Quantitative Structure–Property Relationship Prediction of the Half-Wave Potential for Substituted Nitrobenzenes in Five Nonaqueous Solvents

Mohammad Hossein Fatemi,* Mohammad Reza Hadjmohammadi,
Kamyar Kamel, and Poriya Biparva

Department of Chemistry, University of Mazandaran, Babolsar, Iran

Received June 20, 2006; E-mail: mhfatemi@umz.ac.ir

Half-wave potentials ($E_{1/2}$) are an important electrochemical property of organic compounds, which depends on both solute and solvent. In this work, a quantitative structure–property relationship (QSPR) study was carried out to predict the reduction $E_{1/2}$ behavior of 15 substituted nitrobenzenes in five solvents, including acetonitrile, *N*,*N*-dimethylformamide, dimethyl sulfoxide, methanol, and tetrahydrofuran. For selection of the most relevant descriptors, a stepwise multiple linear regression (MLR) technique was used. Five descriptors that appeared in the MLR model were: minimum charges of the oxygen atom in the nitro group, Guttmann acceptor number of the solvent, dipole moment of solute, dipole moment of the solvent, and Hammet substitution constant (σ). The leave-one out cross validation method was used to evaluate the credibility of the model. Statistical parameters derived from this test are: $R^2 = 0.960$, $Q^2 = 0.959$, and *SPREES* = 0.046. These indicate the suitability of the constructed model.

Nitroaromatic hydrocarbons are widely used as pesticides, explosives, solvents, and pharmaceuticals and comprise an important group of environmental pollutants.[1] Therefore, the investigation of their activities and properties are necessary and interesting. Half-wave potentials ($E_{1/2}$), which are an important electrochemical constant for a reversible oxidation–reduction system, can be useful for predicting other electrochemical properties and activities of organic compounds.[2]

First step in the reduction of many organic compounds is a one-electron process that produces a radical anion, $Q^{\bullet -}$, according to the following equation:

$$Q + e^- \rightarrow Q^{\bullet -} \tag{1a}$$

However, because $Q^{\bullet -}$ is generally reactive, its lifetime varies over a wide range, depending on its intrinsic reactivity and experimental conditions, such as the type of solvent, complex formation, and other chemical reactions.[3]

In recent years, numerous quantitative structure–activity/property relationship (QSAR/QSPR) models have been introduced for calculating the physicochemical properties of molecules from their chemical structures.[4–8] These methods can be used to establish correlations between the similarities among individual compounds and their biological activities/chemical properties.[9–11] There are some reports about applications of QSAR\QSPR in electrochemistry.[12–15] Wei et al. studied the relationship between the reduction characteristics and molecular structure of 87 chlorinated aromatics, such as naphthalenes, biphenyls, benzenes, and phenols.[12] Nesmerak and co-workers employed QSPR approach to investigate the electrooxidation of new benzoxazines as a model of metabolic degradation.[13] Tompe et al. have reported a quantitative structure–electrochemistry relationship study on the half-wave potential of α,β-unsaturated ketones in acetonitrile solution.[14] Also, Shamsipur and Hemmateenejad employed principle component regression (PCR) and principle component–artificial neural network (PC-ANN) models in a QSPR study of some organic compounds.[15] The main aim of the present work was to develop a QSPR model based on multiple linear regression (MLR) to predict $E_{1/2}$ values. To the best of our knowledge, this is the first report published involving prediction of $E_{1/2}$ for nitrobenzene derivatives in five solvents using a unique MLR equation.

Method

Data Set. The data set consists of the 75 experimental values of $E_{1/2}$ for 15 substituted nitrobenzenes in five solvents, which were taken from the work published by Shalev and Evans.[3] These values are measured against a cobaltocenium/cobaltocene reference electrode. The solvents were: acetonitrile (AN), *N*,*N*-dimethylformamide (DMF), dimethyl sulfoxide (DMSO), methanol (MeOH), and tetrahydrofuran (THF). The compounds in the data set with their corresponding $E_{1/2}$ in different solvents are listed in Table 1.

Calculation of Electronic Descriptors. One of the simplest and fundamentally important reactions that a neutral organic molecule can undergo is reduction by one electron to form an anion radical. In the gas phase, the electron affinity (EA) is a measure of the ability of a molecule to accept an electron to form an anion radical. In solution, the tendency of solute to accept an electron is governed in part by gas phase EA, but it is also greatly affected by stabilization of anion through interaction with solvent molecules. Therefore, some electronic parameters, which can quantitatively describe the above interactions, were calculated for molecules and solvents of interest. The quantum-mechanical descriptors for molecules were calculated by using the MOPAC program (AM_1 semi empirical method),[16] based on the minimum energy molecular geometries optimized with the Hyperchem package.[17] The

Published on the web February 13, 2007; doi:10.1246/bcsj.80.303

Table 1. The Experimental Half-Wave Potentials in V in Reference to Cobaltocenium/Cobaltocene Electrode and the Cross-Validation Predicted Potentials for 15 Substituted Nitrobenzenes in Five Solvents Using a Combined Model

		Solvents									
		AN		DMF		DMSO		MeOH		THF	
No	Compounds	Exp.	Pred.	Exp.	Pred.	Exp.	Pred.	Exp.	Pred.	Exp.	Pred.
1	4-Methoxynitrobenzene	−0.322	−0.318	−0.367	−0.358	−0.276	−0.317	0.041	−0.072	−0.494	−0.496
2	4-Methylnitrobenzene	−0.270	−0.254	−0.304	−0.293	−0.216	−0.251	0.074	−0.010	−0.434	−0.441
3	3-Methylnitrobenzene	−0.232	−0.222	−0.271	−0.250	−0.183	−0.213	0.097	0.030	−0.401	−0.400
4	Nitrobenzene	−0.211	−0.205	−0.250	−0.227	−0.163	−0.199	0.099	0.053	−0.380	−0.369
5	3-Methoxynitrobenzene	−0.196	−0.168	−0.235	−0.199	−0.146	−0.158	0.116	0.075	−0.370	−0.337
6	4-Fluoronitrobenzene	−0.202	−0.125	−0.233	−0.140	−0.152	−0.111	0.097	0.153	−0.362	−0.291
7	3-Fluoronitrobenzene	−0.100	−0.060	−0.125	−0.090	−0.046	−0.035	0.185	0.187	−0.256	−0.240
8	4-Chloronitrobenzene	−0.128	−0.097	−0.157	−0.130	−0.081	−0.088	0.151	0.155	−0.285	−0.268
9	3-Chloronitrobenzene	−0.080	−0.081	−0.110	−0.131	−0.036	−0.064	0.205	0.174	−0.238	−0.252
10	3-(Trifluoromethyl)nitrobenzene	−0.058	−0.009	−0.068	−0.049	0.007	−0.021	0.212	0.231	−0.197	−0.192
11	3-Cyanonitrobenzene	−0.017	−0.025	−0.035	−0.049	0.034	−0.004	0.248	0.229	−0.155	−0.188
12	1,3-Dinitrobenzene	0.039	0.031	0.022	−0.011	0.088	0.043	0.269	0.283	−0.101	−0.150
13	4-Cyanonitrobenzene	0.063	0.083	0.066	0.050	0.131	0.097	0.288	0.339	−0.061	−0.092
14	1,4-Dinitrobenzene	0.257	0.238	0.265	0.196	0.321	0.223	0.480	0.483	0.124	0.047
15	3,5-Dinitrobenzonitrile	0.273	0.286	0.272	0.257	0.331	0.297	0.394	0.584	0.147	0.116

Hammett substituent constants for 15 substituted nitrobenzenes were taken from the published work of Hansch et al.[18]

Model Development. Stepwise MLR was used to select the most important descriptors and calculate the coefficients which relate them to the $E_{1/2}$ in each solvent, separately. The descriptors that were derived from the best MLR equations for each solvent were the same and are summarized in Table 2. These descriptors are: minimum charges on the oxygen atoms in the nitro group (*MCOA*), electron density on the nitro group (*EDN*), dipole moment of the solute (*DP*) (Debye), and Hammet substitution constant (σ). The leave-one out cross validation method was used to evaluate the validity of the model. Using this method, one molecule was removed from the data set, and the model was generated with remaining molecules. Then, the value of $E_{1/2}$ for the removed molecule was calculated using the generated model. This procedure was continued until each molecule was predicted once. Q^2 and the square of predicted residual error sum of square (*SPRESS*) were calculated as criterias for evaluating the credibility of model using following equations:[19]

$$Q^2 = 1 - \frac{\sum(Y_\mathrm{p} - Y_\mathrm{exp})^2}{\sum(Y_\mathrm{exp} - Y_\mathrm{mean})^2}. \tag{1b}$$

$$SPRESS = \sqrt{\frac{\sum(Y_\mathrm{p} - Y_\mathrm{exp})^2}{n-k-1}}. \tag{1c}$$

The statistical parameters of these models and also the mean effects of the descriptors are listed in Table 3. The value of mean effect for a given descriptors indicates how important of a descriptor is, and its sign (positive or negative) represents the type of correlation which exist between the selected descriptor and investigated property. From the mean effect of descriptors, *MCOA* and σ have a positive and *EDN* and *DP* have a negative effect on $E_{1/2}$. To develop a combined model for consideration of solvent affects, a few solvent characteristics, such as Guttmann acceptor number (*GAN*), dipole moment (μ) (Debye), and dielectric constant (ε), were added to calculated molecular descriptors. Then, stepwise MLR was applied for selecting the variables and constructing the combined model. Equation 1d shows the obtained MLR equation, which relates the $E_{1/2}$ of nitrobenzene derivatives to their structural descriptors.

$$\begin{aligned} E_{1/2}/\mathrm{V} = 3.223 + 0.160 * \sigma + 0.022 * \mu/\mathrm{D} - 0.027 * DP \\ + 0.012 * GAN + 10.009 * MCOA \end{aligned} \tag{1d}$$

In this equation, *MCOA* is minimum charges on the oxygen atom of the nitro group, *GAN* is the Guttmann acceptor number, *DP* is the dipole moment of the solute, μ is the dipole moment of the solvent, and σ is the Hammet substitution constant. The statistical parameters of this equation are given in Table 4. Finally, the leave-one out cross validation test was employed for the evaluation of the prediction power of obtained MLR model.

Results and Discussion

The main goal of this work was the development of a combined MLR model for predicting the $E_{1/2}$ values of nitrobenzene derivatives in five selected solvents. Table 1 lists experimental and calculated values of $E_{1/2}$ for 15 nitrobenzene derivatives in acetonitrile, *N*,*N*-dimethylformamide, dimethyl sulfoxide, methanol, and tetrahydrofuran using the combined MLR model that we developed. Table 4 also summarizes the specification and statistical parameters of this model. From Table 4, the five descriptors in this model are *MCOA*, *GAN*, *DP*, μ, and σ.

As shown in Table 4, *GAN* has the highest mean effect. *GAN* of a solvent is a measure a solvent's acidity to accept an electron from a solute molecule or anion radical, such as a Lewis acid, and the scale of *GAN* is based on ^{31}P NMR chemical shifts of the solute triethylphosphine oxide in the various solvents.[20] The positive sign of mean effect for *GAN* means that the value of $E_{1/2}$ increased as the value of this descriptor increased. From Table 4, μ, which is the solvent's ability to form dipole-interaction, has a positive effect on $E_{1/2}$. A solvent with a higher μ has a stronger interaction with an anion radical and, thus, stabilizes it, and therefore its mean

Table 2. Coefficients and Their Standard Errors of MLR Models for Different Solvents

Descriptor	Notation	Coefficients				
		AN	DMF	DMSO	MeOH	THF
Constant	—	5.169 ± (1.496)	5.683 ± (1.628)	5.433 ± (1.467)	5.246 ± (2.319)	5.308 ± (1.599)
Minimum charges on oxygen atom in nitro group	MCOA	13.751 ± (4.101)	15.207 ± (4.462)	14.304 ± (4.022)	13.176 ± (6.355)	14.563 ± (4.382)
Electron density on nitro group	EDN	−0.052 ± (0.017)	−0.053 ± (0.066)	−0.052 ± (0.017)	−0.047 ± (0.027)	−0.053 ± (0.018)
Dipole moment of solute	DP	−0.022 ± (0.006)	−0.026 ± (0.007)	−0.025 ± (0.006)	−0.020 ± (0.010)	−0.023 ± (0.007)
Hammet substitution constant	σ	0.122 ± (0.060)	0.123 ± (0.019)	0.116 ± (0.059)	0.094 ± (−0.003)	0.143 ± (0.065)

Table 3. Statistical Parameters for MLR Models Derived for Different Solvents

Descriptors	Mean effect				
	AN	DMF	DMSO	MeOH	THF
MCOA	3.353	3.408	3.557	2.073	3.323
EDN	−2.996	−2.818	−3.084	−1.758	−2.869
DP	−3.548	−3.875	−4.010	−2.092	−3.436
σ	2.020	1.872	1.952	−0.032	2.222
R^2	0.990	0.990	0.991	0.950	0.990
Q^2	0.990	0.990	0.990	0.950	0.990
SPRESS	0.0004	0.0005	0.0004	0.001	0.0005

Table 4. Specification of MLR Models for All Molecules in All Solvents

Descriptor	Coefficient	Mean effect
MCOA	10.009 ± (3.405)	2.940
GAN	0.012 ± (0.000)	27.886
DP	−0.027 ± (0.005)	−5.125
μ	0.022 ± (0.006)	3.546
σ	0.160 ± (0.050)	3.163
Constant	3.223 ± (1.208)	—
R^2	0.960	
SE	0.045	
N	75	

effect has a positive sign. Another descriptor, which appears in combined MLR model, is *MCOA*. A decrease in the negative charge density increases the ability of a molecule to accept an electron and to form an anion radical; therefore, $E_{1/2}$ become more positive. The forth descriptor in the model is *DP*, of which the mean effect has a negative sign. Therefore, an increase in *DP* causes a decrease in $E_{1/2}$. The final descriptor indicates the effect of the substituent on the change in free energy for the formation of anion radicals. Changing the substituent can cause a change in polarization and electron density around the ring in both reactant and product, and it is apparent from experimental data that the $E_{1/2}$ value of each nitrobenzene derivative is affected by the type of substituent on the benzene ring. In all solvents, $E_{1/2}$ increases with an increase in σ because substituent with a higher σ value can more easily accept an electron, i.e., an electron-withdrawing group stabilizes the anion radical and causes the potentials increases.

In order to evaluate the credibility of combined MLR model, the leave-one-out cross validation method was employed. The values of Q^2 and SPRESS are 0.959 and 0.046, respectively, and support the validity of our model.

The $E_{1/2}$ values calculated by this model are reported in Table 1. Figure 1 shows a plot of combined MLR predicted values versus experimental values of $E_{1/2}$. The residual of MLR model calculated values are plotted against the experimental values of $E_{1/2}$ in Fig. 2. The propagation of residuals on both sides of the zero line indicates that no systematic error exist in the MLR model.

Conclusion

In QSPR studies for simplicity, a linear model use to correlate a specific property to structural parameters (descriptors) of the molecule. If the obtained model does not have suitable stat-

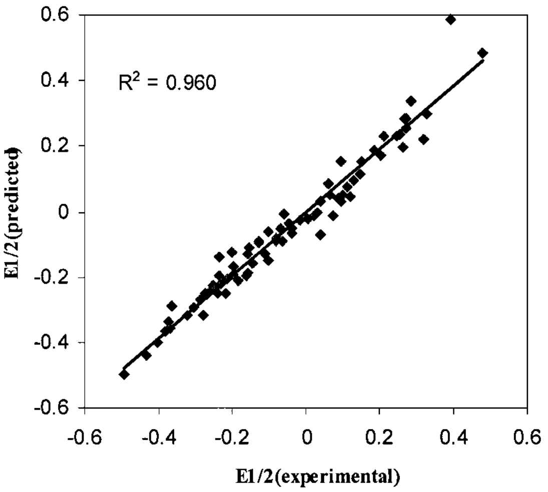

Fig. 1. Plot of predicted $E_{1/2}$ from MLR model against experimental $E_{1/2}$ values.

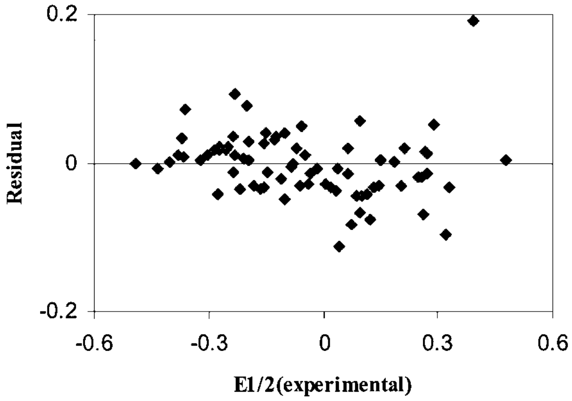

Fig. 2. Plot of the residuals against experimental value of $E_{1/2}$.

istical parameters, then non-linear terms or non-linear modeling techniques, such as artificial neural networks are used. In the present study, MLR was used to develop a QSPR model to predict the reduction $E_{1/2}$ values for nitrobenzene derivatives in five different solvents. Results showed that the QSPR method using MLR can generate suitable models for the prediction of reduction $E_{1/2}$ values for these compounds using electronic descriptors for solutes and solvents, and therefore non-linear terms or non-linear techniques were not needed in our QSPR studies.

References

1 S. A. Boyd, G. Sheng, B. J. Teppen, C. T. Johnston, *Environ. Sci. Technol.* **2001**, *35*, 4227.

2 R. L. McNaughton, A. A. Tipton, N. D. Rubie, R. R. Conry, M. L. Kirk, *Inorg. Chem.* **2000**, *39*, 5697.

3 H. Shalev, D. H. Evans, *J. Am. Chem. Soc.* **1989**, *111*, 2667.

4 M. H. Fatemi, *J. Chromatogr., A* **2003**, *1002*, 221.

5 M. H. Fatemi, *Anal. Chim. Acta* **2006**, *556*, 355.

6 M. Karelson, A. Perkson, *Comput. Chem.* **1999**, *23*, 49.

7 M. H. Fatemi, N. Goudarzi, *Electrophoresis* **2005**, *26*, 2968.

8 T. F. Woloszyn, P. C. Jurs, *Anal. Chem.* **1992**, *64*, 3059.

9 R. Kaliszan, A. Kaliszan, T. A. G. Noctor, *J. Chromatogr., A* **1992**, *609*, 69.

10 P. Gramatica, E. Papa, *QSAR Comb. Sci.* **2003**, *22*, 374.

11 C. Hansch, A. Kurup, R. Garg, H. Gao, *Chem. Rev.* **2001**, *101*, 619.

12 D. Wei, H. Hu, K. Hiramoto, K. Fujie, *Huanjing Kexue* **2003**, *24*, 19.

13 K. Nesmerak, I. Nemece, M. Sticha, K. waisser, K. palaf, *Electrochim. Acta* **2005**, *50*, 1431.

14 P. Tompe, G. Clementis, I. Petnehazy, L. Toke, *Anal. Chim. Acta* **1995**, *305*, 295.

15 B. Hemmateenejad, M. Shamsipur, *Internet Electron. J. Mol. Des.* **2003**, *1*, 316.

16 J. J. P. Stewart, *MOPAC, Semi Empirical Molecular Orbital Program*, QCPE, **1983**, Vol. 455; *MOPAC, Semi Empirical Molecular Orbital Program*, Research Laboratory, United States Air Force Academy, **1990**, version 6.

17 *Hyperchem 5.02*, Hypercube, Inc., Gainesville, U.S.A., **1994**.

18 C. Hansch, A. Leo, R. W. Taft, *Chem. Rev.* **1991**, *91*, 165.

19 Y. S. Prabhakar, *QSAR Comb. Sci.* **2003**, *22*, 583.

20 V. Gutmann, *Electrochim. Acta* **1976**, *21*, 661.