The Islamic University Journal (Series of Natural Studies and Engineering)
Vol.18, No.1, pp 43 - 60 , **2010**, ISSN 1726-6807, http://www.iugaza.edu.ps/ara/research/

A Computational Study of Toxicity of Nitrobenzenes Using QSPR and DFT-Based Molecular Surface Electrostatic Potential

Fakhr M. Abu-Awwad
Chemistry Department, Islamic University of Gaza, Box 108 – Gaza, Palestinian Authority
dr-fakhr@live.com

Abstract: In the present study, the density functional B3LYP/6-311G** level of theory was used to compute and map the molecular surface electrostatic potentials of a group of substituted nitrobenzenes to identify common features related to their subsequent toxicities. Several statistical properties including potentials' extrema (V_{min}, V_{max}), molecular volume, surface area, polar surface area, along with different energies were computed. A little linear correlation was revealed between V_{min} and surface area, and systems' toxicities. Another computations employed quantitative structure–property relationships model in CODESSA package to correlate toxicities with calculated descriptors. Statistically, the most significant correlation is a five-parameter equation with correlation coefficient, R^2 values of 0.962, and the cross-validated correlation coefficient, R^2_{CV}=0.950. The obtained models allowed us to reveal toxic activity of nitrobenzenes.

Keywords: toxicity; nitrobenzene; surface potential; B3LYP; Codessa; QSAR.

دراسة حاسوبية لسمية مركبات النيتروبنزين باستخدام كل من العلاقة الكمية بين التركيب والنشاط، والجهد الالكتروستاتيكي على السطوح الجزيئية المشتق من نظرية دالة الكثافة

ملخص: في هذه الدراسة، تستخدم دالة الكثافة عند B3LYP/6-311G** لحساب وتخطيط الجهد الكهروستاتيكي على السطوح الجزيئية لمجموعة من مركبات النيتروبنزين المستبدلة بهدف تحديد السمات المشتركة التي قد تعزى لسميتها. لقد تم احتساب العديد من الخصائص الإحصائية على السطوح ومنها القيم الصغرى والكبري للجهد (V_{min}, V_{max})، وحجم الجزيئ، المساحة السطحية، المساحة السطحية القطبية، بالإضافة إلى طاقات مختلفة. لقد وجدت علاقة خطية مقبولة بين كل من القيمة الصغرى للجهد ومساحة السطح من ناحية، وسمية المركبات من ناحية أخرى. وكذلك، أجريت حسابات أخرى باستخدام العلاقة الكمية بين التركيب والنشاط المتوفرة في برنامج CODESSA للربط بين السمية والمتغيرات المحسوبة. من الناحية الإحصائية، لقد تم التوصل لمعادلات خطية ذات خمسة متغيرات لها معامل ارتباط R^2 يساوي 0.962 ومعامل تصديق

الارتباط R^2_{CV} يساوي 0.950. إن المعادلات التي تم الحصول عليها ذات فائدة في الكشف عن النشاط السمي لمركبات النيتروبنزين.

Introduction

Computational methods have been widely used to predict physicochemical properties in fields of physical organic chemistry and drug design.[1-3] Several approaches of quantitative structure–activity relationship (QSAR) have been used to correlate experimental biological activities or binding constants with various computed descriptors.[4-6] For quantitative structure-property relationship (QSPR), physicochemical properties are correlated with structural and/or electronic descriptors.[7-9]

As an example of nitro-substituted aromatics, nitrobenzenes are released into the biosphere when produced by incomplete combustion of fossil fuels or used as synthetic intermediates, pesticides, and explosives. As hazardous chemicals of a high potential for environmental pollution, nitrobenzenes commonly show symptoms of toxicity including immunotoxicity and a conjectured carcinogenicity.[10] A common measure of their toxic activities is LD_{50} dose (mg/kg) for rats, which is equivalent to molar $-\log(LD_{50})$ response variables.[11] The di- substituted nitrobenzenes show higher electrophilic reactive toxicity than the mono-substituted, where halogenated nitrobenzenes are metabolized in vivo to such strong electrophiles. Along the same line, some nitrophenols are working as oxidative phosphorylation uncouplers. It is obvious that the nature and degree of the aromatic substitution have a key effect on the toxicity of the nitroaromatic compounds.[12]

For many years both chemists and environmental scientists have been employing QSPR and QSAR to investigate and predict toxicity of nitrobenzenes.[13-14] The hydrophobicity and electrophilicity were employed to investigate the structural and functional features behind that toxicity.[15] Several computational models of different number of parameters were obtained including constitutional, topological, geometrical, electrostatic, and/or quantum chemical descriptors.[16-17] Constitutional descriptors are related to the number of atoms and bonds in a molecule, while topological descriptors include valence and non-valence molecular connectivity indices concluded from size, composition, and degree of branching of the molecule. While geometrical descriptors are calculated from 3D atomic molecular coordinates, and electrostatic descriptors reflect characteristics of the charge distribution of the molecule, quantum-chemical descriptors are mostly derived from the partial charge distribution in a molecule or from the electron densities on particular atoms.

A Computational Study of Toxicity of Nitrobenzenes Using

In the last few years several papers utilized the density functional B3LYP at the basis set 6-311G** to investigate the toxicity of niroaromatics toxicity to specific biological species. Xiao et al concluded that for the golden orfe fish, the nitro group is the main toxic group for both nitrobenzenes and nitroanilines and wrapping or reducing the nitro groups decreases the toxicity of the subject chemicals.[18] Also, Xiao correlated the toxicity of mono- and di-nitroaromatics to both the algae and fathead minnow with the hydrophobicity, the energy of frontier molecular orbitals, and the charge of the nitro group.[19,20] Leszczynski, utilized the toxicity LD_{50} parameter for rats and found that the toxicity of nitroaromatic compounds appears to be governed by a number of factors, such as the number of nitrogroups, the electrotopological state, the presence of certain fragments and the electrophilicity/reactivity parameter.[21] In his comparative research, Zhang et al employed three different theoretical approaches to better seek a QSAR modeling.[22] Obviously, the unique part of this study is employing the computed electrostatic potential-based descriptors along with CODESSA – based descriptors to generate the sought QSAR model.

Molecular Surface Electrostatic Potentials (MSEP)

The electrostatic potential V(**r**) is a real physical property that has been successfully employed experimentally and computationally to explore chemical reactivity in various types of chemical, physical, and biological applications.[23-25] As revealed by equation (1), V(**r**) measures the cumulative charge impact of both nuclei and electrons in the surrounding space of the concerned system. Z_A is the charge on nucleus *A*, located at $\mathbf{R}_A$.[26]

$$V(r) = \sum_A \frac{Z_A}{|R_A - r|} - \int \frac{r(r')dr'}{|r' - r|}$$

(1)

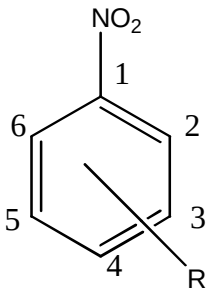
In a series of studies, several properties of noncovalent interacting systems such as heats of phase transitions, solubilities, boiling points, viscosities, and surface tensions are expressed analytically in terms of statistically defined quantities that characterize MSEP.[27-28] In identifying quantities the electrostatic potential $V_{S(\mathbf{r})}$ on a molecular surface, both extrema are characterized; the most positive value, $V_{S,max}$ and the most negative, $V_{S,min}$ in the space of the molecule, where each V_{min} is a site-specific quantity.

Though the units of the electrostatic potential should be energy/charge, $V_S(r)$, $V_{S,min}$ and $V_{S,max}$ are given as simply an energy, where each can be

interpreted as the interaction energy of the system with a unit positive charge placed at the point **r**.

The Polar Surface Area (PSA) is defined as the surface sum over all polar atoms, usually oxygen and nitrogen, including also attached hydrogens. Molecules with a polar surface area of greater than 140 A^2 are usually poor at permeating cell membranes. For molecules to penetrate the blood-brain barrier, and thus acting on receptors in the central nervous system, PSA should be less than 60 A^2.[29]

In this study, the toxicity of several mono- and di-substituted nitrobenzene derivatives is investigated by both electronic calculations and QSPR. The computed MSEPs nitrobenzenes are mapped to identify any common features that are related to their subsequent toxicities. Hence, the concern here is not with the potential throughout the space around a molecule but rather with its pattern on the molecule's surface. Therefore, our focus shall be primarily upon the most positive and most negative values of $V_{S(r)}$, the $V_{S,max}$ and the $V_{S,min}$, respectively. In some instances, we will also refer to the overall most negative potentials (not limited to the molecule's surface), V_{min}, that are associated with nitrogen and oxygen lone pairs.



Scheme 1: Parent structure of mono- and di- substituted nitrobenzenes

Computational Methodology

1- Potential surface calculations:

Molecular geometries of 21 structures (Scheme 1) were fully optimized using Gaussian 03 for windows packages[30] at the density functional theory (DFT) with Becke's three-parameter hybrid functional using the LYP correlation functional (B3LYP) [31] and the standard Pople's 6-311G(d,p) basis set. The optimized structures were properly attributed to their local minima where the matrices of the energy second derivatives were checked at the same level of theory to have zero imaginary values. Electrostatic potential, local charges at each atom, dipole moment, HOMO and LUMO energies were calculated for each of the compounds. Then, electrostatic

potentials on the optimized molecular surfaces were mapped using eq. (1) at the same level of theory utilizing Spartan 08 package for windows.[32] The values of the $V_{S,min}$, $V_{S,max}$, surface area ($Å^2$), molecular volume ($Å^3$) and the polar surface area, PSA ($Å^2$), were calculated.

2- CODESSA calculations:

The optimized structures were analyzed by the CODESSA program[33], where about 300 descriptors including constitutional, topological, geometrical, electrostatic, and quantum mechanical were exploited. The CODESSA heuristic method (HM) was applied to the whole dataset of 21 nitrobenzenes seeking a pre-selection step for the many available descriptors and to select the rough starting regression models. Thus good estimates about the quality of expected correlation could be quickly given through identifying both the more or less significant descriptors from the standpoint of a single-parameter correlation, and then the highly inter-correlated descriptors. This information help much in optimizing the number of descriptors involved in the subsequent step of model development. Accordingly, HM was applied separately once again to the two datasets of 12 mono- and 9 di- nitrobenzenes.

Starting from computing all possible one-parameter regression models, a stepwise addition of descriptors is examined to find the best multiparameter regression models with optimal values of the statistical parameters including highest values of R^2, the cross-validated R^2_{cv}, Fisher F-criterion value and s^2, the standard deviation of the regression. Descriptors for which values could not be calculated and/or descriptors of low variance in each dataset were discarded.

Results and Discussion

The toxic activity data was obtained from the literature for the 21 studied compounds.[12] They are listed in Table 1 along with the computed properties including HOMO, LUMO, dipole moments μ, molecular volume A^3, polar and surface area PSA, A^2, $V_{S,min}$, and $V_{S,max}$, along with several molecular surface based statistically derived descriptors. Following Bader *et al*.[34] in taking the surface to be a particular outer contour of the molecule's electronic density, figure 1 (1(a), 1(b), 1(c), and 1(d)) shows examples of electrostatic potentials on $\rho(r) = 0.002$ au molecular surfaces of different nitrobenzenes; (1), (5) (13) and (21) subsequently.

Table 1: Calculated Statistical related properties to molecular Surface Electrostatic Potential of substituted nitrobenzene:

Toxicity **(**LD_{50} dose), surface area ($Å^2$), molecular volume ($Å^3$), minimum potential ($V_{S,min}$), maximum potential ($V_{S,max}$), energy of highest occupied orbital (HOMO), energy of lowest occupied orbital (LUMO), the energy gap (HOMO-LUMO), and dipole moment (debye)

Sys	**R**	**toxicity**	**area**	**volume**	$\mathbf{V_{min}}$	$\mathbf{V_{max}}$	**PSA**	$\mathbf{e_{HOMO}}$	$\mathbf{e_{LUMO}}$	$\mathbf{e_{HOMO-LUMO}}$	**dipole**
1	1-NO_2	-2.778	137.630	122.430	-157.922	124.918	30.210	-7.8230	-2.6327	5.4010	2.0180
2	2- CH3	-2.950	154.120	141.280	-159.658	118.949	24.650	-7.4890	-2.4861	5.0031	4.2629
3	4-CH3	-3.331	157.360	142.180	-164.348	117.558	28.130	-7.5733	-2.5095	5.0638	5.2005
4	2-OH	-2.524	144.090	130.120	-139.989	155.077	21.520	-7.0260	-2.9097	6.2652	1.9748
5	3-OH	-2.516	146.720	131.330	-154.607	308.996	33.120	-7.0258	-2.6586	4.3672	3.5553
6	4-OH	-2.305	146.720	131.280	-167.797	323.267	40.610	-7.1433	-2.4191	4.7242	5.3072
7	2-Cl	-2.428	152.390	139.820	-154.357	140.796	38.220	-7.6220	-2.7797	4.8423	5.0364
8	3-Cl	-2.591	155.320	140.530	-142.944	140.540	28.650	-7.6827	-2.9029	4.7797	3.8336
9	4-Cl	-2.623	155.240	140.500	-144.740	136.682	29.390	-7.7787	-2.8616	5.1854	6.3823
10	3-COOH	-2.833	167.110	152.510	-145.442	289.685	83.630	-8.2198	-3.0344	4.9925	2.5476
11	4-COOH	-3.292	167.030	152.520	-142.968	291.356	33.860	-8.2742	-3.2817	5.0913	3.8132
12	4-CH_2Cl	-3.257	175.560	160.850	-148.269	141.386	43.600	-7.9768	-2.8855	5.0913	3.8132
13	2-, 6-CH_3	-3.301	168.850	159.910	-164.014	115.421	23.220	-7.2195	-2.3032	4.9163	3.6499
14	3-, 5-CH3	-3.387	173.390	161.060	-165.539	114.924	24.240	-7.1730	-2.4564	4.7166	4.4927
15	3-, 4-Cl	-2.808	170.620	158.230	-135.061	151.502	22.650	-7.7901	-3.0610	4.7291	2.6400
16	3-NO_2	-1.919	162.900	148.060	-128.004	162.190	39.750	-8.6201	-3.3166	5.3035	4.2107

17	2- CH_3 3- NO_2	-2.255	175.010	165.520	-131.130	151.121	34.100	-8.1009	-3.0254	5.0755	2.9381
18	3- NO_2 5- CH_3	-2.428	179.450	166.890	-136.737	156.360	41.310	-8.3063	-3.1623	5.1441	4.8780
19	3- NO_2 4- OH	-1.851	169.400	155.770	-136.440	207.149	34.930	-7.8318	-3.4967	4.3351	3.3981
20	3- NO_2 4- F	-1.699	167.340	152.280	-125.523	177.084	41.870	-8.5768	-3.3598	5.2170	3.8634
21	3- NO_2 5- NO_2	-2.439	188.600	173.980	-100.647	158.997	66.270	-9.1605	-3.8502	5.3103	0.0015

The mapped 3D-isosurfaces represent the superimposed electrostatic potentials in the electronic density. Obviously, the code-colored value of each isosurface identifies the molecular volume and both $V_{S,min}$ and $V_{S,max}$. The most negative potential is colored red, while the most positive potential is colored blue. Intermediate potentials are assigned colors according to the color spectrum red < orange < yellow < green < blue.[35] In general, the most positive potentials of blue regions are associated with nitro nitrogens and ring hydrogens, while the most negative potential of the red regions are due to nitro oxygens, nitrogen lone pairs, and halogens' substituents.

Though the aromatic ring is a key structural constituent in all instances, there are substantial differences in the listed properties. There is a considerable range of sizes; the surface areas being between 137 and 188 A^2 and the molecular volumes being between 122 and 174 A^3.

The positive regions of the surface electrostatic potentials of these molecules provide further contrasts. As mentioned above, the strongest positive potentials, with $V_{S,max}$ between 114.92 and 323.27 kJ/mole, are produced by hydroxyl, or ring hydrogens. However there is no correlation between number of available hydrogens and their molecules subsequent $V_{S,max}$, indicating that the positive regions on their surfaces are relatively weak.

On the other hand, the negative surface regions, while less extensive in area, are much more uniform in strength. The $V_{S,min}$ are all within a relatively narrow range, –100.65 to –167.80 kJ/mole, It seems realistic to conclude that the negative potentials are of primary importance in toxic activity of nitrobenzenes.

A particularly interesting point of similarity among the molecules in Table 1 is the polar surface area, PSA, which could be a measure of local polarity. It varies between 21.52 A^2 and 83.63 A^2; most often, however, it is less than 40 A^2. What is notable in Table 1 is that 16 of the 21 PSA values

are between 21 and 40 A^2. Thus the internal charge separations in these molecules are quite significant, but are rather strictly circumscribed in magnitude. This suggests a need for a substantial but degree of hydrophilic character.

Linear regression analysis was used to estimate the possible correlation of nitrobenzenes toxicities with the previous quantities using SPSS for Windows[36], where both $V_{S,min}$ and molecular surface area showed minor correlation with toxicity with R^2 about 0.5.

Employing HM of Codessa QSAR analysis in the study, The results shown in Table 2 for the 21 molecules are rather amazing, where five types of indices has evicted in the correlated models, but not any geometrical indices has driven out. The five molecular descriptors (two quantum mechanical, one electrostatic, one constitutional, and one topological) involved in the selected model are: *the relative number of aromatic bonds (RNAB)*, a constitutional descriptor, which is correlated with the density of electron cloud; *Min electroph. react. index for a C atom*, a quantum-chemical descriptor; *Information content (order 2)*, a topological indexes which reflects the branching of the molecule and reflects how information rich the molecule is; *PNSA-1 Partial negative surface area [Zefirov's PC]*, an electrostatic descriptor, which is the sum of the surface area of negative atoms; and *RNCS Relative negative charged surface area (SAMNEG*RNCG) [Zefirov's PC]*, a quantum-chemical descriptor, which depends directly on the quantum-chemically calculated charge.

Table 2: Calculated Codessa based descriptors:

RNBR: Relative number of aromatic bonds; MERICA: Min electroph. react. index for a C atom; IC2: Information content (order 2); PNSA-1: Partial negative surface area [Zefirov's PC]; RNCS Relative negative charged SA (SAMNEG*RNCG) [Zefirov's PC]

Sys	RNCS	PNSA-1	IC2	MERICA	RNBR
1	25.2437	218.6776	32.9384	6.79E-05	0.5714
2	22.2271	190.4189	52.7320	1.01E-03	0.4706
3	22.1231	212.7824	50.7320	1.13E-03	0.4706
4	17.8896	196.4983	46.6034	3.65E-04	0.5333
5	17.9673	208.3365	48.6034	1.70E-03	0.5333
6	17.8586	165.7361	44.6034	7.80E-04	0.5333
7	21.2246	218.5963	41.3030	5.47E-04	0.5714

8	29.9616	222.0062	43.3030	1.81E-03	0.5714
9	30.4718	203.8272	39.3030	1.40E-03	0.5714
10	13.2692	245.3041	59.4869	1.13E-04	0.4706
11	15.0281	250.3032	55.4869	2.29E-04	0.4706
12	29.2992	226.0356	53.4869	3.62E-04	0.4706
13	15.3447	165.8929	58.1739	8.13E-05	0.4000
14	15.5293	182.3019	60.1739	1.18E-03	0.4000
15	18.5910	268.3116	42.5481	7.05E-05	0.5714
16	16.4056	286.188	42.0000	3.30E-03	0.6250
17	13.0684	273.1216	57.2008	1.00E-04	0.5263
18	15.4146	277.2625	61.2008	1.01E-04	0.5263
19	12.9256	295.2116	54.7320	1.09E-04	0.5882
20	12.5492	297.5677	49.2451	4.60E-04	0.6250
21	10.9574	408.0962	40.5293	3.94E-03	0.6667

Five correlated equations of the best one-, two-, three-, four-, and five-parameters were selected as models depicted in Table 3. In these models, the correlation coefficient, R^2, measures the fit of the regression equation, while, F, the Fisher test value, reflects the ratio of the variance explained by the model and the variance due to the error in it. s^2 is the standard deviation of the regression. The maximum number of descriptors used was set to 5, so that the ratio between the number of descriptor exploited and the available known molecules was about 1:5, as previously suggested for correlations employing multiple linear regressions (MLR).[37] A higher ratio may lead to over-correlated equations of poorly reliable predictions. In the first step, only models for which $R^2 > 0.6$ were considered for further selection, followed by a subsequent validation through discarding several models from the initial ones. Lastly, the simplest among these last validated models is selected as the most suitable model of correlation.

Table 3: Regression parameters and statistical quality of the correlations of the toxicity of nitrobenzenes (LD_{50} dose) in the present study

Models	**Descriptors involved**	**t-test**	***B* (intercept)**	Statistical Parameters			
				R^2	R^2_{cv}	F	S^2
Eq. 1	Relative number of aromatic bonds	5.77	-10.81	0.637	0.547	33.27	0.094
Eq. 2	Relative number of single bonds Min electroph. react. index for a C atom	-9.46 -5.58	2.04	0.845	0.805	48.97	0.042
Eq. 3	Relative number of aromatic bonds Min electroph. react. index for a C atom Structural Information content (order 2)	12.08 -5.29 4.31	-12.62	0.915	0.873	61.08	0.025
Eq. 4	Relative number of aromatic bonds Min electroph. react. index for a C atom Information content (order 2) PNSA-1 Partial negative surface area [Zefirov's PC]	10.27 -2.512 6.32 -3.83	-13.50	0.953	0.915	81.07	0.014
Eq. 5	Relative number of aromatic bonds Min electroph. react. index for a C atom Information content (order 2) PNSA-1 Partial negative surface area [Zefirov's PC] RNCS Relative negative charged SA (SAMNEG*RNCG) [Zefirov's PC]	12.78 -12.13 6.90 -5.50 -3.03	-14.50	0.971	0.943	99.75	0.010

The HM based one-parameter regression expression is as in eq. (2), where RNAB is the relative number of aromatic bonds:

$$\text{Toxicity} = 5.586*\text{RNAB} - 5.614 \quad (2)$$

$$N = 21;\ R^2 = 0.636;\ R^2_{cv} = 0.547;\ F=33.27;\ s2 = 0.0939$$

Here and thereafter, N is the number of compounds; R^2 is the correlation coefficient, R^2_{CV} is the 'leave one out' (LOO) cross-validated coefficient, F is the Fisher-statistic value, and s^2 is the standard deviation of the regression equation. In the above equation, a constitutional descriptor RNAB has a positive-sign coefficient, implying that increasing the magnitude of RNBR would favor the exhibitions of the toxic activity of nitrobenzenes.

Among the obtained two-parameter models, statistically the best one is as follows:

$$\text{Toxicity} = 6.307*\text{RNAB} - 102.429*\text{MERICA} - 5.906 \quad (3)$$

$$N = 21;\ R^2 = 0.675;\ R^2_{cv} = 0.805;\ F=18.65;\ s2 = 0.0423$$

In this model, MERICA is a quantum mechanical descriptor measures the minimum electrophilic reaction index for a C atom in each molecule. A gain RNAB has a coefficient with positive sign while MERICA has a coefficient with a negative sign. According to the models, a further substitution of the benzene ring lowers the toxic activity of the compounds.

Among the obtained three-parameter models, statistically the best one is as follows:

$$\text{Toxicity} = 8.312*\text{RNAB} - 78.891*\text{MERICA} + 0.0282*\text{SIC2} - 8.382 \quad (4)$$

$$N = 21;\ R^2 = 0.783;\ R^2_{cv} = 0.745;\ F=61.077;\ s2 = 0.0245$$

In this model, SIC2 is a topological descriptor stands for structural information content (order 2), which describes the atomic connectivity and branching information of the molecules While RNAB and MERICA have subsequently positive and negative coefficients, SIC2 has a coefficient with a positive sign. This points out that an increase in the magnitude of SIC2 favors the exhibitions of the toxic activity of the compounds.

Among the obtained four-parameter models, the best one is as shown below:

$$\text{Toxicity} = 13.884*\text{RNAB} - 7.742*\text{MERICA} + 0.0607*\text{SIC2} - 0.0066*\text{PNSA} - 11.433 \quad (5)$$

$$N = 21;\ R^2 = 0.935;\ R^2_{cv} = 0.918;\ F=81.017;\ s2 = 0.0144$$

In this model, PNSA is an electrostatic descriptor which refers to partial negative surface area [Zefirov's PC] known as PNSA-1. The negative sign of its coefficient highlights a decrease in the magnitude of PNSA favors the exhibitions of toxic activity of nitrobenzenes.

During the regression analysis using HM, several five-parameter equations was obtained. Out of these equations, equation (6) in the model below consists of RNCS along with the previously correlated descriptors. RNCS is the relative negative charged surface area (SAMNEG*RNCG) [Zefirov's PC].

$$\text{Toxicity} = 13.478*RNAB - 7.428*MERICA + 0.0514*SIC2 - 0.0074*PNSA - 0.0193*RNCS - 10.218 \quad (6)$$

$$N=21;\ R^2=0.962;\ R^2_{CV}=0.950;\ F=99.75;\ s2=0.096.$$

In this model, the sign of coefficients is the same as in that in the previous models and thus they carry the same significance. RNCS is an electrostatic descriptor. It has a negative sign of coefficient, meaning that the decrease in the magnitude of RNCS is the favorable parameter for the exhibition of the toxic activity of the compounds.

Lastly, a correlation was sought between experimental toxicity of nitrobenzenes from one side, and from the other side, a combination of Spartan based V_{min} and molecular surface area and the six Codessa based descriptors. The *enter* method of linear regression in SPSS was employed to seek any possible correlation. The obtained seven-parameter model was as follows:

$$\text{Toxicity} = 13.685*RNAB - 6.959*MERICA + 0.0518*SIC2 - 0.0071*PNSA - 0.0192*RNCS - 0.0019*V_{min} + 0.00058*Area - 10.781 \quad (7)$$

$$N = 21;\ R^2 = 0.963;\ R^2_{cv} = 0.943;\ F=47.87;\ s2 = 0.0245$$

Obviously, as concluded before, both V_{min} and molecular area have insignificant impact on the linear regression with toxicities of our systems. The negative and positive coefficients of both quantities point at their subsequent inverse and direct proportional relation with toxicity.

Conclusion:

In the present study, the descriptors of 21 toxic nitrobenzenes have been correlated with their toxicity, (LD_{50} dose) using Spartan 8, Gaussian 03, and Codessa package. The most negative potential, $V_{S,min}$ in the space of the molecule and the its surface area showed little significance in the toxic activity. Five-parameter equation has been attained that consist of the descriptors; relative number of aromatic bonds, minimum electrophilic reaction index for a C atom, information content (order 2), (PNSA-1) partial negative surface area [Zefirov's PC], and RNCS relative negative charged SA (SAMNEG*RNCG) [Zefirov's PC]. The statistical parameters of this model are the $R^2=0.962$; $R^2_{CV}=0.950$; $F=99.75$; $s^2=0.096$. The equation as

concluded from tables (4, 5) is working properly to predict the toxicity of nitrobenzenes.

Table 4: Correlation matrix for the inter-correlation of various molecular descriptors involved in the obtained models.

	RNAB	**MERICA**	**IC2**	**PNSA-1**	**RNCS**
RNBR	1.0000				
MERIC	-0.0634	1.0000			
IC2	-0.2904	-0.7210	1.0000		
PNSA-1	0.6994	-0.2903	0.3184	1.0000	
RNCS	0.1863	-0.4388	0.1722	0.0924	1.0000

RNBR: Relative number of aromatic bonds; MERICA: Min electroph. react. index for a C atom; IC2: Information content (order 2); PNSA-1: Partial negative surface area [Zefirov's PC]; RNCS Relative negative charged SA (SAMNEG*RNCG) [Zefirov's PC].

Table 5: Experimental and calculated toxicities of nitrobenzenes using five-[1], and four-[2] parameter correlations in eqs. (5) and (6).

Sys	Exp. toxicity	Calc. toxicity[1]	Diff[1]	Calc. Toxicity[2]	Diff[2]
1	-2.7780	-2.8862	-0.1082	-2.8873	-0.1093
2	-2.9500	-2.9922	-0.0422	-2.9599	-0.0099
3	-3.3310	-3.1764	0.1546	-3.1140	0.2170
4	-2.5240	-2.4333	0.0907	-2.4901	0.0339
5	-2.5160	-2.4388	0.0772	-2.4809	0.0351
6	-2.3050	-2.3390	-0.0340	-2.4408	-0.1358
7	-2.4280	-2.4588	-0.0308	-2.4933	-0.0653
8	-2.5910	-2.5352	0.0558	-2.4172	0.1738
9	-2.6230	-2.6153	0.0077	-2.5282	0.0948
10	-2.8330	-2.8630	-0.0300	-2.8787	-0.0457
11	-3.2920	-3.3002	-0.0082	-3.3780	-0.0860
12	-3.2570	-3.3745	-0.1175	-3.2184	0.0386
13	-3.3010	-3.3431	-0.0421	-3.4093	-0.1083
14	-3.3870	-3.3654	0.0216	-3.4052	-0.0182
15	-2.8080	-2.6584	0.1496	-2.6729	0.1351
16	-1.9190	-2.0027	-0.0837	-2.0062	-0.0872
17	-2.2550	-2.4009	-0.1459	-2.3901	-0.1351
18	-2.4280	-2.3035	0.1245	-2.2397	0.1883
19	-1.8510	-1.8973	-0.0464	-1.8909	-0.0399
20	-1.6990	-1.7186	-0.0196	-1.7532	-0.0542
21	-2.4390	-2.4120	0.0270	-2.4609	-0.0219
St. deviation*			**0.0001**		**0.0002**

***The tiny standard deviations reflect the importance of both the two models in calculating nitrobenzenes toxicities according to eqs. (5) or (6).

References:

[1] Kim C.K., Lee K.A, Hyun K. H., Park H. J., Kwack I.Y., Kim C.K., Lee H.W., Lee B.S., J. Comp. Chem., 25 (2004) 2073.
[2] Monasterios M., Avendano M., Amaro M.I., Infante W., Charris J., J. Molec. Struct. 798 (2006) 102.
[3] Chai X., Zhang J., Yu S., Hu H., Zou Y., Zhao Q., Dan Z., Zhang D., Wu Q., Bioorg. Med. Chem. Lett., 19 (2009) 1811.
[4] Moldovan D.C., Costescu A., Katona G., Diudea V.M., J. Math. Chem., 45 (2009) 442.
[5] Sahu V.K., Khan A.K.R., Singh R.K., Singh P.P., Int. J. quant. chem., 109 (2009) 1243.
[6] Tetko I.V., Jaroszewicz I., Platts J.A., Kuduk-J J., J. Inorg. Biochem., 102 (2008) 1424.
[7] Xu H.Y., Zou J.W., Yu Q.S., Wang Y.H., Zhang J.Y., Jin H.X., J. Chemosph., 66 (2007) 1998.
[8] Cheng Y.Y., Yuan H., J. Molec. Graph. Mod., 24 (2006) 219.
[9] Mager D.E., Adv. Drug Del. Rev., 58 (2006) 1326.
[10] Niazi A., Bozorghi J.S., Nori-Shargh D., J. Hazard. Materials, 151 (2008) 603.
[11] Dearden J.C., Cronin M.T.D., Lin T.W., Lin D.T., Quant. Struct. Act. Relat., 14 (1995) 427.
[12] Katrizky A.R., Oliferenko P., Oliferenko A., Lomaka A., Karelson M., J. Phys. Org. Chem. 16 (2003) 811.
[13] Lozano S., Lescot E., Halm M.P., Lepailleur A., Bureau R., Rault S., J. Enzyme Inhib. Med. Chem. (2009)
[14] Toropov, A., Rasulev B., Leszczynski J., QSAR & Comb. Sc., 26 (2007) 686.
[15] Wang X., Yin, C., Wang L., J. Chemosphere, 46 (2002) 1045.
[16] Agrawal V.K., Khadikar P.V., Bioorg. Med. Chem. 9 (2001) 3035.
[17] Warne M.A., Osborn D., Lindon J.C., Nicholson J.K., Chemosphere, 38 (1999) 3357.
[18] Yan X.F., Shu Y.J., Wang L.J., Xiao H.M., Acta Chimica Sinica, 65-17 (2007) 1789.
[19] Yan X.F., Xiao H.M., Gong X.D., Ju C.H., Chemosphere, 59-4 (2005) 467.
[20] Yan X.F., Xiao H.M., Ju C.H., Gong X.D., Chinese J. Chemistry, 23-8 (2005) 947.
[21] Isayev O., Rasulev B., Gorb L., Leszezynski J., Mol. Diversity, 10-2 (2006) 233.

[22] Cui X.J., Zhang Z.Y., Yuan X., Zhang J.P., Liu S.D., Guo L.P., Harrington P.D., Chem. Res. Chinese Universities, 22-4 (2006) 439.
[23] Toro-Labbe A., Gutie´rrez-Oliva S., Concha M.C., Murray J.S., Politzer P., J. chem. phys., 121 (2004) 4570.
[24] Joshi K.A., Patil D.D., Gejji S.P., J. Molec. Mod. 15 (2009) 383.
[25] Yuan P.Q., Kong N., Cheng Z.M., Semiat R., Desalination, 238 (2009) 246.
[26] Politzer P., Murray J.S., in: Murray J.S., Sen K.D. (Eds.), Molecular Electrostatic Potentials: Concepts and Applications, Elsevier, Amsterdam, (1996) 649.
[27] Politzer P, Murray J.S., Abu-Awwad F.M., Int. J. Quant. Chem., 76 (2000) 643.
[28] Abu-Awwad F.M., Chem. Phys. Lett., 360 (2002) 340.
[29] Ertl, P. Polar Surface Area, in Molecular Drug Properties, R. Mannhold (ed), Wiley-VCH, (2007) 111.
[30] Frisch M.J. et al.; *Gaussian 03*, revision E.01; Gaussian, Inc.: Pittsburgh, PA, 2003
[31] Becke A.D., J. Chem. Phys. 98 (1993) 5648.
[32] Spartan version 8.0 from Wavefunction**,** Inc., Irvine, CA**.** USA
[33] CODESSA$_{TM}$, References Manual, V. 2.7.10 (PC)., Semichem, 7204, Mullen, Shawnee, KS, USA, Copyright Semichem and the University of Florida, 2002.
[34] Bader R.F.W., Carroll M., Cheeseman J.R., Chang C., J. Am. Chem. Soc., 109 (1987) 7968.
[35] Spartan Physical Chemistry Edition: Tutorial and Activities, Wavefunction Inc., Japan Branch Office, (2005) 81.
[36] SPSS for Windows – Professional Statistics. Release 10.0, Chicago: SPSS Inc.
[37] Katritzky A.R., Lobanov V.S., Karelson M., CODESSA: Ref. Manual; Ver. 2, University of Florida, 1994.

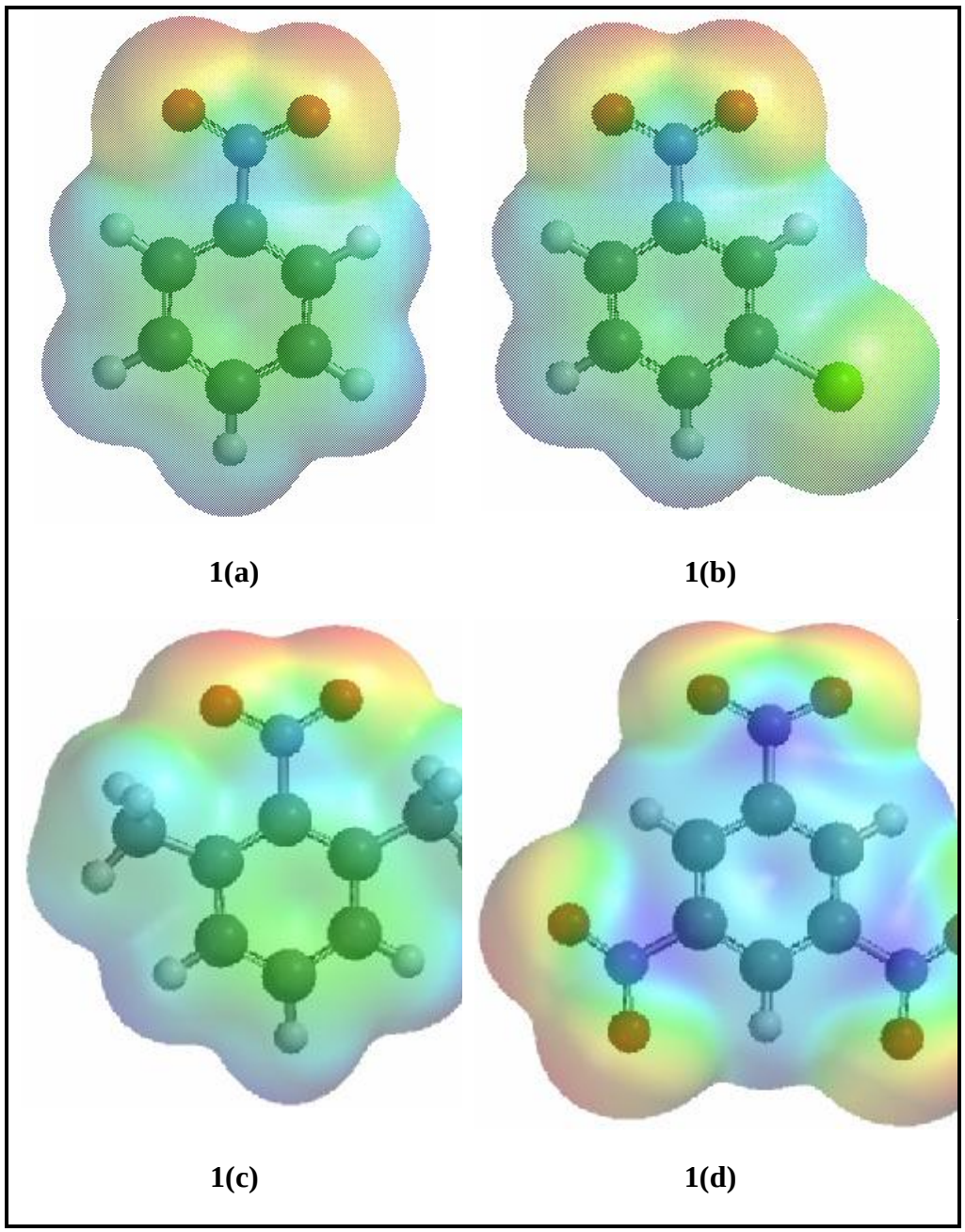


Figure 1. Computed electrostatic potential on the molecular surface of nitrobenzenes (1, 8, 13, 21). The potential ranges, in kJ/mole, according to the color code: red (most negative) < orange < yellow < green < blue (most positive).

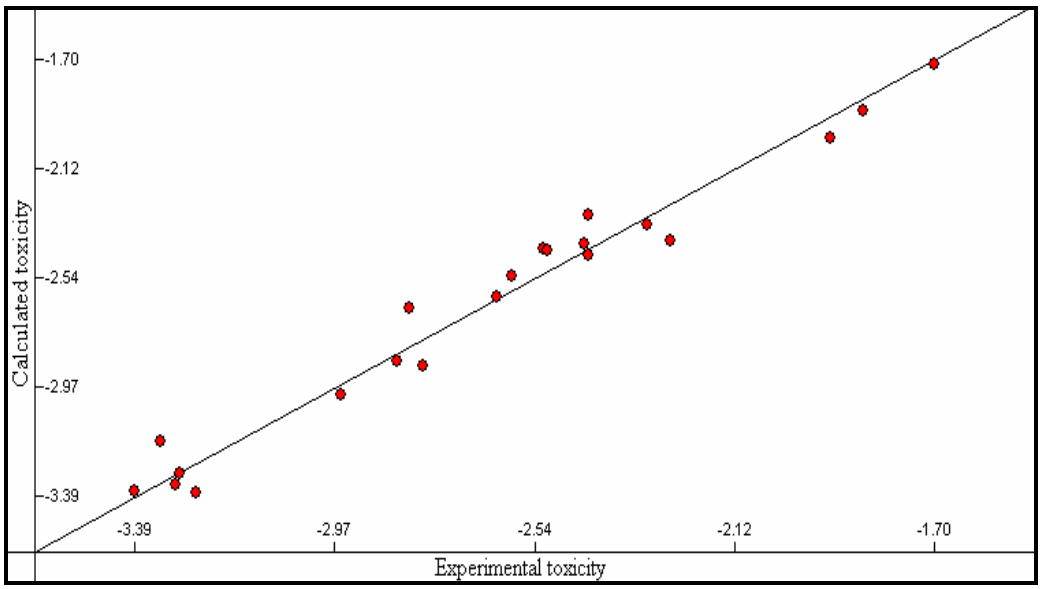


Fig. 2: Comparison of experimental and calculated toxicity from the regressional analysis in eq. (6). R^2=0.962; R^2_{CV}=0.950; F=99.75; s^2=0.096 for 21 nitrobenzenes.

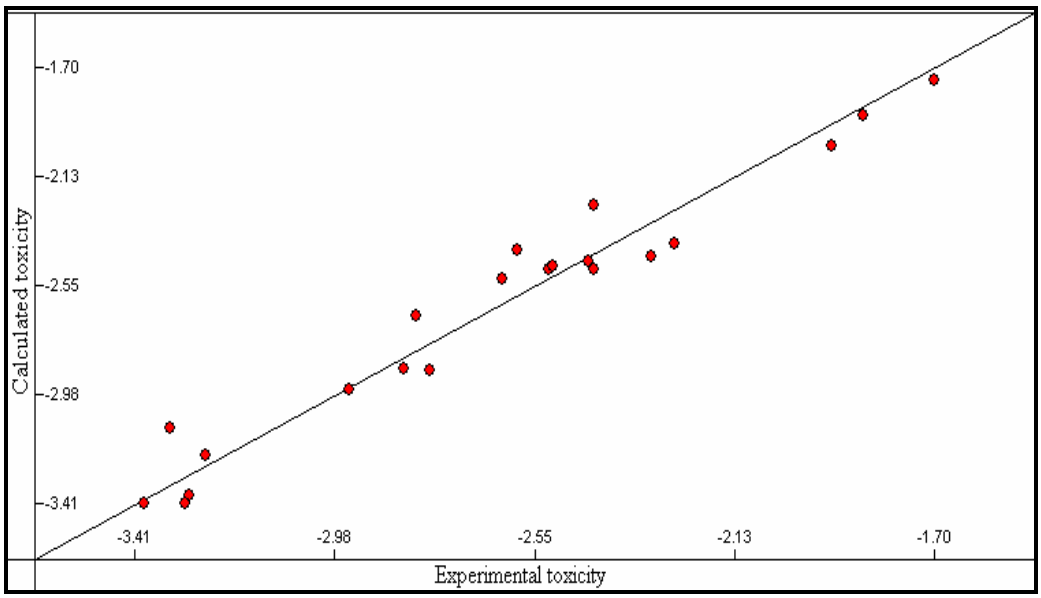


Fig. 3: Comparison of experimental and calculated toxicity from the regressional analysis in eq. (5). R^2 = 0.935; R^2_{cv} = 0.918; F=81.017; s^2 = 0.0144 for 21 nitrobenzenes.