Gas phase carbon acidity and correlations with aqueous pK_a values: A comparison of B3LYP/6-311++G(d,p), Gaussian-4 (G4), and experimental approaches

Sierra Rayne^{a,*} and Kaya Forest^b

The gas and aqueous phase pK_a values of carbon Bronsted acids are of considerable theoretical and applied interest.[1-3] The pK_a values of carbon acids can vary widely (up to >60 units) depending on the electron withdrawing or releasing ability of substituents attached to the reaction center.[4,5] Because of the difficulty and expense in directly measuring carbon acid pK_a values, several theoretical approaches have been pursued to better predict corresponding acidity constants in the gas and aqueous phases.[6-10]

In recent work, Charif et al. [10] have employed the B3LYP/6-311++G(d,p) density functional level of theory to estimate gas phase standard state (298.15 K, 1 atm) free energies of acid dissociation ($\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$) for a range of 21 carbon acids, and subsequently correlated the B3LYP/6-311++G(d,p) $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ values to experimental aqueous pK_a measurements. A number of the carbon acids investigated by Charif et al. [10] have experimental $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ values in the NIST database [11], although these authors did not compare their B3LYP/6-311++G(d,p) estimates to the experimental reports. To better understand the accuracy of the B3LYP/6-311++G(d,p) method for gas phase carbon acidity estimation and potential quantitative structure-property relationship (QSPR) application towards aqueous phase pK_a prediction, we have conducted both B3LYP/6-311++G(d,p) [12-19] and Gaussian-4 (G4) [20] level calculations using Gaussian 09 [21] on a subset of 10 carbon acids from Charif et al. [10] having reliable experimental data for comparison (Figure 1). Calculations used gas phase starting geometries obtained with the PM6 semiempirical method [22] as implemented in MOPAC 2009 (<http://www.openmopac.net/>; v. 9.281). All molecular enthalpies and free energies include zero point and thermal corrections. Only the lowest energy conformation of each molecule was considered. All optimized structures were confirmed as true minima by vibrational analysis at the same level.

Fig. 1. Structures of the compounds under consideration.

Excellent agreement having effective chemical accuracy was obtained between the G4 $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ estimates and the experimental dataset, with a mean signed deviation (MSD) of 2.4 kJ/mol, a mean absolute deviation (MAD) of 4.6 kJ/mol, and a root mean squared deviation (RMSD) of 5.2 kJ/mol (Table 1). Our B3LYP/6-311++G(d,p) calculations yielded higher $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ errors, with MSD / MAD / RMSD of -7.8 / 8.9 / 11.3 kJ/mol. By comparison, the B3LYP/6-311++G(d,p) calculations of Charif et al. [10] on these 10 compounds yield MSD / MAD / RMSD of 2.1 / 19.4 / 34.2 kJ/mol.

Table 1. B3LYP/6-311++G(d,p) and G4 estimated and corresponding experimental gas phase standard state (298.15 K, 1 atm) free energies of acid dissociation ($\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$) for various carbon acids. Values are in kJ/mol.

carbon acid	B3LYP/ 6-311++G(d,p)	G4	expt. ^a
ethane	1718.7 ^b / 1720.2 ^c	1728.0	1723.0±8.8
methane	1712.3 / 1712.2	1723.0	1715±15
cyclopropane	1695.9 / 1695.6	1696.5	1690±30
ethylene	1669.3 / 1672.2	1678.4	1677.8±2.1
toluene	1665.7 / 1560.6	1568.7	1564.0±8.4
trifluoromethane	1530.0 / 1529.8	1555.9	1549.0±6.3
acetylene	1543.2 / 1544.9	1553.0	1550±20
cyclopentadiene	1447.9 / 1447.8	1450.5	1455.0±8.4
dimedone	1361.7 / 1361.5	1378.7	1385.0±8.4
isopropylidene malonate	1344.5 / 1344.9	1359.1	1359.0

^a from ref. [11]. ^b from ref. [10]. ^c current work.

There is clearly a problem with the reported B3LYP/6-311++G(d,p) $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ for toluene by Charif et al. [10] (underlined for emphasis in Table 1). These authors

give a $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ of 1665.7 kJ/mol that is in error by 101.7 kJ/mol compared to the experimental value [11] of 1564.0±8.4 kJ/mol. Our B3LYP/6-311++G(d,p) calculations obtain a $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ of 1560.6 kJ/mol, or 105.1 kJ/mol lower than reported by Charif et al. [10] at this level of theory, and only 3.4 kJ/mol lower than the experimental value. The problem appears to be with the molecular free energy Charif et al. [10] report for the benzylic anion of toluene. Our B3LYP/6-311++G(d,p) free energy for the undissociated toluene (-271.540626 Hartrees) agrees with the corresponding value reported in the Supplementary Materials by Charif et al. [10] (-271.540629 Hartrees), but our free energy of the benzylic anion at this level of theory (-270.936151 Hartrees) is substantially lower than that of Charif et al. [10] (-270.896168 Hartrees). With the exception of the toluene outlier, our B3LYP/6-311++G(d,p) results differed from those of Charif et al. by between -0.3 and +2.8 kJ/mol. We also note that Charif et al. [10] reported a B3LYP/6-311++G(d,p) $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ of 1408.4 kJ/mol for diethyl malonate, whereas the NIST database [11] cites an experimental value of 1432.0±8.4 kJ/mol.

The G4 enthalpies of acid dissociation ($\Delta_{\text{acid}}H^{\circ}_{(\text{g})}$) are also in excellent agreement with the experimental dataset, having MSD / MAD / RMSD of 1.1 / 4.8 / 5.7 kJ/mol (Table 2). As with the $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ estimates, the B3LYP/6-311++G(d,p) method displays a lower level of accuracy, having $\Delta_{\text{acid}}H^{\circ}_{(\text{g})}$ MSD / MAD / RMSD of -9.0 / 10.6 / 12.9 kJ/mol.

Table 2. B3LYP/6-311++G(d,p) and G4 estimated and corresponding experimental gas phase standard state (298.15 K, 1 atm) enthalpies of acid dissociation ($\Delta_{\text{acid}}H^{\circ}_{(\text{g})}$) for various carbon acids. Values are in kJ/mol.

carbon acid	B3LYP/ 6-311++G(d,p)	G4	expt. ^a
ethane	1752.6	1759.8	1758.0±8.4
methane	1743.6	1754.4	1749±15
cyclopropane	1728.1	1728.9	1720±20
ethylene	1704.0	1710.2	1712.9±2.5
toluene	1596.1	1604.0	1600.0±5.0
trifluoromethane	1563.6	1589.7	1582.0±5.9
acetylene	1576.8	1584.0	1590±20
cyclopentadiene	1479.8	1482.4	1481.0±9.2
dimedone	1391.3	1408.1	1418.0±8.8
isopropylidene malonate	1374.2	1388.6	1388.7±0.4

^a from ref. [11].

One application of theoretical $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ estimates is their potential correlation with pK_a values,[10] allowing less expensive gas phase calculations of carbon acids as a proxy for estimated aqueous acidity constants. For the 10 carbon acids being studied, the experimental pK_a values range over 45 units between 50 and 5 (Table 3).

Table 3. Aqueous pK_a values of the compounds under consideration. Values were taken from ref. [10], citing ref [5,23].

carbon acid	expt. pK _a
ethane	50.0
methane	48.0
cyclopropane	46.0
ethylene	44.0
toluene	41.0
trifluoromethane	32.0
cyclopentadiene	29.0
acetylene	25.0
cyclopentadiene	15.0
dimedone	5.3
isopropylidene malonate	4.8

We have plotted the B3LYP/6-311++G(d,p) $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ reported by Charif et al. [10] (including their erroneous toluene $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$; Figure 2(a)) and our B3LYP/6-311++G(d,p) $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ (Figure 2(b)), as well as the G4 (Figure 2(c)) and experimental (Figure 2(d)) $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$, against the corresponding experimental pK_a values. The following best fit linear regression equations are obtained (error bars are standard errors):

$$\text{pK}_a = 0.120 \pm 0.005 \Delta_{\text{acid}}G^{\circ}_{(\text{g})} - 158 \pm 9 \text{ (B3LYP/6-311++G(d,p); ref. [10])} \dots (1)$$

$$r = 0.991, n = 10, s = 2.46$$

$$\text{pK}_a = 0.121 \pm 0.010 \Delta_{\text{acid}}G^{\circ}_{(\text{g})} - 158 \pm 16 \text{ (B3LYP/6-311++G(d,p); current work)} \dots (2)$$

$$r = 0.973, n = 10, s = 4.33$$

$$\text{pK}_a = 0.124 \pm 0.010 \Delta_{\text{acid}}G^{\circ}_{(\text{g})} - 164 \pm 16 \text{ (G4)} \dots (3)$$

$$r = 0.975, n = 10, s = 4.12$$

$$\text{pK}_a = 0.127 \pm 0.011 \Delta_{\text{acid}}G^{\circ}_{(\text{g})} - 168 \pm 17 \text{ (experimental)} \dots (4)$$

$$r = 0.973, n = 10, s = 4.31$$

Fig. 2. Correlations between B3LYP/6-311++G(d,p), Gaussian-4 (G4), and experimental $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ and the corresponding experimental aqueous pK_{a} values (error bars on the experimental $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ are from the original literature sources). The datapoint for toluene is highlighted in each plot as a filled circle.

While the G4 and our corrected B3LYP/6-311++G(d,p) theoretical $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ both display strong correlations with experimental aqueous pK_{a} values in a manner consistent with the experimental $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ dataset, the lower $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ prediction accuracy of the B3LYP/6-311++G(d,p) method relative to higher level G4 calculations, as demonstrated herein, requires consideration when constructing carbon acid aqueous pK_{a} QSPRs. In addition, lower carbon acid pK_{a} prediction accuracy via $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ based QSPRs than implied by Charif et al. [10] should be expected using both low and high level theoretical methods. For example, with our reduced set of carbon acids, we obtain a regression standard error at the B3LYP/6-311++G(d,p) level that is nearly twice as large as obtained by Charif et al. [10] at the same level of theory for the same compounds, [10] and about equal to that we obtained using both high level G4 results and with the experimental $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ dataset. This difference at the B3LYP/6-311++G(d,p) level between the two studies is primarily due to the erroneous $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ these authors obtained for toluene, which was highly favorable in terms of their QSPR model quality-of-fit (i.e., note different locations of the toluene datapoint in Figure 2(a) and (b) between the two studies). Thus, further theoretical efforts are warranted to determine what other carbon acid functional groups may be outliers that confound accurate applications of computationally estimated $\Delta_{\text{acid}}G^{\circ}_{(\text{g})}$ based carbon acid aqueous pK_{a} QSPRs.

Acknowledgements

This work was made possible by the facilities of the Western Canada Research Grid (WestGrid:www.westgrid.ca; project 100185), the Shared Hierarchical Academic Research Computing Network (SHARCNET:www.sharcnet.ca; project sn4612), and Compute/Calcul Canada.

References and Notes

- ^a Ecologica Research, 412-3311 Wilson Street, Penticton, British Columbia, V2A 8J3, Canada
- ^b Department of Chemistry, Okanagan College, 583 Duncan Avenue West, Penticton, British Columbia, V2A 8E1, Canada
- * Corresponding author: rayne.sierra@gmail.com

- [1] Cram, D.J. "Fundamentals of Carbanion Chemistry.", Academic Press: New York, NY, USA, 1965.
- [2] Jones, J.R. "The Ionization of Carbon Acids.", Academic Press: London, UK, 1973.
- [3] Carey, F.A.; Sundberg, R.J. "Advanced Organic Chemistry: Part A: Structure and Mechanisms.", Springer: New York, NY, USA, 2007.
- [4] Serjeant, E.P.; Dempsey, B. "Ionization Constants of Organic Acids on Aqueous Solution (IUPAC Chemical Data Series).", Pergamon: New York, NY, USA, 1979.

- [5] Bordwell, F.G. "Equilibrium acidities in dimethylsulfoxide solution." *Acc. Chem. Res.* **1988**, 21, 456-463.
- [6] Choho, K.; Van Lier, G.; Van de Woude, G.; Geerlings, P. "Acidity of hydrofullerenes: A quantum chemical study." *J. Chem. Soc., Perkin Trans. 2* **1996**, 1723-1732.
- [7] Rezende, M.C. "A theoretical HSAB study of the acidity of carbon acids CH_3Z ." *J. Braz. Chem. Soc.* **2001**, 12, 73-80.
- [8] Rezende, M.C. "The acidity of carbon acids in aqueous solutions: correlations with theoretical descriptors." *Tetrahedron* **2001**, 57, 5923-5930.
- [9] Duchowicz, P.R.; Castro, E.A. "QSPR study of the acidity of carbon acids in aqueous solutions." *Mendeleev Comm.* **2002**, 12, 187-189.
- [10] Charif, I.E.; Mekelleche, S.M.; Villemain, D.; Mora-Diez, N. "Correlation of aqueous pK_{a} values of carbon acids with theoretical descriptors: A DFT study" *J. Mol. Struct. THEOCHEM* **2007**, 818, 1-6.
- [11] Lias, S.G.; Bartmess, J.E.; Liebman, J.F.; Holmes, J.L.; Levin, R.D.; Mallard, W.G. "Ion energetics data." In: *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899, <http://webbook.nist.gov>, (retrieved April 19, 2010).
- [12] Becke, A.D. "Density functional thermochemistry. III. The role of exact exchange." *J. Chem. Phys.* **1993**, 98, 5648-5652.
- [13] Lee, C.; Yang, W.; Parr, R.G. "Development of the Colle-Salvetti correlation energy formula into a functional of the electron density." *Phys. Rev. B* **1988**, 37, 785-789.
- [14] Ditchfield, R.; Hehre, W.J.; Pople, J.A. "Self-consistent molecular orbital methods. IX. An extended Gaussian-type basis for molecular orbital studies of organic molecules." *J. Chem. Phys.* **1971**, 54, 724-728.
- [15] Hehre, W.J.; Ditchfield, R.; Pople, J.A. "Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules." *J. Chem. Phys.* **1972**, 56, 2257-2261.
- [16] Hariharan, P.C.; Pople, J.A. "Accuracy of AHn equilibrium geometries by single determinant molecular orbital theory." *Mol. Phys.* **1974**, 27, 209.
- [17] Gordon, M.S. "The isomers of silacyclopentane." *Chem. Phys. Lett.* **1980**, 76, 163-168.
- [18] Hariharan, P.C.; Pople, J.A. "The influence of polarization functions on molecular orbital hydrogenation energies." *Theor. Chem. Acc.* **1973**, 28, 213-222.
- [19] Franci, M.M. "Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements." *J. Chem. Phys.* **1982**, 77, 3654-3654.
- [20] Curtiss, L.A.; Redfern, P.C.; Raghavachari, K. "Gaussian-4 theory." *J. Chem. Phys.* **2007**, 126, 084108.
- [21] Gaussian 09, Revision A.1, Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J.L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J.A.; Peralta, J.E.; Ogliaro, F.; Bearpark, M.; Heyd, J.J.; Brothers, E.; Kudin, K.N.; Staroverov, V.N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J.C.; Iyengar, S.S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N.J.; Klene, M.; Knox, J.E.; Cross, J.B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R.E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.; Ochterski, J.W.; Martin, R.L.; Morokuma, K.; Zakrzewski, V.G.; Voth, G.A.; Salvador, P.; Dannenberg, J.J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J.B.; Ortiz, J.V.; Cioslowski, J.; Fox, D.J. Gaussian, Inc., Wallingford, CT, USA, 2009.
- [22] Stewart, J.J.P. "Optimization of parameters for semiempirical methods. V. Modification of NDDO approximations and application to 70 elements." *J. Mol. Model.* **2007**, 13, 1173-1213.
- [23] Alkorta, I.; Elguero, J. "Carbon acidity and ring strain: A hybrid HF-DFT approach (Becke3LYP/6-311++G**)." *Tetrahedron* **1997**, 53, 9741-9748.

Appendix Table S1. Gas phase standard state (298.15 K; 1 atm) molecular enthalpies and free energies at the B3LYP/6-311++G(d,p) and G4 levels of theory for the compounds under consideration.

	B3LYP/6-311++G(d,p)				G4			
	molecular acid		dissociated anion		molecular acid		dissociated anion	
	H° _(g) (Hartrees)	G° _(g) (Hartrees)	H° _(g) (Hartrees)	G° _(g) (Hartrees)	H° _(g) (Hartrees)	G° _(g) (Hartrees)	H° _(g) (Hartrees)	G° _(g) (Hartrees)
ethane	-79.777838	-79.805388	-79.112681	-79.140121	-79.733660	-79.761239	-79.065736	-79.093024
methane	-40.485610	-40.509087	-39.823866	-39.846881	-40.461489	-40.484972	-39.795635	-39.818631
cyclopropane	-117.845716	-117.874327	-117.189869	-117.218425	-117.785822	-117.814463	-117.129672	-117.158238
ethylene	-78.560770	-78.586936	-77.914113	-77.939969	-78.517874	-78.544054	-77.868840	-77.894711
toluene	-271.505365	-271.540626	-270.899819	-270.936151	-271.371096	-271.406459	-270.762531	-270.798889
trifluoromethane	-338.319827	-338.350393	-337.726626	-337.757653	-338.150202	-338.180734	-337.547079	-337.578063
acetylene	-77.325928	-77.349289	-76.727704	-76.750784	-77.286490	-77.310021	-76.685517	-76.708444
cyclopentadiene	-194.059102	-194.090803	-193.497849	-193.529303	-193.964215	-193.995967	-193.401967	-193.433422
dimedone	-462.469044	-462.514789	-461.941485	-461.986135	-462.250871	-462.296890	-461.716894	-461.761700
isopropylidene malonate	-534.383331	-534.428282	-533.862285	-533.905962	-534.122050	-534.167103	-533.595514	-533.639378