DE GRUYTER
OPEN

Theoretical study of a series of phenol derivatives: molecular properties *vs.* cytotoxicity

Martin Michalík, Lívia Sádecká, Vladimír Lukeš

Institute of Physical Chemistry and Chemical Physics, Slovak University of Technology in Bratislava, Radlinského 9, SK-812 37 Bratislava, Slovakia
martin.michalik@stuba.sk

Abstract: The quantum chemical calculations using DFT were performed for 2-alkyl-4X and 2,6-dialkyl-4-X substituted phenols. Based on the optimal geometries the bond dissociation enthalpies (BDEs), proton enthalpies (PAs) and the lipophilicities were computed. Additionally, simple geometry parameter was found correlating well with experimental leukemia cell toxicity of substituted phenols. Next, we have found no linear dependence between PA or BDE values and log1/*C* values in gas phase or in water despite the radical toxicity mechanism proposed in the literature.

Keywords: model compounds, drug development, QSAR, substituent effect

Introduction

Phenol derivatives show various biological activities including many of them that are still not completely understood. They may act like radical scavengers as well as cytotoxic compounds. Some evidences even suggest catechol and hydroquinone, the components in cigarette tar, as the cause of lung damage in smokers (Ong et al., 1994). Recent biological studies examined the cytotoxicity, i.e. the molar concentration of substituted phenols that induces 50 % inhibition of fast-growing murine leukemia cell line (L1210). Important structural and electronic features that might be responsible for the biological activity were pointed out for twenty one of 2-alkyl-4-X-phenols and 2,6-dialkyl-4-X-phenols (Selassie et al., 2002). Authors supposed a radical mechanism of the observed toxicity. For 14 derivatives, the linear dependence of toxicity on relative homolytic O—H bond dissociation energy (BDE) was found with the coefficient of determination $R^2 = 0.936$. The BDEs were obtained by quantum chemical calculations based on the gas-phase density functional calculations in 6-31G** basis set using AM1 optimal geometries. On the other hand, the correlations of partition coefficients (log *P*) and/or Taft's steric parameters with the cytotoxicity concentrations for more than 10 compounds lead to worse dependencies. The log *P* values were obtained from the Quantitative Structure-Activity Relationship (QSAR) model. Authors concluded that overall toxicity of the phenol derivatives involves a balance between electron density on the aromatic ring and lipophilicity. Also, electron withdrawing substituents induced toxicity, but much less than their electron donating counterparts (Selassie et al., 1998; Ghamali et al., 2015).

From the theoretical point of view, the quantum chemical approaches are commonly used in drug design. Very often, the molecular properties are calculated for a set of model compounds or real drugs and they are correlated with macroscopic biological activities. In this context we can mention the density functional theory. In comparison with QSAR models the quantum chemical calculations are able to account for the conformation effects and specific geometry effects which are important in drug docking.

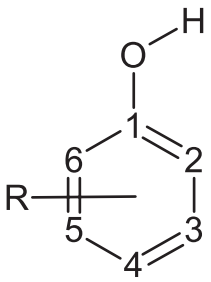
With respect to the above mentioned we will perform the DFT calculations for 21 phenolic compounds. The schematic structure of studied derivatives is presented in Scheme 1 and the substituents are collected in Tab. 1. The main goals of this work can be defined as follows:

i) to obtain optimal geometries in the gas phase and solvents (water, n-octanol);
ii) to calculate the BDEs and proton affinities (PAs) for phenolic O—H bond.

Next, we employ the statistical tools and explore various correlations of experimental phenol derivatives toxicity and obtained parameters.

Computational details

DFT calculations were accomplished using Gaussian 09 program package (Frisch et al., 2009). For our calculations, the 6-311+G(d) basis sets (Krishnan et al., 1980; McLean et al., 1980; Clark et al., 1983) and B3LYP hybrid functional (Lee et al., 1988; Becke, 1988) were employed. At first the optimal geometries of the studied species (in neutral, ionic and radical form) were found in gas phase (with *opt = tight* keyword). Vibrational analysis showed no imaginary frequencies confirming the real geometry of the energy minima. BDE



Scheme 1. Schematic structure of phenol derivatives with atom numbering.

values of hydroxyl group cleavage

$$BDE = H(\mathrm{ArO^{\bullet}}) + H(\mathrm{H^{\bullet}}) - H(\mathrm{ArO{-}H}) \quad (1)$$

were based on the enthalpies of the phenol derivative H(ArO—H) in the neutral form, hydrogen radical enthalpy $H(\mathrm{H^{\bullet}})$ and the enthalpy of the phenol derivative radical $H(\mathrm{ArO^{\bullet}})$ that is created after homolytic O—H bond breaking. Proton affinities were calculated in similar fashion

$$PA = H(\mathrm{ArO^{-}}) + H(\mathrm{H^{+}}) - H(\mathrm{ArO{-}H}) \quad (2)$$

where $H(\mathrm{ArO^{-}})$ is the enthalpy of phenol derivative ion formed after the proton abstraction from O—H bond. The symbol $H(\mathrm{H^{+}})$ stands for the proton enthalpy. All the enthalpies were estimated for T = 298.15 K. Solvent effects for theoretical hydrophobicity estimation were accounted by the continuum Solvation Model based on the quantum mechanical charge Density (SMD) (Marenich et al., 2009) of a solute molecule interacting with a continuum implemented in Gaussian 09. Lipophilicity calculations involve the abstraction of the Gibb's energy of the derivative in water G_{W}(ArO—H) and in *n*-octanol G_{O}(ArO—H)

$$\text{DFT-Log}P = \frac{G_{\mathrm{w}}(\mathrm{ArO{-}H}) - G_{\mathrm{o}}(\mathrm{ArO{-}H})}{2.303\,\mathrm{R}\,T} \quad (3)$$

Finally, the BDE and PA values in solvent approximated by SMD model were calculated using Eqs. (1) and (2). The hydrogen atom ($\mathrm{H^{\bullet}}$) and proton ($\mathrm{H^{+}}$) hydration enthalpies were taken from the literature: $\Delta_{\mathrm{hydr}}H(\mathrm{H^{\bullet}})$ = –4 kJ mol^{-1} (Parker, 1992; Bizarro, 1999), $\Delta_{\mathrm{hydr}}H(\mathrm{H^{+}})$ = –1090 kJ mol^{-1} (Atkins, 1998).

Tab. 1. The experimental cytotoxicity of the studied substituted phenols log 1/C (Selassie et al., 2002), B3LYP/6-311+G* gas phase and SMD(Water) BDE and PA values, predicted lipophilicity DFT-Log*P*, theoretical lipophilicity model MiLog*P* (Molinspiration, 2016) and the C1—C2 distance $R_{C1—C2}$.

#	Substituent	log 1/*C*	BDE/kJ mol^{-1}		PA/kJ mol^{-1}		DFT-Log*P*	MiLog*P*	$R_{C1—C2}$/Å
			gas	water	gas	water			
1	2,6-di-Me	3.02	311	304	1427	111	1.97	2.52	1.40220
2	2,6-di-OMe	3.86	311	284	1458	107	0.40	1.34	1.40200
3	2,4,6-tri-Me	3.20	304	293	1432	113	2.36	2.92	1.40218
4	2,6-di-CMe_3	3.85	296	292	1393	114	4.03	5.04	1.41459
5	2,6-di-CMe_3-4-Me	4.04	289	281	1397	119	4.48	5.43	1.41456
6	2,6-di-C_2H_5	3.26	309	304	1418	113	2.55	3.45	1.40382
7	2,6-di-$CHMe_2$	3.25	306	301	1412	113	3.49	3.70	1.40695
8	2,4,6-tri-CMe_3	3.90	290	283	1393	119	5.99	6.69	1.41365
9	2-CMe_3-6-Me	3.73	308	298	1414	115	3.49	3.78	1.41089
10	2,6-di-CMe_3-4-NO_2	4.90	313	318	1306	59	3.53	4.95	1.41883
11	2,6-di-CMe_3-4-C_2H_5	3.91	289	283	1394	120	4.55	5.90	1.41328
12	2,6-di-CMe_3-4-Br	4.11	294	294	1366	105	3.11	5.80	1.41430
13	2,4-di-CMe_3	4.24	312	301	1418	116	4.36	5.27	1.40645
14	2-CMe_3-4-Me	3.80	308	300	1420	119	3.42	4.01	1.41182
15	2,4-di-Me	3.04	310	299	1433	112	1.89	2.75	1.40089
16	2-Me-4-F	3.09	312	308	1415	106	1.33	2.47	1.40401
17	2-Me-4-NO_2	3.49	336	341	1327	60	–0.13	2.26	1.40870
18	2-Me-4-Br	3.46	316	310	1397	102	1.21	3.11	1.40383
19	2-Me-4-OMe	3.39	297	286	1437	114	0.99	2.36	1.40578
20	2-Me-4-COMe	3.14	325	323	1366	85	0.97	2.20	1.40512
21	2-CMe_3-4-C_2H_5	3.80	311	300	1419	118	3.75	4.47	1.41048

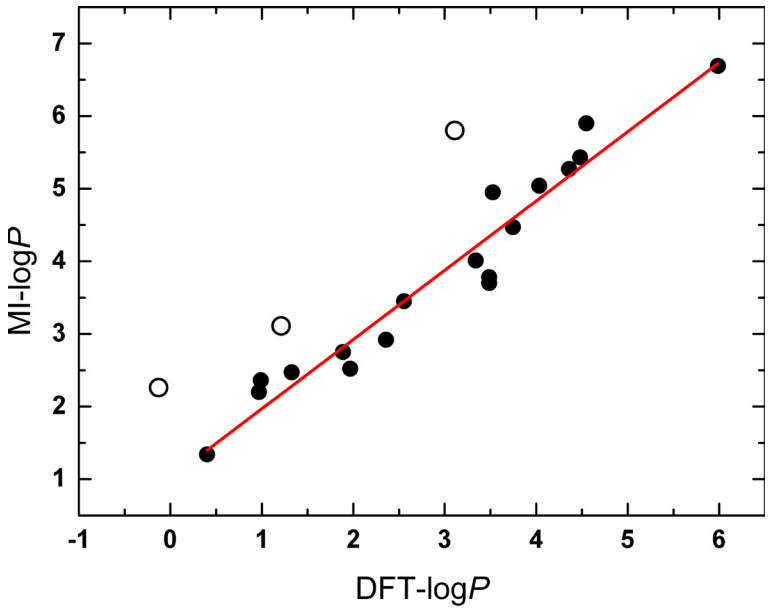


Fig. 1. Correlation of theoretical lipophilicity values — calculated by Molinspiration online tool (MI-log$P$) and on B3LYP/6-311+G* level of DFT theory (DFT-log$P$). The omitted substituents are denoted as empty circles.

Results and Discussion

The selected compounds examined here include the commercial antioxidant BHT (#05) and its analogs possessing similar structure of a sterically hindered phenolic group. These molecules are able to scavenge free radicals and to form relatively stable radicals after hydrogen abstraction from the O—H bond. Therefore, radical mechanism of the toxicity has been anticipated by several works (Selassie et al., 1998, 2002).

The calculated BDE values are collected in Tab. 1. The solvent has the minimal effect on these values (from 0 to 11 kJ mol^{-1} except for the derivative No. 2 containing methoxy groups in *ortho* positions). Despite above mentioned literature findings, an attempt to obtain a linear relationship between experimental toxicity concentration (log 1/C) values and DFT BDEs resulted in a poor correlation. This discrepancy can indicate the preference of another mechanism. While in gas phase the hydrogen atom cleavage may be thermodynamically preferred, in water solvent the proton abstraction is by ca. 200 kJ mol^{-1} less energy demanding (Tab. 1). Nevertheless, PAs in gas phase as well as in water show no linear correlation with toxicity. The lipophilicity is the next calculated thermodynamical property connected with the solubility. Since experimental log P values to our best knowledge are not available the DFT calculations in water and octanol solvents allowed the theoretical estimation of a DFT-log$P$ dependence. The check for coarse errors can be done by correlating the obtained values with the Molinspiration data (MI-log$P$), which is a simple online tool for hydrophobicity prediction (Cheminformatics, 2016).

$$\text{MI-log}P = 1.0(2) + 0.95(6) \cdot \text{DFT-log}P \qquad (4)$$

We found correlation coefficient R = 0.973 after excluding slightly off linear points for 4-bromo-2,6-di-*tert*-butylphenol, 2-methyl-4-nitrophenol and 2-methyl-4-bromophenol. Cytotoxicity uses to be directly proportional to the partition coefficient, because the site of action is usually a lipid membrane. Indeed, the linear trend of the experimental toxicity expressed as log1/C with DFT calculated values (DFT-log$P$) according to the Eq. (5)

$$\log 1/C = 2.4(2) + 0.37(4) \cdot \text{DFT-log}P \qquad (5)$$

can be found with the correlation coefficient R = 0.955. This linear dependence covers 57 % of the derivatives. It seems that the specific solvent effects neglected in the implicit models and steric hindrance may be important here and are difficult to be described theoretically.

Besides the energetical quantities the molecular structure should have the direct influence on the cytotoxicity. In our model we have investigated the role of the bonds of the aromatic ring. For the sake of simplicity, we have defined the gas phase geometry parameter R_d

$$R_d = (R_i - R_{\text{PhOH},\,i})^2 \qquad (6)$$

where R_i is carbon-carbon distance of i-th derivative in neutral form and $R_{\text{PhOH},\,i}$ represents the corresponding bond in the parent phenol molecule. The

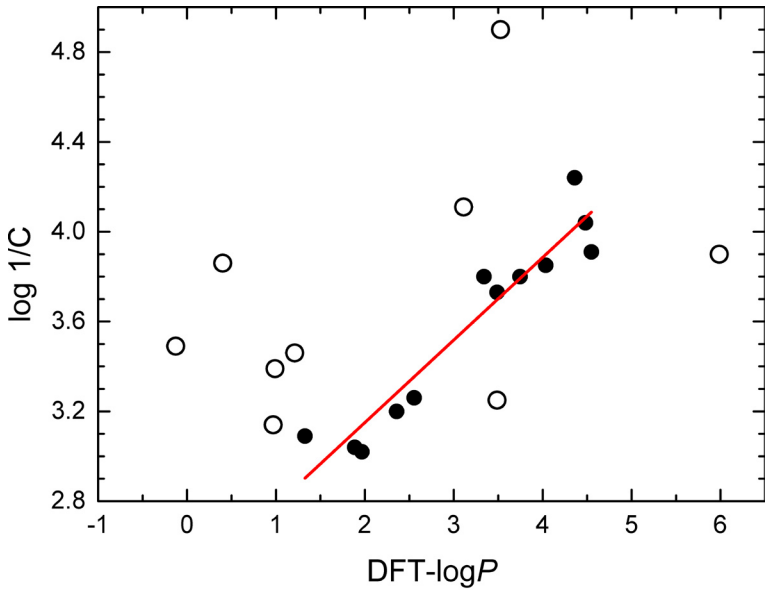


Fig. 2. Dependence of experimental phenol toxicity (log 1/C) on theoretical B3LYP/6-311+G* lipophilicity values (DFT-log*P*). The omitted substituents are denoted as empty circles.

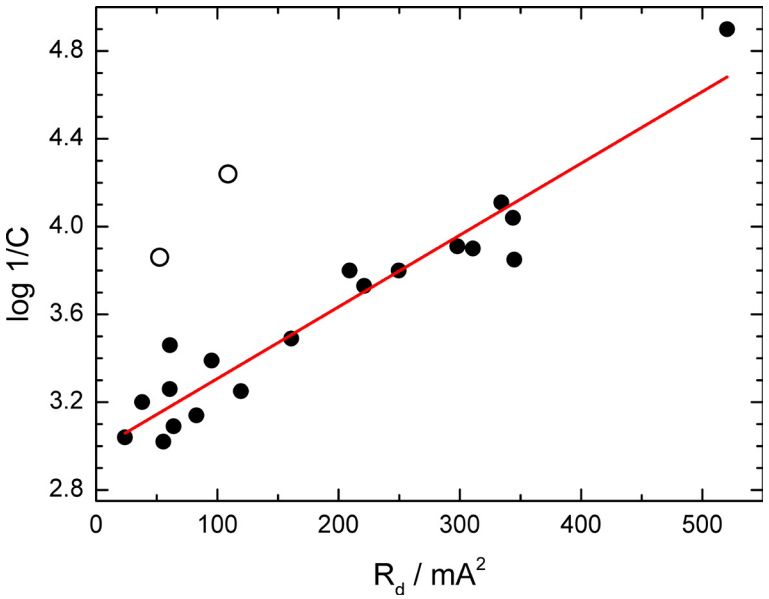


Fig. 3. Relationship between experimental phenol toxicity (log 1/C) and theoretical B3LYP/6-311+G* relative geometrical parameter R_d (see eq. 6) for C1—C2 distantance. The omitted substituents are denoted as empty circles.

performed statistical analysis indicated that only the C1—C2 bond (see Scheme 1 for atom numbering) is relevant as good linear relationship (R = 0.961) of its R_d parameter with cytotoxicity was obtained. Regression analysis yielded the following equation for this bond (standard deviations in parentheses)

$$\log 1/C = 2.98(6) + 3.3(3)10^3 \cdot R_d \qquad (7)$$

Two compounds (2,6-dimethoxyphenol, and 2,4-di-*tert*-butylphenol) were omitted (see Fig. 3). The latter one seems to be generally difficult to characterize (Borges dos Santos, 1998; Selassie et al., 2002). The reference C1—C2 distance in phenol is 1.39476 Å and Tab. 1 contains $R_{C1—C2}$ values for the studied derivatives. It seems that in order to make phenol more cytotoxic, we have to pick the substituents which can stretch the C1—C2 bond and *vice-versa*.

Conclusions

In this paper, the set of 21 di- and trisubstituted phenols with most substituents in the *ortho* position to the OH group were examined, and the corresponding experimental toxicity was correlated with

theoretical parameters calculated at the DFT level of theory. Our results render toxicity, represented by log $1/C$ values, closely related to lipophilicity as the DFT-log$P$ data explained the cytotoxicity of most studied derivatives. It should be mentioned that the accuracy of the theoretical values is highly dependent on the choice of the solvation model. Adding the solvent molecules explicitly may improve the description of the solvent effects, but at the cost of the significantly higher CPU time requirements. Despite various insights regarding radical mechanism of the toxicity that can be found in the literature, we found no linear dependence of BDE values and log $1/C$ in gas phase or in water. The same conclusion holds for proton affinities as well. It seems that the mechanism of the toxicity cannot be easily described by thermodynamical quantities and various subtle steric and substituent effects may be either in synergy or in antagony. Additionally, we have shown that the simple gas phase geometrical parameter based on the C1—C2 bond provides better correlation with experimental toxicity values. The most cytotoxic derivative had also the longest C1—C2 bond. Found linear dependences can be used to predict biological activities of other phenolic compounds.

Acknowledgment

The work has been supported by Slovak Grant Agency (1/0594/16) and Slovak Research and Development Agency (APVV-15-0053). We are grateful to the HPC center at the Slovak University of Technology in Bratislava, which is a part of the Slovak Infrastructure of High Performance Computing (SIVVP project, ITMS code 26230120002, funded by the European region development funds, ERDF) for the computational time and resources made available.

References

Atkins PW (1998) Physical Chemistry, 6th ed. Oxford University Press, Oxford.

Becke AD (1988) Phys Rev A 38: 3098—3100.

Bizarro MM, Cabral BJC, dos Santos RMB, Simões JAM (1999) Pure Appl Chem 71: 1249—1256.

Borges dos Santos RM, Martinho Simões JA (1998) J Phys Chem Ref Data 27: 707—739.

Cheminformatics, Molinspiration. Bratislava, Slovak Republic. 2016.

Clark T, Chandrasekhar J, Spitznagel GW, Schleyer PVR (1983) J Comp Chem 4: 294—301.

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr., Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Keith T, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J and Fox DJ (2009) Gaussian 09, Revision D.01, Gaussian, Inc. Wallingford CT.

Ghamali M, Chtita S, Adad A, Hmamouchi R, Bouachrine M, Lakhlifi T (2015) J Mater Environ Sci 6: 280—288.

Krishnan RBJS, Binkley JS, Seeger R, Pople JA (1980) J Chem Phys 72: 650—654.

Lee C, Yang W, Parr RG (1988) Phys Rev B 37: 785—789.

Marenich AV, Cramer CJ, Truhlar DG (2009) J Phys Chem B 113: 6378.

McLean AD, Chandler GS (1980) J Chem Phys 72: 5639—5648.

Ong CN, Lee BL, Shi CY, Ong HY, Lee HP (1994) Int J Cancer 59: 177—180.

Parker VD (1992) J Am Chem Soc 114: 7458—7462.

Selassie CD, De Soyza TV, Rosario M, Gao H, Hansch C (1998) Chem-bio Interact 113: 175—190.

Selassie C, Shusterman A, Verma R (1999) J Chem Soc Perk Trans 2: 2729—2733.

Selassie CD, Verma RP, Kapur S, Shusterman AJ, Hansch C (2002) J Chem Soc Perk Trans 2: 1112—1117.