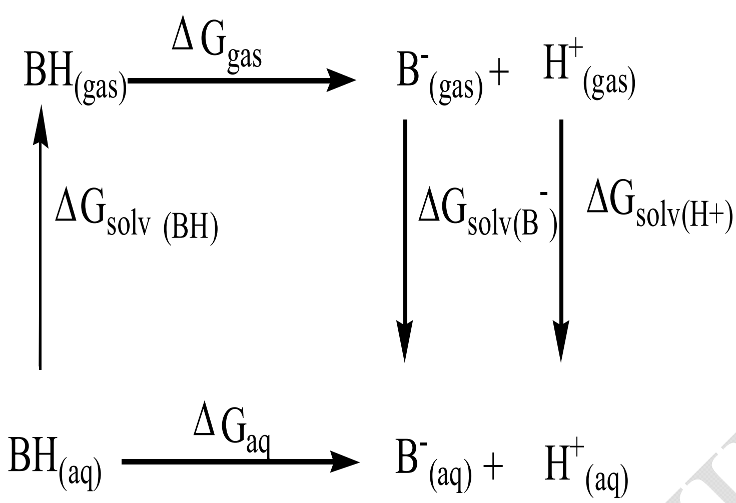
The results of this study showed a strong relationship, (R^2=0.99) between the calculated and experimental pka, also showed correlations between the activity of the enzyme and some of the studied descriptors. Moreover, the results of the study revealed that by using the SPSS program, there is a correlation between LUMO, Softness, Nucleofugality and Electrofugality as dependent variables and Cal. pKa as an independent variable.

Keywords: Acetoxy, QSAR, pKa, HOMO, and LUMO

1-Introduction

In this study, we have calculated pKa value for acetoxy group using CBS-Q method which is one of the Complete Basis Set methods to find accurate energies [2-6]. Acetoxy derivatives thought to be molecules that may activate paraoxonase1 (PON1). We used the following thermodynamic cycles as [7].

1.1.Cycle 1: The thermodynamic cycle 1 were given at Scheme 1.



Scheme 1 Thermodynamic Cycle 1 Interrelationship between the gas phase and solution thermodynamic parameters.

Theoretical pKa values are commonly obtained by using the thermodynamic cycles. Experimental solvation free energy of H^+ was used to calculate pKa value in thermodynamic cycles.

pKa values were obtained by Eq.1

$$pKa = \Delta G_{aq} / 2.303RT \qquad (1)$$

$$\Delta G_{aq} = \Delta G_{gas} + \Delta\Delta G_{solv} \qquad (2)$$

In this cycle, ΔGgas can be calculated as in Eq 3

$$\Delta G_{gas} = G_{gas}(H+) + G_{gas}(B-) - G_{gas}(BH) \qquad (3)$$

Since proton electronic energy is zero, Hgas (H+) has been obtained by adding up the translational energy (E = 3/2RT) and PV = RT. This value reported 1.48 kcal/mol at 298 K. Entropy, $S(H^+)$, was calculated by the Sackur–Tetrode equation for gas-phase monoatomic species, so TS = -7.76 kcal/mol at 298 K and 1 atm. Then $G_{gas}(H^+)$ equals -6.28 kcal/mol [8].

Since pKa determination employs the standard free deprotonation energy in solution, 1 M aqueous phase calculations also uses a reference state of 1 M and gas-phase free deprotonation energies are calculated to a reference state of 1 atm, gas-phase free energy difference, ΔG_{gas}, must be referred to 1 M by taking into account the factor RTln 24.46.

$$\Delta G_{gas}\ (1M) = \Delta\Delta G_{gas}\ (1atm) + RTln24.46 \qquad (4)$$

the calculation of $\Delta\Delta G_{solv}$ were obtained with the Eq.5

$$\Delta\Delta G_{solv} = \Delta G_{solv}(H^+) + \Delta G_{solv}(B^-) - \Delta G_{solv}(BH) \qquad (5).$$

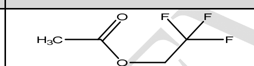
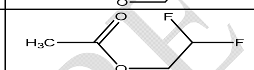
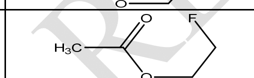
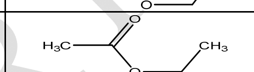
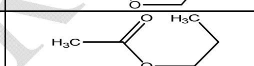
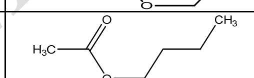
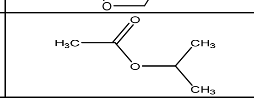
This work is aimed to study the quantum chemical descriptors of acetoxy derivatives, calculate pKa and compared to the experimental pKa for the same compounds that may activate paraoxonase1.

2- Materials and Methods

2.1 Molecules of study

Khersonsky, Tawfik [9] studied a group of molecules including these 7 molecules of study.

Table 1: some of acetoxy derivative compounds [9].

Mol. No.	Name	Structure	Pka
1	Trifluoroethyl acetate		12.4
2	2,2-difluoroethyl acetate		13.3
3	2- fluoroethyl acetate		14.2
4	Ethyl acetate		16.1
5	Propyl acetate		16
6	Butyl acetate		16.1
7	Isopropyl acetate		17.1

CBS-Q was used for the calculation of pKa, these calculations have been used for gas and water phase calculations of acetoxy derivatives. all calculations were performed using Gaussian 09W program [10].

2.2 Calculation Methods

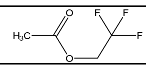
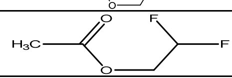
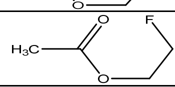
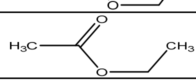
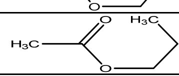
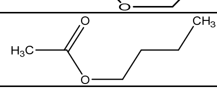
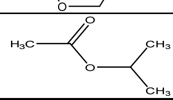
All calculations were performed on Intel core-i7 Sony laptop Computer, using Gaussian 09W program [10]. The CBS-Q method has been used for all gas and water phase calculations for acetoxy molecules. All calculation results have no imaginary frequency at gas and water phase.

The CBS-Q method is one of the effective methods of The Complete Basis Set Methods which were developed by Petersson and coworkers [11, 12, 6]. The CBS methods include some corrections for ab-inito calculation errors. These methods use relatively large basis sets for the structure calculation, medium-sized basis sets for the second-order correlation correction, and small-sized basis sets for higher order correlation corrections. Thus the CBS methods can compute energies for the molecules very accurately [13, 14, 11, 12].

3-Results

The pKa values for the 7 acetoxy derivatives were calculated, thermodynamic data for those molecules were calculated in the case of syn-periplanar position and in antiperiplanar position, these data are presented in Tables 2 and 3.

Table 2 calculated pKa to the experimental pka of some acetoxy derivatives compounds(syn-periplanar -positions)

Mol. No.	Mol. Structure	CH2-g (kcal/mol)	gp-an (kcal/mol)	CH2-aq (kcal/mol)	aq-an (kcal/mol)	Cal. pKa	Exp. pKa
1		-604.71715	-604.179206	-604.724876	-604.269088	10.15	12.40
2		-505.531756	-504.597594	-505.540674	-505.083327	10.87	13.30
3		-406.353085	-405.807454	-406.360854	-405.901682	11.71	14.20
4		-307.193869	-306.638776	-307.200251	-306.737631	13.29	16.10
5		-346.42217	-345.867373	-346.428935	-345.966237	13.33	16.10
6		-385.650767	-385.096011	-385.657331	-385.19437	13.45	16.10
7		-346.426101	-345.870991	-346.432491	-345.968931	13.72	17.10

The calculated pka for the syn-periplanar -position compounds were near to the experimental pKa, and when testing the correlation between them, a significant relationship has been observed, as presented by Figure 1.

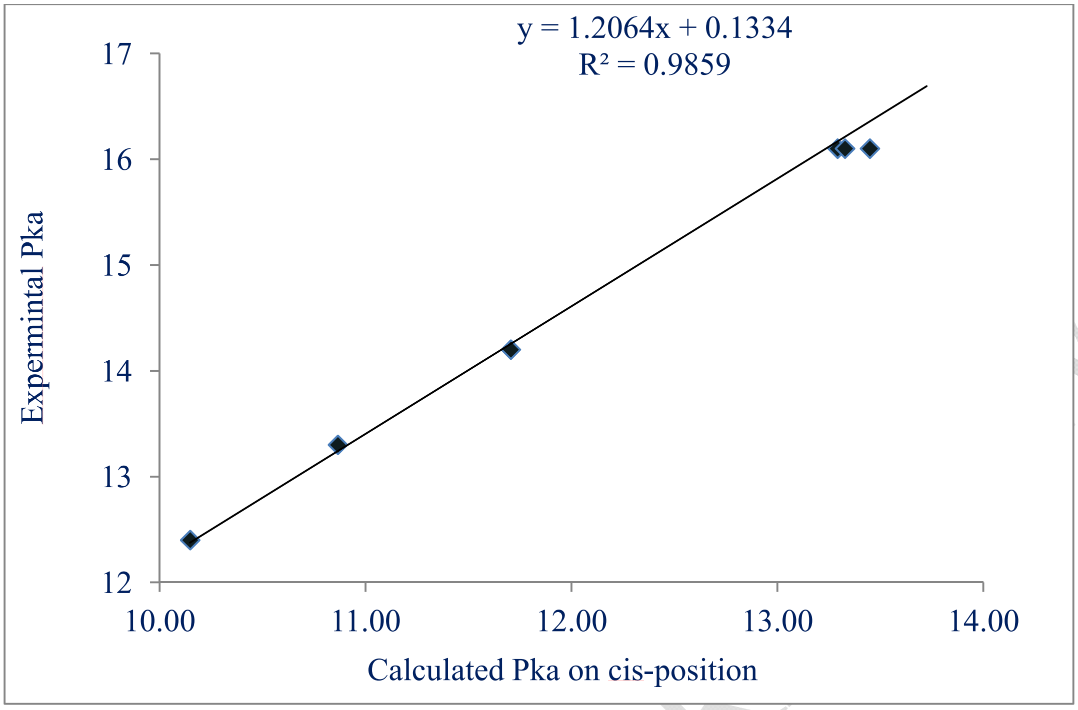


Figure 1: the relationship between the calculated and experimental pka in the syn-periplanar - position

While in antiperiplanar-position the calculated pka was little bit far from experimental one, but also it gave a good relationship.

Table 3 calculated pka to the experimental pKa of some acetoxy derivatives compounds(antiperiplanar-positions)

Mol. No.	CH2-g (kcal/mol)	gp-an (kcal/mol)	CH2-aq (kcal/mol)	aq-an (kcal/mol)	Cal. pka	Exp. pka
1	-604.7132	-604.105244	-604.720172	-604.2715	6.87	12.40
2	-505.528058	-504.999049	-505.534058	-505.08518	6.97	13.30
3	-406.346338	-405.811861	-406.354372	-405.90227	8.45	14.20
4	-307.189504	-306.642444	-307.196235	-306.73863	10.99	16.10
5	-346.416057	-345.871429	-346.423042	-345.96616	10.65	16.10
6	-385.645342	-385.10061	-385.652389	-385.19636	10.26	16.10
7	-346.470306	-345.877325	-346.429248	-345.97077	11.38	17.10

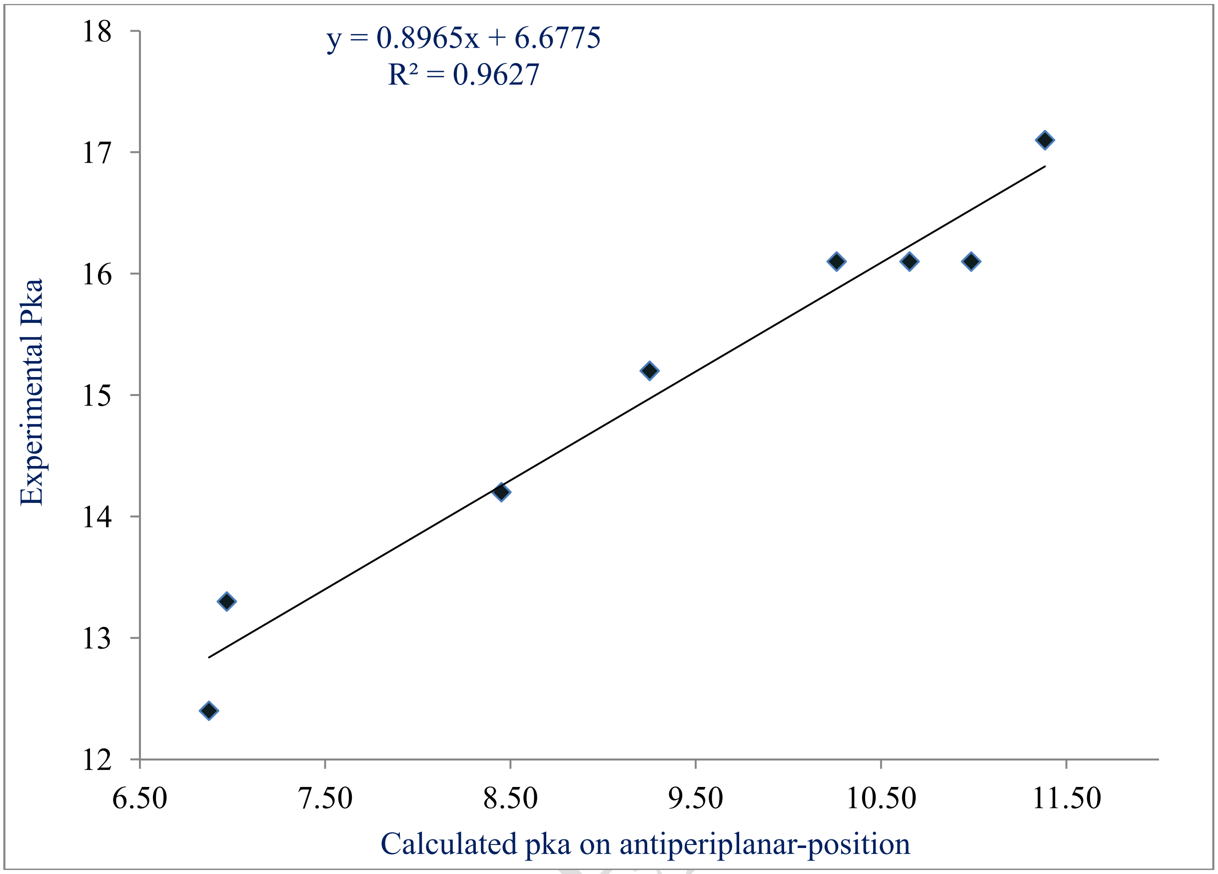


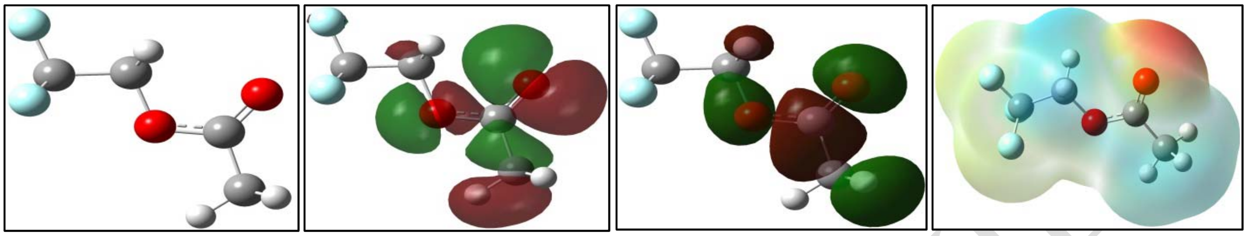
Figure 2: the relationship between the calculated and experimental pka in the antiperiplanarposition

Moreover, the quantum chemical calculations have been carried out at the CBS-Q level of theory using Gaussian-09 series of program package. Some descriptors for the same molecules like E_{HOMO}, E_{LUMO}, Energy gap, Hardness, Softness, Electronegativity, Chemical potential, Electrophilicity index, Electrofugality, and Nucleofugality were calculated, as shown in Table 4.

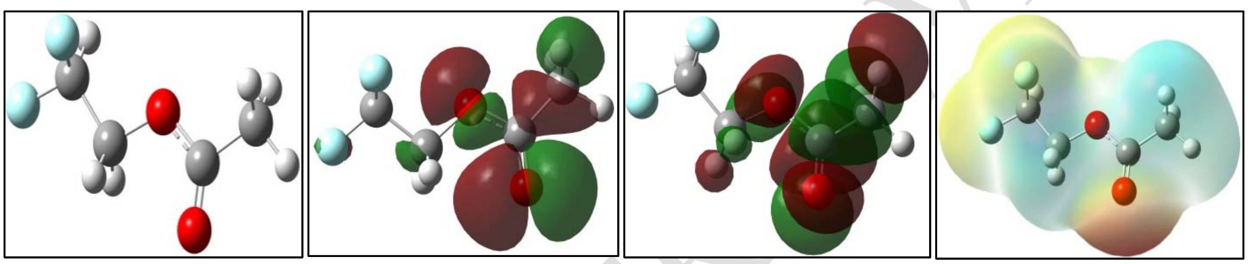
Table 4: calculated descriptors for the acetoxy derivative compounds (syn-periplanar -position)

Mol. No.	HOMO (eV)	LUMO (eV)	Energy Gap(eV)	Hardness (eV)	Softness (eV^{-1})	Electro-negativity (eV)	Chemical potential (eV)	electrophili-city index(ω)	Nucleo-fugality	Electr-fugality	Cal. pka	exp. pka
1	-12.5081	4.428144	16.93623	8.468114	0.059045	4.03997	-4.03997	0.963695	1.157782	9.237722	10.2	12.4
2	-12.3201	4.594951	16.915	8.457501	0.059119	3.862551	-3.86255	0.882016	1.248216	8.973317	10.9	13.3
3	-12.1377	4.761214	16.89895	8.449474	0.059175	3.68826	-3.68826	0.804977	1.341454	8.717974	11.7	14.2
4	-11.7905	5.11959	16.9101	8.455052	0.059136	3.335462	-3.33546	0.657909	1.549973	8.220897	13.3	16.1
5	-11.757	5.079589	16.83663	8.418317	0.059394	3.338727	-3.33873	0.662074	1.532505	8.20996	13.3	16.1
6	-11.7252	5.092379	16.81759	8.408793	0.059462	3.316414	-3.31641	0.653994	1.541977	8.174804	13.4	16.1
7	-11.7472	5.137006	16.88425	8.442127	0.059227	3.305121	-3.30512	0.646983	1.562925	8.173167	13.7	17.1

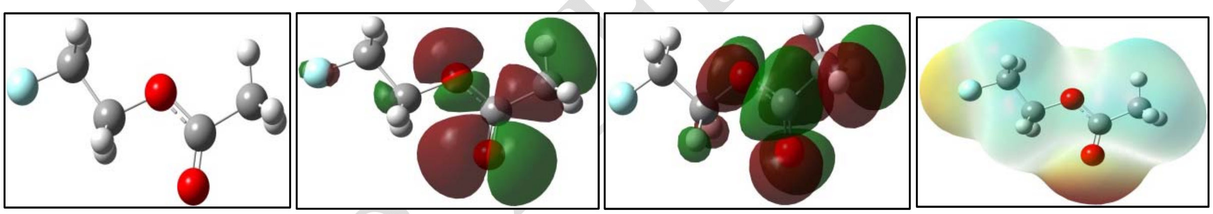
E_{HOMO} is associated with the ability of a molecule to donate an electron, the high E_{HOMO} value indicates the tendency of the molecule to donate electrons to an appropriate acceptor molecule with lower energy MO [15-19]. HOMO and LUMO orbitals of the 7 molecules were obtained from the quantum chemical calculation by the DFT using cbs-q bases sets as shown in Figure 1:



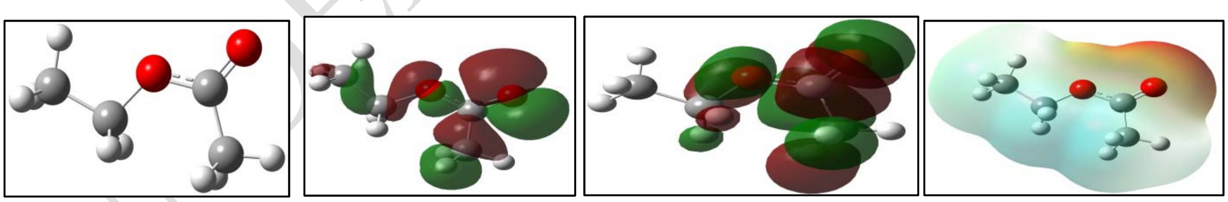
1-OPT 1-HOMO 1-LUMO 1-ESP



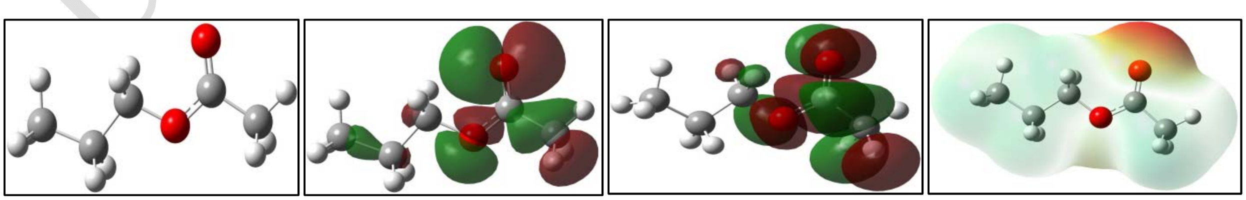
2-OPT 2-HOMO 2-LUMO 2-ESP



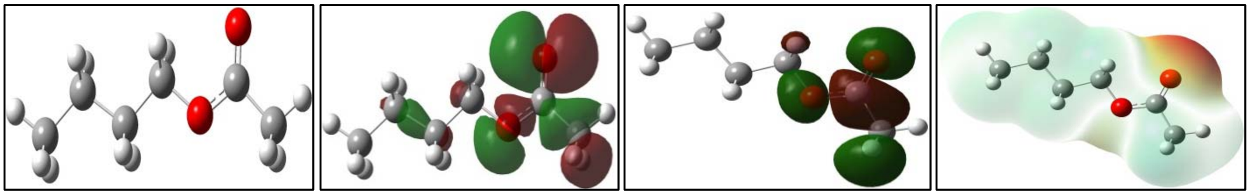
3-OPT 3-HOMO 3-LUMO 3-ESP



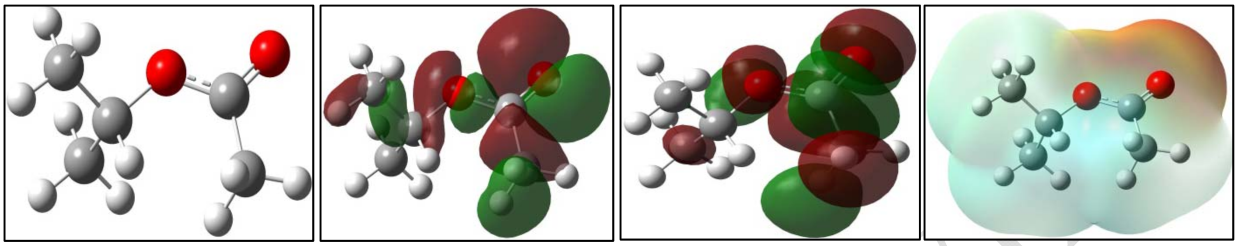
4-OPT 4-HOMO 4-LUMO 4-ESP



5-OPT 5-HOMO 5-LUMO 5-ESP



6-OPT 6-HOMO 6-LUMO 6-ESP



7-OPT 7-HOMO 7-LUMO 7-ESP

Figure 3: Molecular structure, HOMO, LUMO & ESP of the 7 molecules

The highest E_{HOMO} among the 7 syn-periplanar -molecules were: -11.7252 eV, -11.7472 eV, -11.757 eV, -11.7905 eV, -12.1377 eV, -12.3201 eV and -12.5081 eV, recorded with molecules: 6, 7, 5, 4, 3, 2 and 1, respectively. Whereas the E_{LUMO} of these molecules were: 5.137006 eV, 5.11959 eV, 5.092379 eV, 5.079589 eV, 4.761214 eV, 4.594951 eV and 4.428144 eV, these values for 7, 4, 6, 5, 3, 2 and 1 molecules, respectively. The Energy Gap among these molecules were: 16.93623 eV, 16.9150 eV, 16.9101 eV, 16.89895 eV, 16.88425 eV, 16.83663 eV and 16.81759 eV, these for molecules 1, 2, 4, 3, 7, 5 and 6. As the molecules with smaller E_{HOMO}-E_{LUMO} energy gap lead to lower kinetic stability and higher chemical reactivity, so the molecules that have high activity are 1, 2, 4, 3, 7, 5 and 6 respectively [15].

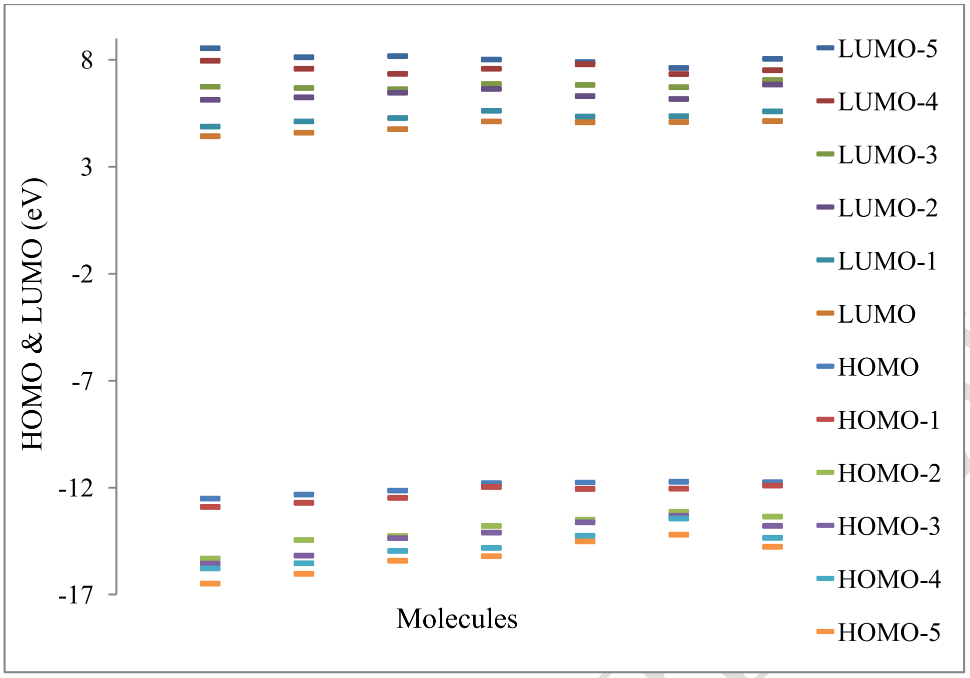


Figure 4: HOMO & LUMO for syn-periplanar -molecules

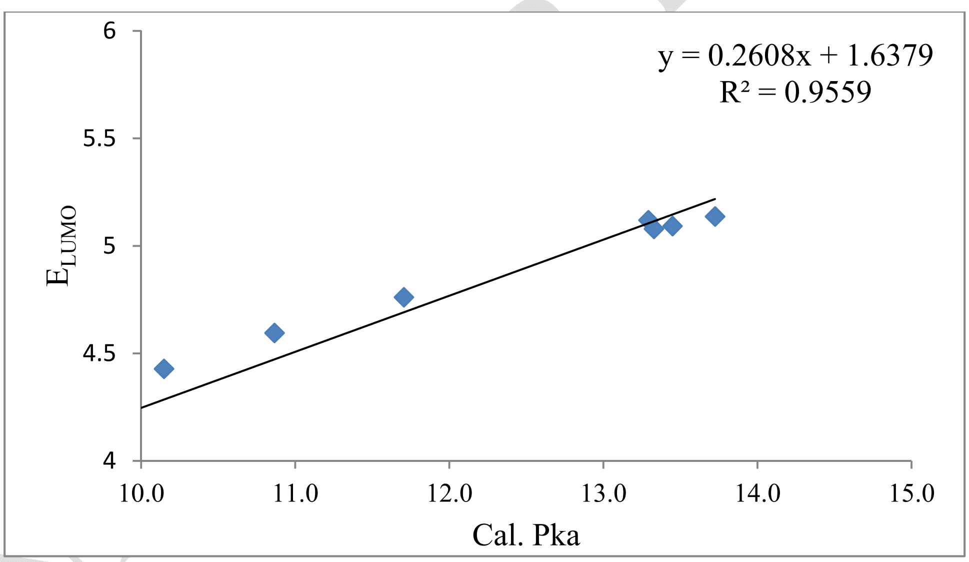


Figure 5: The correlation between calculated pka and LUMO in case of syn-periplanar -position

Nucleofugality is defined as the propensity of an atom or group of them to depart bearing the bonding electron pair in a heterolytic cleavage process [18,19], the highest nucleofugality of the 7 molecules were: 1.562925, 1.549973, 1.541977, 1.532505, 1.341454, 1.248216 and 1.157782. these results for molecules: 7, 4, , 6, 5, 3, 2 and 1. According to these values, those molecules have the activation activity to PON1. Figure 6 represents the relationship between the calculated Pka and nucleofugality, which also confirmed by the results of the statistical analysis.

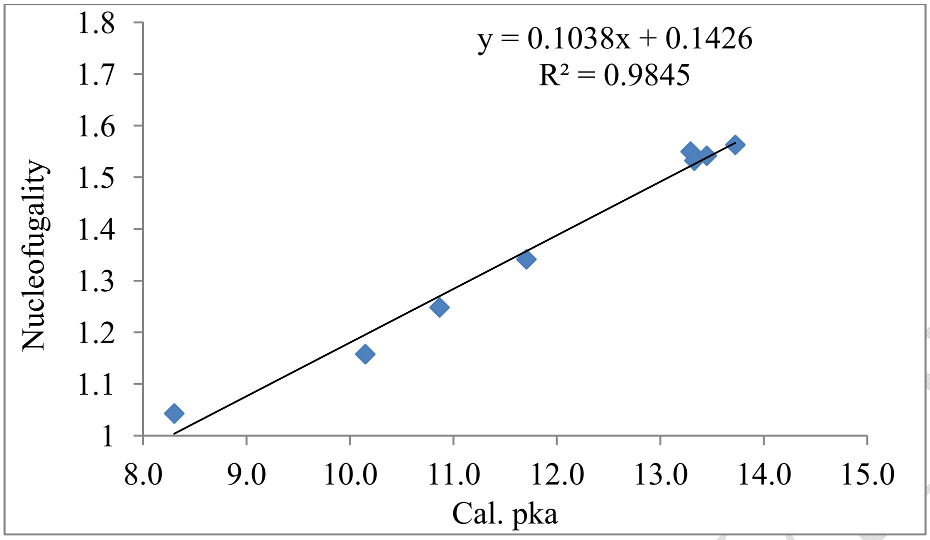


Figure 6: The correlation between calculated pka and Nucleofugality in case of syn-periplanar - position

Table 5: calculated descriptors for the acetoxy derivative compounds (antiperiplanarposition)

Mol. No.	HOMO (eV)	LUMO (eV)	Energy Gap(eV)	Hardness (eV)	Softness (eV^{-1})	Electro-negativity (eV)	Chemical potential (eV)	electrophili-city index(ω)	Nucleo-fugality	Electr-fugality	Cal. pka	exp. pka
1	-12.6836	4.4488	17.132423	8.566212	0.058369	4.117387	-4.117387	0.98952	1.155239	9.390013	6.87	12.4
2	-12.5206	4.6831	17.203718	8.601859	0.058127	3.918743	-3.918743	0.892629	1.274816	9.112301	6.97	13.3
3	-12.3105	4.7901	17.100586	8.550293	0.058478	3.760235	-3.760235	0.826835	1.341747	8.862217	8.45	14.2
4	-11.8975	5.0527	16.950106	8.475053	0.058997	3.422403	-3.422403	0.691019	1.506142	8.350948	10.99	16.1
5	-11.9315	5.1596	17.091062	8.545531	0.05851	3.385939	-3.385939	0.670794	1.55762	8.329499	10.65	16.1
6	-11.9029	5.1724	17.075279	8.53764	0.058564	3.365259	-3.365259	0.663237	1.566799	8.297316	10.26	16.1
7	-11.7818	5.0573	16.839082	8.419541	0.059386	3.362265	-3.362265	0.671345	1.51885	8.243381	11.38	17.1

In case of antiperiplanar position, the calculated descriptors were presented in Table 5, the values were not far of the syn-periplanar -position case, but there was a rearrangement of the molecules, especially for the HOMO and LUMO, Energy Gap and nucleofugality as shown in Table 6, as well as Figures 7 and 8.

Table 6: comparison between antiperiplanar and syn-periplanar -positions highest E_{HOMO}, E_{LUMO} and Energy Gap.

Syn-periplanar - position						**Antiperiplanar-position**					
Mol. No.	E_{HOMO}	Mol. No.	E_{LUMO}	Mol. No.	Energy Gap	Mol. No.	E_{HOMO}	Mol. No.	E_{LUMO}	Mol. No.	Energy Gap
6	-11.7252	7	5.137006	1	16.93623	7	-11.7818	6	5.1724	2	17.20372
7	-11.7472	4	5.11959	2	16.9150	4	-11.8975	5	5.1596	1	17.13242
5	-11.757	6	5.092379	4	16.9101	6	-11.9029	7	5.0573	3	17.10059
4	-11.7905	5	5.079589	3	16.89895	5	-11.9315	4	5.0527	5	17.09106
3	-12.1377	3	4.761214	7	16.88425	3	-12.3105	3	4.7901	6	17.07528
2	-12.3201	2	4.594951	5	16.83663	2	-12.5206	2	4.6831	4	16.95011
1	-12.5081	1	4.428144	6	16.81759	1	-12.6836	1	4.4488	7	16.83908

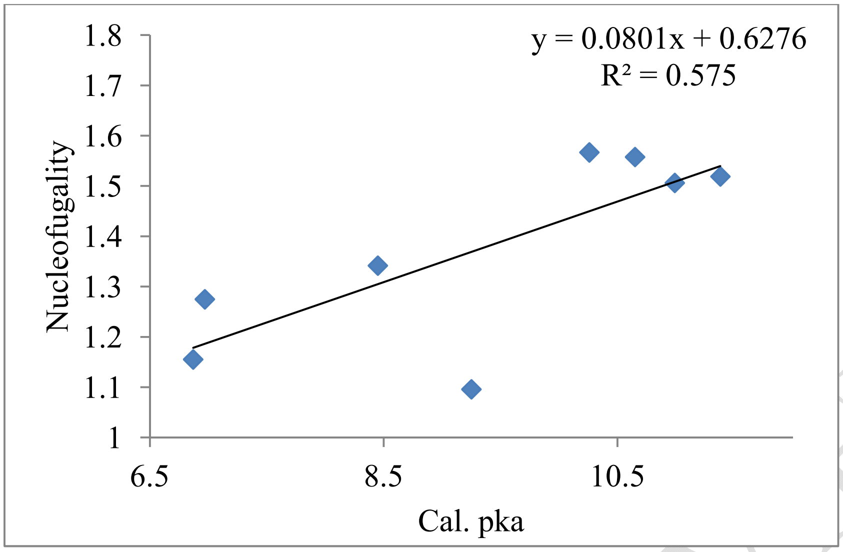


Figure 7: The correlation between calculated pka and Nucleofugality in case of trasposition

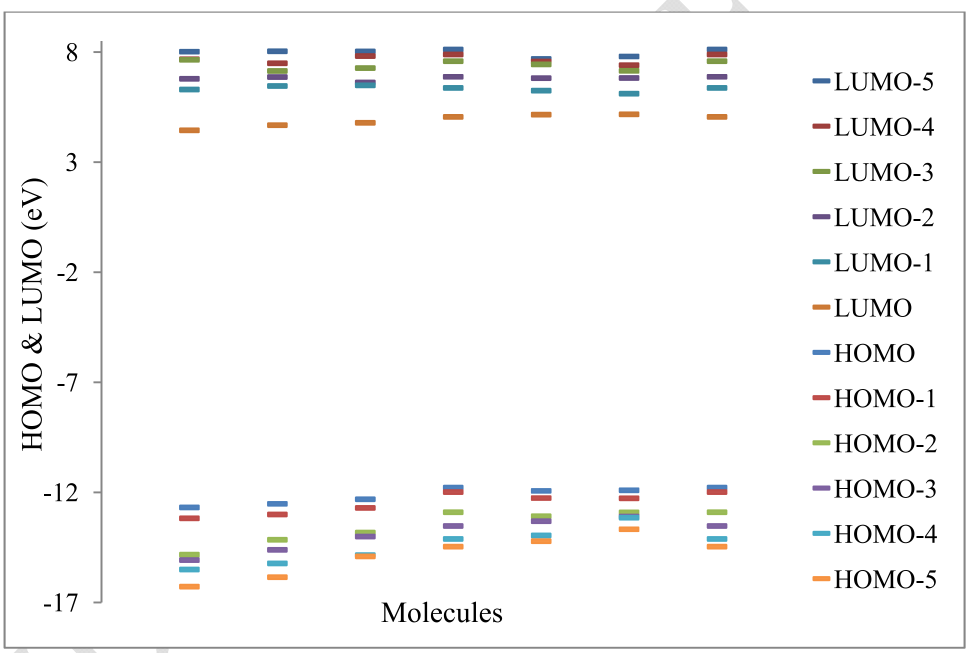


Figure 8: Frontier orbital energies near HOMO and LUMO near for antiperiplanar-molecules

Statistical analysis

To detect whether there is a relationship or effect between the independent variable (Calculated pka) and dependent variables (E_{HOMO}, E_{LOMO}, Hardness, Energy Gap, Softness, Electronegativity, Chemical potential, electrophilicity index(ω), Nucleofugality and Electrfugality), and, which of these variables is more effective, multiple regression analysis has been used, as it is the standard method and used to enter all independent variables not excluding any variables.

Given the correlation matrix between all independent and dependent variables, it has been shown that the variables (E_{HOMO}, E_{LOMO}, Energy Gap, Softness Hardness, Electronegativity and Chemical potential) have no correlation between them, so these variables would be excluded as they were not effective. While the remaining independent variables (electrophilicity index (ω), Nucleofugality, Electrfugality) were associated with the dependent variable, this is in case of syn-periplanar -position molecules.

The correlation coefficient, value between the dependent variable and the independent variables under study was R= 0.999 intermediate value indicating a relationship between these variables. The coefficient of determination R^2 = 0.998, which revealed that, the independent variables were able to explain 100% of the differences and changes in (Cal. pka).

Table 7: the model summary of the statistical analysis in syn-periplanar -position

Model	R	R^2	Adjusted R^2	Std. Error of the Estimate
1	.999[a]	.998	.996	.1205

a. Predictors: (Constant), Electrfugality, Nucleofugality, electrophilicity index(ω).

Anova test showed that there was a very strong relationship between the independent variable and dependent variables, which confirmed the high explanatory power of the statistically multiple linear regression model. From the coefficients table, it can be concluded that the statistically independent variables and T-test at the significant level ($P \leq 0.05$) had no significant effect on the multiple regression model, although there was a correlation between these variables and the independent variable.

Regression equation was obtained using non-standard beta (fixed limit) as follows:

Cal. pka= -3.010+2.720 electrophilicity index(ω)+ 10.259 Nucleofugality-.145 Electrfugality.

In case of antiperiplanarposition, the independent variables (LOMO, Softness, Nucleofugality, Electrfugality) was associated with the dependent variable. The value of the correlation coefficient between the dependent and independent variables under study was R= 0.682 intermediate value and indicated a relationship between these variables. The coefficient of determination R^2 = 0.465 this means that the independent variables were able to explain 47% of the differences and changes in (Cal. pKa).

Table 8: the model summary of the statistical analysis in antiperiplanarposition

Model	R	R^2	Adjusted R^2	Std. Error of the Estimate
1	.682[a]	.465	.363	2.55

a. Predictors: (Constant), Electrfugality, LOMO, Nucleofugality, Softness)

Anova test showed that there was a relationship between the independent variable and dependent variables which confirmed the high explanatory power of the statistically multiple linear regression model. From the coefficients table, it can be concluded that the statistically independent variables and T-test at the significant level ($P \leq 0.05$) had no significant effect on the multiple regression model, although there was a correlation between these variables and the independent variable.

Conclusion

This paper aimed to study the relation of calculated pKa with the experimental for 7 molecules and to study the effect of some descriptors on the above-mentioned molecules and their correlation with calculated pKa. The values of calculated pKa revealed that there was a strong relationship between the calculated and experimental pKa. The calculated values were nearby the expermintal values of syn-periplanar -position molecules, calculation of pKa using other methods could be more close to the experimental. Extensive comparative studies by the other methods, rather than CBS-Q to confirm which is a more acceptable method.

References:

1. Nielsen J. E. & Mccammon J. A. Calculating pKa values in enzyme active sites. Protein Science. 2003;(12):1894–1901.

2. Nyden, M. R.; Petersson, G. A. Complete basis set correlation energies. I. The asymptotic convergence of pair natural orbital expansions. J Chem Phys 1981;(75):1843–1862.

3. Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. A complete basis set model chemistry. I. The total energies of closed□shell atoms and hydrides of the first□row elements. J Chem Phys. 1988;(89):2193–2218.

4. a- Petersson, G. A.; Al-Laham, M. A. A complete basis set model chemistry. II. Open□shell systems and the total energies of the first□row atoms. J Chem Phys 1991;(94):6081–6090.

5. b- Petersson, G. A.; Tenfeldt, T. G.; Mongomery, J. A. A complete basis set model chemistry. III. The complete basis set□quadratic configuration interaction family of methods. J Chem Phys. 1991;(94): 6091–6101.

6. Petersson G. A., Malick D. K., Wilson W. G, Ochterski J. W., Montgomery J. A. and Frisch J. M. Calibration and comparison of the Gaussian-2, complete basis set, and density

functional methods for computational thermochemistry. J. Chem. Phys. 1998;(109): 10570-10579.

7. Casasnovas, R.; Frau J.; Ortega-Castro J.; Salvà A.; Donoso J.; Muñoz F. Absolute and relative pKa calculations of mono and diprotic pyridines by quantum methods, Journal of Molecular Structure: THEO. CHEM. 2009, 912: 5–12.

8. Casasnovas, R.; Fernandez, D.; Ortega-Castro, J.; Frau, J.; Donoso J.; Muñoz F. (). Avoiding gas-phase calculations in theoretical pKa predictions, Theor Chem Acc. 2011 130: 1–13.

9. Khersonsky O; Tawfik DS. Structure-reactivity studies of serum paraoxonase PON1 suggest that its native activity is lactonase. Biochemistry 2005;(44): 6371-6382.

10. Gao, D.Q.; Svoronos, P.; Wong, P.K.; Maddalena, D.; Hwang, J.; Walker, H. pK(a) of acetate in water: A computational study. J. Phys. Chem. 2005;109(47):10776—85.

11. Mongomery, J. A.; Ochterski, J. W.; Petersson, G. A. A complete basis set model chemistry. IV. An improved atomic pair natural orbital method. J Chem Phys 1994;(101): 5900–5909.

12. Montgomery J. A., Frisch M. J., Ochterski J. W., Petersson G. A. A complete basis set model chemistry. VI. Use of density functional geometries and frequencies. J. Chem. Phys. 1999;(110): 2822-2827.

13. Ochterski J. W., Petersson G. Montgomery, J. A. A complete basis set model chemistry. V. Extensions to six or more heavy atoms. J. Chem.Phys. 1998;(104):2598-2619.

14. Liptak M. D. and Shields G. C. Accurate pKa Calculations for Carboxylic Acids Using Complete Basis Set and Gaussian-n Models Combined with CPCM Continuum Solvation Methods, J. Am. Chem. Soc. 2001;(123): 7314-7319.

15. Liptak M.D., Gross K.C., Seybold P.G., Feldgus S. and Shields G.C. Absolute pKa Determinations for Substituted Phenols. J. Am. Chem. Soc. 2002;(124):6421-6427.

16. Aihara, J.; "Reduced HOMO&LUMO Gap as an Index of Kinetic Stability for Polycyclic Aromatic Hydrocarbons". J. Phys. Chem. 1999;(103): 7487–7495.

17. Kim KH, Han YK, Jung J. Basis set effects on relative energies and HOMO–LUMO energy gaps of fullerene C36. Theoretical Chemistry Accounts. 2005;(113): 233–237.

18. Abbaz T., Bendjeddou A. , Villemin D. Molecular structure, HOMO, LUMO, MEP, natural bond orbital analysis of benzo and anthraquinodimethane derivatives. Pharmaceutical and Biological Evaluations, vol. 5:2, pp. 27-39, 2018.

19. Sayiner HS, Abdalrahim AAS., Başaran MA., Kovalishyn V., Kandemirli F. "The Quantum Chemical and QSAR studies on acinetobacter baumannii Oxphos inhibitors." Medicinal Chemistry, vol. 14, pp. 253-268, 2018.

20. Osman ODFT Study of the Structure, Reactivity, Natural Bond Orbital and Hyperpolarizability of Thiazole Azo Dyes. Int J Mol Sci 2017;18.

21. Campod´onico P.R.; Ormaz´abal-T, R.; Aizman A.; Contreras R. “Permanent group effect on nucleofugality in aryl benzoates,” Chemical Physics Letters, 2010;(498): 221–225.

22. Mishra V., Pandeya S. N., Pannecouque C., Witvrouw M. and De Clercq E. Anti-HIV Activity of Thiosemicarbazone and Semicarbazone Derivatives of (±)-3-Menthone. Arch. Pharm. 2002;335(5):183-186.