

Accuracy of computational solvation free energies for neutral and ionic compounds: Dependence on level of theory and solvent model

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

^a Ecologica Research, 301-1965 Pandosy Street, Kelowna, British Columbia, Canada V1Y 1R9

^b Department of Chemistry, 583 Duncan Avenue West, Okanagan College, Penticton, British Columbia, Canada V2A 8E1

* Corresponding author. Tel.: +1 250 487 0166. E-mail address: rayne.sierra@gmail.com (S. Rayne).

Abstract

Gas to aqueous phase standard state (1 atm→1 mol/L; 298.15 K) free energies of solvation ($\Delta G^{\circ}_{\text{solv}}$) were calculated for a range of neutral and ionic inorganic and organic compounds using various levels and combinations of Hartree-Fock and density functional theory (DFT) and composite methods (CBS-Q//B3, G4MP2, and G4) with the IEFPCM-UFF, CPCM, and SMD solvation models in Gaussian 09 (G09). For a subset of highly polar and generally polyfunctional neutral organic compounds previously identified as problematic for prior solvation models, we find significantly reduced $\Delta G^{\circ}_{\text{solv}}$ errors using the revised solvent models in G09. The use of composite methods for these compounds also substantially reduces their apparent $\Delta G^{\circ}_{\text{solv}}$ errors. In contrast, no general level of theory effects between the B3LYP/6-31+G** and G4 methods were observed on a suite of simpler neutral, anionic, and cationic molecules commonly used to benchmark solvation models. Further investigations on mono- and polyhalogenated short chain alkanes and alkenes and other possibly difficult functional groups also revealed significant $\Delta G^{\circ}_{\text{solv}}$ error reductions by increasing the level of theory from DFT to G4. Future solvent model benchmarking efforts should include high level composite method calculations to allow better discrimination of potential error sources between the levels of theory and the solvation models.

Keywords: Solvation free energy; Solvation models; Hartree-Fock; Density functional theory; Composite methods

The continuing development of implicit solvent models for predicting solvation free energies (ΔG_{solv}) is driven by their importance in better understanding and estimating reaction rates, mechanisms, and equilibria in solution, partitioning in biological, environmental, and engineered systems, and fundamental aspects of biological and medicinal chemistry.^[1, 2] In recent work, a set of 54 highly polar and generally polyfunctional organic compounds with available experimental solvation free energies was proposed against which current and future solvation models should be benchmarked.^[3] Using various self-consistent reaction field (SCRF) based solvation models (PCM, CPCM, DPCM, IEFPCM, IPCM, and SCIPCM) in Gaussian 03 (G03) and Gaussian 98 (G98), relatively poor ΔG_{solv} prediction performance was reported for the models, notably with halogenated and heteratom substituted compounds. However, many of the proposed compounds contain substantial conformational complexity, resulting in significant work to ensure that the lowest energy gas and aqueous phase conformers (which could differ from each other) are found when evaluating all potential combinations of energetic contributions from low-lying states towards the net solvation free energy. In addition, some compounds are hydrolyzable (e.g., methanesulfonyl chloride; carboxylate, phosphate, and sulfate esters), which could preclude accurate assessment of their experimental solvation energies. Numerous other members contain aliphatic and aromatic amines (e.g., imidazole), which can have relevant acid/base behavior at near neutral aqueous pH values, further complicating reliable experimental solvation energy measurements against which to assess theoretical estimates.

From this dataset of 54 compounds, we identified a subset of 9 compounds (Table 1) that maintain a reasonable degree of polyfunctionality and polarity, are relatively constrained with regard to low energy conformational freedom, lack potentially confounding acid/base behavior in near neutral aqueous solutions that solvation models attempt to mimic, and which display approximately the same error distribution as the parent dataset. For example, at the HF/6-31+G** and B3LYP/6-31+G** levels

of theory^[4-12] with the parent n=54 dataset, Guthrie and Povar^[3] found root mean squared deviations (RMSDs) from experimental gas to aqueous phase free energies of solvation ($\Delta G^{\circ}_{\text{solv}}$; denoting transfer of solute at 298.15 from 1 atm in the ideal gas phase to 1 mol/L in the ideal dilute solution aqueous phase) of 2.13 (IEFPCM-UAHF/HF/6-31+G**), 2.31 (IEFPCM-UAHF/B3LYP/6-31+G**), 2.09 (CPCM/HF/6-31+G**), and 2.02 kcal/mol (CPCM/B3LYP/6-31+G**) using the IEFPCM^[13] and CPCM^[14, 15] solvation models in G03.^[16] These authors state they typically used the default settings with the UAHF radii for specifying the IEFPCM molecular cavity in G03 and G98; the default PCM radii has changed to UFF in Gaussian 09 (G09; http://www.gaussian.com/g_tech/g_ur/k_scrf.htm). The corresponding $\Delta G^{\circ}_{\text{solv}}$ RMSDs for the 9 compound subset using these two levels of theory and two solvent models from Guthrie and Povar^[3] are 3.1, 3.2, 3.1, and 3.1 kcal/mol, respectively. Our subset thus does not discriminate against difficult compounds in the parent dataset based on the error metrics given in Guthrie and Povar;^[3] arguably, it concentrates the relative presence of difficult compounds. Our subset does omit some functional group classes discussed above due to uncertainty regarding the quality of the experimental $\Delta G^{\circ}_{\text{solv}}$ data being benchmarked against, and issues over conformational complexity and aqueous reactivity. The $\Delta G^{\circ}_{\text{solv}}$ span is 10.8 kcal/mol in the n=9 subset versus 13 kcal/mol in the original 54 compound dataset.

Our first investigation was to examine whether changes in the default IEFPCM^[13] and CPCM^[14, 15] solvent models from G03^[16] to G09^[17] resulted in different $\Delta G^{\circ}_{\text{solv}}$ prediction accuracies at the HF/6-31+G** and B3LYP/6-31+G** levels of theory^[4-12] for these 9 compounds (Table 1). Calculated ΔG^*_{solv} (transfer of solute at 298.15 from 1 mol/L in the ideal gas phase to 1 mol/L in the ideal dilute solution aqueous phase) were converted to $\Delta G^{\circ}_{\text{solv}}$ via the relationship, $\Delta G^{\circ}_{\text{solv}} = \Delta G^*_{\text{solv}} + RT \ln(\hat{R}T)$ where the term $RT \ln(\hat{R}T)$ is about 1.9 kcal/mol at 298.15 K. All calculations were conducted using G09,^[17] only the lowest energy conformation was considered, no imaginary frequencies were present,

free energies include both zero-point and thermal corrections, and geometries were optimized in the respective phases of interest. For all sets of data (both G03 [from ref. ^[3]] and G09 [present study]), the IEFPCM and CPCM solvent models yield effectively equivalent $\Delta G^{\circ}_{\text{solv}}$ and associated error metrics. The G09 versions of IEFPCM-UFF and CPCM also give substantially lower $\Delta G^{\circ}_{\text{solv}}$ errors than their G03 default counterparts, with reductions in MAD and RMSD by up to 50%. For our IEFPCM-UFF/B3LYP/6-31+G** calculations using G09, the MAD and RMSD are near chemical accuracy at 1.2 and 1.4 kcal/mol, respectively, with a maximum individual absolute error (MIAE) of 1.7 kcal/mol, reduced from the corresponding G03 data at the IEFPCM-UAHF/B3LYP/6-31+G** level from ref. ^[3] of 2.2 (MAD), 3.1 (RMSD), and 7.5 (MIAE) kcal/mol, respectively. Based on the updated error metrics at the IEFPCM-UFF/B3LYP/6-31+G** level in G09, one would no longer consider these particularly “difficult” compounds for solvation energy modeling. Thus, users of G09 solvation models with default settings will need to be aware that historical $\Delta G^{\circ}_{\text{solv}}$ benchmarking studies using G03 and earlier versions (see also, e.g., ^[18]) are not necessarily applicable to the most recent version of this software. We note that G09 has a “G03Defaults” option for solvent settings to allow users to reproduce default G03 solvent model calculations as closely as possible, but perfect agreement is not always possible due to revisions to the software between these two versions.

In general, prior $\Delta G^{\circ}_{\text{solv}}$ benchmarking studies have used DFT methods, and concluded that increasing basis set completeness either has no significant effect on the quality of $\Delta G^{\circ}_{\text{solv}}$ predictions, or in some cases may even decrease the predictive accuracy (see, e.g., ^[3, 18-22]). Consequently, we performed calculations using the HF and B3LYP model chemistries with the larger 6-311++G(d,p) basis set, and also added calculations at these levels using the SMD solvent model ^[20] in G09. To probe the effect of density functional on the results, we also employed the M062X/6-311++G(d,p) level of theory ^[7-12, 23] with the IEFPCM-UFF and SMD solvent models (Table 2). At the HF and DFT level, we see no

substantial difference in $\Delta G^{\circ}_{\text{solv}}$ prediction accuracy with increasing level of theory (although the IEFPCM-UFF/B3LYP/6-311++G(d,p) method reaches closer to near chemically accurate MAD/RMSD of 1.1/1.2 kcal/mol), but we observe a significant difference in error metrics between the B3LYP and M062X functionals at basis set equivalence using the SMD solvent model. The SMD solvent model also does not perform as well as the IEFPCM-UFF model using any level of theory considered for these compounds.

Calculations were then conducted using the CBS-Q//B3,^[24, 25] G4MP2,^[26] and G4^[27] composite methods in G09 with the IEFPCM-UFF and SMD solvent models (Table 3). At these levels of theory, chemical accuracy is obtained by the IEFPCM-UFF solvent model using all three methods, with G4 MAD/RMSD of 0.7/0.9 kcal/mol and no systematic error (MSD=0.0 kcal/mol). The maximum errors for outliers are also reduced from up to 4 kcal/mol for HF and DFT levels of theory to no larger than 1.5 kcal/mol with the composite methods. While still maintaining higher $\Delta G^{\circ}_{\text{solv}}$ errors for this suite of compounds, the SMD model errors are substantially reduced (RMSD reduction of about 33%, and loss of the systematic low $\Delta G^{\circ}_{\text{solv}}$ bias) in moving from B3LYP/6-311++G(d,p) up to the composite methods. The error difference between the M062X/6-311++G(d,p) and composite method calculations is more pronounced, with the RMSD reduced by more than 50% and the elimination of a substantial systematic underbias of 2.0 kcal/mol.

We found these results intriguing, and sought to examine whether the significant reduction in $\Delta G^{\circ}_{\text{solv}}$ errors moving from HF/DFT to composite methods with the IEFPCM-UFF and SMD solvent models for the subset of “difficult” neutral compounds would hold when applied against sets of “simpler” neutral and ionic compounds often employed in previous benchmarking studies. For this, we used a set of 18 anionic, 15 cationic, and 25 neutral inorganic and organic compounds and their corresponding

experimental ΔG_{solv}^* data from the Gaussian 98 (G98) and G03 CPCM benchmarking study of Takano and Houk^[18] and ΔG_{solv}^* compilation of Pliego and Riveros.^[28] ΔG_{solv}^* were converted to $\Delta G_{\text{solv}}^\circ$ as described above. Unlike the subset of “difficult” compounds, we find little difference in the overall error metrics for the 25 “simpler” neutral compounds between B3LYP/6-31+G** and G4 $\Delta G_{\text{solv}}^\circ$ estimates using either the IEFPCM-UFF or SMD solvent models (Table 4). However, substantial $\Delta G_{\text{solv}}^\circ$ differences (>1 kcal/mol) are observed at the B3LYP/6-31+G** and G4 $\Delta G_{\text{solv}}^\circ$ levels for the following two compounds (all values in kcal/mol): acetone, -3.5 (IEFPCM-UFF/B3LYP/6-31+G**), -1.3 (IEFPCM-UFF/G4), -4.3 (SMD/B3LYP/6-31+G**), -1.3 (SMD/G4); and water, -3.8 (IEFPCM-UFF/B3LYP/6-31+G**), -1.9 (IEFPCM-UFF/G4), -7.5 (SMD/B3LYP/6-31+G**), -5.0 (SMD/G4). The error trend with increasing level of theory for water also reverses with a change in solvent model, with B3LYP/6-31+G** \rightarrow G4 substantially decreasing the prediction accuracy (by 1.8 kcal/mol) using the IEFPCM-UFF model while substantially increasing the prediction accuracy (by 2.5 kcal/mol) with SMD model. These apparently isolated anomalous error trendings with level of theory and solvent model on simple and fundamental compounds complicate the generalizability of $\Delta G_{\text{solv}}^\circ$ benchmarking efforts.

All methods have substantial difficulty with anions (MAD \sim 12.5 kcal/mol for IEFPCM-UFF and \sim 10 to 10.5 kcal/mol for SMD) and cations (MAD \sim 15 kcal/mol for IEFPCM-UFF and \sim 5 to 5.5 kcal/mol for SMD), but display good predictive accuracy for neutral compounds (MAD $<$ 2 kcal/mol). In particular, the SMD solvent model exhibits chemical accuracy without systematic bias for the neutral subclass at the B3LYP/6-31+G** and G4 levels, with MSD/MAD of 0.0/0.8 and 0.9/1.0 kcal/mol, respectively. A modest improvement (MAD reduction of \sim 2 kcal/mol compared to IEFPCM-UFF) is obtained from the SMD model for anions, whereas the SMD model achieves an almost 3-fold reduction in MAD for cations versus the IEFPCM-UFF model. There is generally good agreement between the

two levels of theory within a particular solvent model for both anions and cations, with the exceptions of Br^- (IEFPCM-UFF), CH_3COO^- (SMD), $\text{HC}\equiv\text{C}^-$ (IEFPCM-UFF and SMD), OH^- (IEFPCM-UFF and SMD), and OOH^- (IEFPCM-UFF and SMD) for which differences of >1 kcal/mol were observed. The bulk error metrics for the ionic species also mask large individual $\Delta G^\circ_{\text{solv}}$ differences between the solvation models. Although the MAD between the IEFPCM-UFF and SMD models for anions differs by only ~ 2 kcal/mol across both levels of theory, the individual differences range up to 14 kcal/mol (and up to 19 kcal/mol for the cations). We also note that as part of their CPCM benchmarking study, Takano and Houk^[18] found similarly poor performance for the UFF cavity with anions and cations at the HF/6-31+G(d)//B3LYP/6-31+G(d) level, with the best performance (MAD ~ 3 to 5 kcal/mol) using the UAHF and UAKS cavities.

Finding that the IEFPCM-UFF model outperformed the SMD model for the subset of neutral “difficult” compounds, the SMD model outperformed the IEFPCM-UFF model for the neutral “simpler” compounds, and that level of theory effects were only clearly evident with the “difficult” compounds, we needed an additional suite of potentially “difficult” compounds upon which to draw some overall conclusions. A set of 25 compounds was chosen from the compendium in Marenich et al.,^[21] and calculations were conducted with the B3LYP/6-31+G** and G4 methods using the IEFPCM-UFF and SMD solvent models (Table 5). For the 16 mono- and polyhalogenated hydrocarbons (fluoro-, chloro-, and bromo- substituted methanes, ethanes, and ethenes) increasing the level of theory results in significant error reductions, much as we observed with the subset of 9 “difficult” compounds from ref.^[3] as presented in Table 3. The MAD is reduced by $\sim 50\%$ and 25% with the SMD and IEFPCM-UFF models, respectively, when the G4 method is used compared to the B3LYP/6-31+G** level of theory. The systematic $\Delta G^\circ_{\text{solv}}$ underbias is also eliminated at the SMD/G4 level, and substantially reduced at the IEFPCM-UFF/G4 level. No large changes in $\Delta G^\circ_{\text{solv}}$ (i.e., >1 kcal/mol) were observed when the

theoretical level was increased; rather, the overall error reduction with the G4 method generally involved relatively small error reductions for most compounds with no clear outliers.

With the known branching errors in DFT functionals (particularly B3LYP) for relative energies of linear and branch alkane isomers, one may expect analogous $\Delta G^\circ_{\text{solv}}$ errors present at the DFT level which are reduced with composite methods. Although we only considered n-hexane and 2-methylpentane, we know that the B3LYP branching error manifests itself already at the C₆ alkane homolog between these two isomers.^[29] However, we observe no difference in $\Delta G^\circ_{\text{solv}}$ (or the relative $\Delta G^\circ_{\text{solv}}$; $\Delta\Delta G^\circ_{\text{solv}}$) between these two compounds using the B3LYP/6-31+G** and G4 methods with either solvent model. This does not rule out branching errors in $\Delta G^\circ_{\text{solv}}$ at higher homologs or with more highly branched congeners, which should be systematically investigated as part of future solvent benchmarking efforts. There is no $\Delta G^\circ_{\text{solv}}$ prediction improvement with increasing level of theory for the cyclic compounds, and for cyclopentane, the G4 calculations result in a substantial decrease of accuracy. As with nitromethane in the initial 9 compound “difficult” subset, three additional nitroalkanes (nitroethane, 1-nitropropane, and 2-nitropropane) collectively display a significant improvement in $\Delta G^\circ_{\text{solv}}$ accuracy at the G4 level as compared to the B3LYP/6-31+G** method, particularly with the IEFPCM-UFF solvent model. These examples are sufficient to unambiguously demonstrate the possibility for HF/DFT methods to suggest problems in solvent models for some functional groups, whereas the use of higher level methods appears to lower the perceived $\Delta G^\circ_{\text{solv}}$ inaccuracies.

Overall, the findings show that the default IEFPCM-UFF and CPCM solvation models in G09 have significantly improved $\Delta G^\circ_{\text{solv}}$ prediction accuracy for a subset of highly polar/polyfunctional organic compounds relative to previous software versions. In addition, while increasing level of theory does not

improve $\Delta G^{\circ}_{\text{solv}}$ estimation performance for a representative standard set of neutral and ionic monofunctional common benchmarking compounds having alkanol, amine, thiol, nitrile, aldehyde, ketone, carboxylic acid, ester, ether, phenol, peroxide, and main group polyhydride moieties, higher level methods appear to significantly improve the $\Delta G^{\circ}_{\text{solv}}$ prediction capability against more polar and polyfunctional molecules containing nitro, sulfoxide, sulfonyl, and halogenated moieties. The results are perhaps intuitive, as low levels of HF/DFT methods can obtain reasonably accurate geometries and energies for simple organic compounds. On the other hand, a number of previous studies have shown that high-level ab initio and composite methods are required to achieve suitable performance against more problematic functionalities, particularly moieties such as polyhalogenated groups and highly branched hydrocarbons.^[30-45] Difficult compounds are not only troublesome from the theoretical perspective. Often, these compounds (notably the polyhalogenated members) pose experimental challenges for determining partitioning properties such as vapor pressure, solubility, and Henry's law constants (from which $\Delta G^{\circ}_{\text{solv}}$ is calculated, either directly or by proxy). Thus, as benchmarking $\Delta G^{\circ}_{\text{solv}}$ moves into difficult compounds at high levels of theory, as with other areas of computational thermodynamics (e.g., enthalpies of formation/isomerization), we may find the high level theoretical values are more reliable than the experimental estimates, while HF/DFT approaches do not reach the required accuracy.

Consequently, the rationale for continuing to use only HF/DFT calculations when benchmarking solvent models is not clear. If the goal is to isolate and benchmark the accuracy of solvation models, particularly across the broadest suite of functional groups possible, the highest practical levels of theory should be employed. We concede that high level calculations cannot be performed on larger molecular and macromolecular systems for which future solvation model development is often intended. In these cases, solvation inaccuracies will likely result from errors in both the reduced level of theory needed

for the computations to remain practical, as well as errors in the solvent model, and these sources will be difficult to separate (as well as the inherently more difficult and unreliable experimental $\Delta G^{\circ}_{\text{solv}}$ determinations for large molecules). When lower levels of theory are applied, it is not clear whether the model chemistry, basis set, or the solvation model is at fault when discrepancies from experimental data are encountered. Furthermore, the wide diversity of density functionals and basis sets, and likely equal diversity of $\Delta G^{\circ}_{\text{solv}}$ prediction capabilities, also complicates the task. Such ambiguity prevents discriminating whether the solvent model is not accurately modeling a particular functional group, or whether the level of theory is at issue. Similarly, good agreement with experimental data using HF/DFT approaches could result from error cancellation, thereby masking underlying fundamental problems with the solvent model.

Since the potentially problematic moieties (out of the large set of all possible functional groups) whose $\Delta G^{\circ}_{\text{solv}}$ accuracy is dependent on the use of high level methods cannot readily be identified in advance, and with the advent of high performance computing platforms and the relatively cost efficient CBS-Q//B3, G4MP2, and G4 composite methods for compounds with <10 to 15 heavy atoms (and practicality of the W1BD method ^[46] for systems with <6 heavy atoms), the sole use of HF and DFT approaches for benchmarking solvation models against neutral and ionic species should be discontinued. Instead, high level calculations should complement HF/DFT solvent benchmarking efforts to the greatest degree possible. These types of multiple theory level benchmarks will help the community best identify the combination of method and solvent model suitable for the task at hand, and better assist in the accurate parametrization of existing and proposed solvation models.

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 aqn-965), and Compute/Calcul Canada.

References

- [1] C.J. Cramer, D.G. Truhlar, *Chem. Rev.* 99 (1999) 2161.
- [2] J.P. Guthrie, *J. Phys. Chem. B* 113 (2009) 4501.
- [3] J.P. Guthrie, I. Povar, *Can. J. Chem.* 87 (2009) 1154.
- [4] A.D. Becke, *J. Chem. Phys.* 98 (1993) 5648.
- [5] C. Lee, W. Yang, R.G. Parr, *Phys. Rev. B* 37 (1988) 785.
- [6] B. Miehlich, A. Savin, H. Stoll, H. Preuss, *Chem. Phys. Lett.* 157 (1989) 200.
- [7] R. Ditchfield, W.J. Hehre, J.A. Pople, *J. Chem. Phys.* 54 (1971) 724.
- [8] P.C. Hariharan, J.A. Pople, *Mol. Phys.* 27 (1974) 209.
- [9] P.C. Hariharan, J.A. Pople, *Theor. Chem. Acc.* 28 (1973) 213.
- [10] W.J. Hehre, R. Ditchfield, J.A. Pople, *J. Chem. Phys.* 56 (1972) 2257.
- [11] A.D. McLean, G.S. Chandler, *J. Chem. Phys.* 72 (1980) 5639.
- [12] K. Raghavachari, J.S. Binkley, R. Seeger, J.A. Pople, *J. Chem. Phys.* 72 (1980) 650.
- [13] J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* 105 (2005) 2999.
- [14] V. Barone, M. Cossi, *J. Phys. Chem. A* 102 (1998) 1995.
- [15] M. Cossi, N. Rega, G. Scalmani, V. Barone, *J. Comp. Chem.* 24 (2003) 669.
- [16] Gaussian 03, Revision C.02, M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery, T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui,

A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, and J.A. Pople, Gaussian, Inc., Wallingford CT, 2004.

[17] Gaussian 09, Revision A.1, M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery, J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, and D.J. Fox, Gaussian, Inc., Wallingford CT, 2009.

[18] Y. Takano, K.N. Houk, *J. Chem. Theory Comput.* 1 (2005) 70.

[19] C.P. Kelly, C.J. Cramer, D.G. Truhlar, *J. Phys. Chem. B* 110 (2006) 16066.

[20] A.V. Marenich, C.J. Cramer, D.G. Truhlar, *J. Phys. Chem. B* 113 (2009) 6378.

[21] A.V. Marenich, C.J. Cramer, D.G. Truhlar, *J. Chem. Theory Comput.* 5 (2009) 2447.

[22] C.P. Kelly, C.J. Cramer, D.G. Truhlar, *J. Chem. Theory Comput.* 1 (2005) 1133.

[23] Y. Zhao, D. Truhlar, *Theor. Chem. Acc.* 120 (2008) 215.

[24] J.A. Montgomery, M.J. Frisch, J.W. Ochterski, G.A. Petersson, *J. Chem. Phys.* 110 (1999) 2822.

[25] J.A. Montgomery, M.J. Frisch, J.W. Ochterski, G.A. Petersson, *J. Chem. Phys.* 112 (2000) 6532.

[26] L.A. Curtiss, P.C. Redfern, K. Raghavachari, *J. Chem. Phys.* 127 (2007) 124105.

- [27] L.A. Curtiss, P.C. Redfern, K. Raghavachari, *J. Chem. Phys.* 126 (2007) 84108.
- [28] J.R. Pliego, J.M. Riveros, *Phys. Chem. Chem. Phys.* 4 (2002) 1622.
- [29] S. Rayne, K. Forest, *J. Mol. Struct. THEOCHEM* 941 (2010) 107.
- [30] J. Tirado-Rives, W.L. Jorgensen, *J. Chem. Theory Comput.* 4 (2008) 297.
- [31] S. Grimme, M. Steinmetz, M. Korth, *J. Org. Chem.* 72 (2007) 2118.
- [32] S. Rayne, K. Forest, *J. Mol. Struct. THEOCHEM* 948 (2010) 102.
- [33] G.A. Shamov, P.H.M. Budzelaar, G. Schreckenbach, *J. Chem. Theory Comput.* 6 (2010) 477.
- [34] K.U. Goss, G. Bronner, *J. Phys. Chem. A* 110 (2006) 9518.
- [35] K.U. Goss, G. Bronner, T. Harner, M. Hertel, T.C. Schmidt, *Environ. Sci. Technol.* 40 (2006) 3572.
- [36] H.P.H. Arp, C. Niederer, K.U. Goss, *Environ. Sci. Technol.* 40 (2006) 7298.
- [37] S. Rayne, K. Forest, *J. Environ. Sci. Health A* 44 (2009) 1145.
- [38] C.E. Check, T.M. Gilbert, *J. Org. Chem.* 70 (2005) 9828.
- [39] S. Grimme, *Angew. Chem. Int. Ed.* 45 (2006) 4460.
- [40] K.W. Sattelmeyer, J. Tirado-Rives, W.L. Jorgensen, *J. Phys. Chem. A* 110 (2006) 13551.
- [41] M.D. Wodrich, C. Corminboeuf, P.V.R. Schleyer, *Org. Lett.* 8 (2006) 3631.
- [42] P.R. Schreiner, A.A. Fokin, R.A. Pascal, A. de Meijere, *Org. Lett.* 8 (2006) 3635.
- [43] Y. Zhao, D.G. Truhlar, *Org. Lett.* 8 (2006) 5753.
- [44] S. Grimme, M. Steinmetz, M. Korth, *J. Chem. Theory Comput.* 3 (2007) 42.
- [45] M.D. Wodrich, C. Corminboeuf, P.R. Schreiner, A.A. Fokin, P.V.R. Schleyer, *Org. Lett.* 9 (2007) 1851.
- [46] E.C. Barnes, G.A. Petersson, J.A. Montgomery, M.J. Frisch, J.M.L. *J. Chem. Theory Comput.* 5 (2009) 2687.
- [47] S. Cabani, P. Gianni, V. Mollica, L. Lepori, *J. Sol. Chem.* 10 (1981) 563.

Table 1. Comparison of experimental solvation free energies ($\Delta G^\circ_{\text{solv}}$) for 9 organic compounds and corresponding theoretical estimates at the HF/6-31+G** and B3LYP/6-31+G** levels of theory using the IEFPCM and CPCM solvent models in Gaussian 03 (G03) and Gaussian 09 (G09). Experimental and G03 $\Delta G^\circ_{\text{solv}}$ data taken from ref. [3]. Values in brackets represent deviations from experimental $\Delta G^\circ_{\text{solv}}$. Values are in kcal/mol.

compound	expt. [3]	HF/6-31+G**				B3LYP/6-31+G**			
		IEFPCM-UAHF-G03 [3]	IEFPCM-UFF-G09	CPCM-G03 [3]	CPCM-G09	IEFPCM-UAHF-G03 [3]	IEFPCM-UFF-G09	CPCM-G03 [3]	CPCM-G09
1,4-dioxane	-3.2	-6.8 [-3.6]	-2.6 [0.5]	-6.9 [-3.7]	-2.7 [0.5]	-6.4 [-3.2]	-2.2 [0.9]	-6.4 [-3.4]	-2.2 [0.9]
2,2,2-trifluoroethanol	-2.4	-6.7 [-4.3]	-4.5 [-2.1]	-6.8 [-4.4]	-4.6 [-2.2]	-6.3 [-3.9]	-4.1 [-1.7]	-6.4 [-4.0]	-4.2 [-1.8]
cyanobenzene	-2.2	-2.5 [-0.3]	-4.1 [-1.9]	-2.5 [-0.3]	-4.1 [-1.9]	-1.6 [0.6]	-3.3 [-1.1]	-1.6 [0.6]	-3.3 [-1.1]
dimethyl sulfide	0.4	0.3 [-0.0]	-0.9 [-1.2]	0.3 [-0.1]	-0.9 [-1.3]	0.6 [0.3]	-0.6 [-0.9]	0.6 [0.2]	-0.6 [-0.9]
dimethyl sulfoxide	-6.8	-0.8 [6.0]	-8.8 [-2.0]	-0.9 [5.9]	-8.8 [-2.0]	0.7 [7.5]	-6.8 [0.0]	0.5 [7.3]	-6.8 [0.0]
methyl t-butyl ether	-0.3	-0.0 [0.3]	-0.8 [-0.5]	-0.1 [0.2]	-0.8 [-0.5]	0.1 [0.4]	-0.6 [-0.3]	0.1 [0.4]	-0.6 [-0.3]
nitrobenzene	-2.2	-2.4 [-0.2]	-4.4 [-2.2]	-2.5 [-0.3]	-4.4 [-2.2]	-1.4 [0.8]	-3.6 [-1.3]	-1.5 [0.7]	-3.6 [-1.4]
nitromethane	-2.3	-5.0 [-2.7]	-3.9 [-1.6]	-5.0 [-2.7]	-5.2 [-2.9]	-3.6 [-1.4]	-4.0 [-1.7]	-3.6 [1.3]	-4.0 [-1.7]
sulfolane	-6.8	-10.5 [-3.7]	-10.1 [-3.3]	-10.6 [-3.8]	-10.2 [-3.4]	-8.7 [1.9]	-8.2 [-1.5]	-8.7 [-2.0]	-8.3 [-1.5]
	MSD ^a	-1.0	-1.6	-1.0	-1.8	-0.1	-0.5	-0.1	-0.5
	MAD ^b	2.4	1.7	2.4	1.9	2.2	1.2	2.2	1.2
	RMSD ^c	3.1	1.9	3.2	2.1	3.1	1.4	3.1	1.4

^a mean signed deviation. ^b mean absolute deviation. ^c root mean squared deviation.

Table 2. Comparison of experimental solvation free energies ($\Delta G^{\circ}_{\text{soln}}$) for 9 organic compounds and corresponding theoretical estimates at the HF/6-31+G**, HF/6-311++G(d,p), B3LYP/6-31+G**, B3LYP/6-311++G(d,p), and M062X/6-311++G(d,p) levels of theory using the IEFPCM-UFF and SMD solvent models in Gaussian 09 (G09). Experimental $\Delta G^{\circ}_{\text{soln}}$ data taken from ref. [3]. Values in brackets represent deviations from experimental $\Delta G^{\circ}_{\text{soln}}$. Values are in kcal/mol.

compound	expt. [3]	IEFPCM-UFF			SMD				
		HF/6-311++G(d,p)	B3LYP/6-311++G(d,p)	M062X/6-311++G(d,p)	HF/6-31+G**	HF/6-311++G(d,p)	B3LYP/6-31+G**	B3LYP/6-311++G(d,p)	M062X/6-311++G(d,p)
1,4-dioxane	-3.2	-2.5 [0.6]	-2.2 [0.9]	-2.4 [0.8]	-4.8 [-1.7]	-4.5 [-1.3]	-4.8 [-1.7]	-4.5 [-1.3]	-4.9 [-1.7]
2,2,2-trifluoroethanol	-2.4	-4.5 [-2.1]	-4.3 [-1.8]	-4.3 [-1.9]	-6.2 [-3.8]	-6.2 [-3.8]	-6.2 [-3.8]	-4.6 [-2.1]	-5.6 [-3.2]
cyanobenzene	-2.2	-4.0 [-1.8]	-3.3 [-1.1]	-3.3 [1.1]	-3.0 [-0.8]	-2.9 [-0.7]	-3.0 [-0.8]	-1.8 [0.4]	-2.0 [0.2]
dimethyl sulfide	0.4	-0.8 [-1.2]	-0.5 [-0.8]	-0.6 [-0.9]	0.7 [0.3]	0.7 [0.4]	0.7 [0.3]	1.0 [0.7]	0.9 [0.5]
dimethyl sulfoxide	-6.8	-9.1 [-2.3]	-7.0 [-0.2]	-7.6 [-0.7]	-12.8 [-6.0]	-13.4 [-6.6]	-12.8 [-6.0]	-10.3 [-3.5]	-11.3 [-4.5]
methyl t-butyl ether	-0.3	-0.8 [-0.5]	-0.6 [-0.3]	-0.6 [-0.3]	-0.3 [0.0]	-0.2 [0.1]	-0.3 [0.0]	-1.3 [-1.0]	0.0 [0.3]
nitrobenzene	-2.2	-4.4 [-2.2]	-3.5 [-1.3]	-3.5 [-1.3]	-3.1 [-0.9]	-3.1 [-0.9]	-3.1 [-0.9]	-1.5 [0.7]	-1.6 [0.6]
nitromethane	-2.3	-4.0 [-1.7]	-4.0 [-1.7]	-3.7 [-1.4]	-4.0 [-1.6]	-4.0 [-1.7]	-4.0 [-1.6]	-3.1 [-0.8]	-3.7 [-1.4]
sulfolane	-6.8	-10.4 [-3.6]	-8.4 [-1.6]	-9.7 [-2.9]	-15.9 [-9.1]	-16.6 [-9.8]	-15.9 [-9.1]	-13.1 [-6.3]	-16.2 [-9.4]
	MSD ^a	-1.6	-0.9	-1.1	-2.6	-2.7	-1.3	-1.5	-2.0
	MAD ^b	1.8	1.1	1.3	2.7	2.8	1.9	1.9	2.4
	RMSD ^c	2.0	1.2	1.4	4.0	4.2	2.5	2.6	3.7

^a mean signed deviation. ^b mean absolute deviation. ^c root mean squared deviation.

Table 3. Comparison of experimental solvation free energies ($\Delta G^\circ_{\text{solv}}$) for 9 organic compounds and corresponding theoretical estimates at the CBS-Q//B3, G4MP2, and G4 levels of theory using the IEFPCM-UFF and SMD solvent models in Gaussian 09 (G09). Experimental $\Delta G^\circ_{\text{solv}}$ data taken from ref. ^[3]. Values in brackets represent deviations from experimental $\Delta G^\circ_{\text{solv}}$. Values are in kcal/mol.

compound	expt. ^[3]	IEFPCM-UFF			SMD		
		CBS-Q//B3	G4MP2	G4	CBS-Q//B3	G4MP2	G4
1,4-dioxane	-3.2	-1.9 [1.3]	-1.7 [1.4]	-1.7 [1.4]	-3.7 [-0.5]	-3.4 [-0.2]	-3.4 [-0.2]
2,2,2-trifluoroethanol	-2.4	-3.7 [-1.3]	-3.3 [-0.9]	-3.3 [-0.9]	-3.7 [-1.3]	-4.9 [-2.5]	-4.9 [-2.5]
cyanobenzene	-2.2	-2.8 [-0.6]	-2.6 [-0.4]	-2.6 [-0.4]	-1.1 [1.1]	-0.9 [1.3]	-0.9 [1.3]
dimethyl sulfide	0.4	-0.2 [-0.6]	-0.2 [-0.6]	-0.2 [-0.6]	1.3 [0.9]	1.3 [0.9]	1.3 [0.9]
dimethyl sulfoxide	-6.8	-5.4 [1.4]	-5.3 [1.5]	-5.3 [1.5]	-7.4 [-0.6]	-7.5 [-0.7]	-7.6 [-0.8]
methyl t-butyl ether	-0.3	-0.4 [-0.1]	-0.3 [0.0]	-0.3 [0.0]	1.0 [1.3]	1.2 [1.5]	1.2 [1.5]
nitrobenzene	-2.2	-2.5 [-0.3]	-2.3 [-0.1]	-2.3 [-0.1]	0.2 [2.4]	0.5 [2.8]	0.5 [2.7]
nitromethane	-2.3	-3.4 [1.1]	-3.2 [-0.9]	-3.2 [-0.9]	-2.0 [0.3]	-1.5 [0.8]	-1.5 [0.8]
sulfolane	-6.8	-6.5 [0.3]	-6.4 [0.4]	-6.4 [0.4]	-9.9 [-3.1]	-9.6 [-2.8]	-9.6 [-2.8]
	MSD ^a	-0.1	0.0	0.0	0.1	0.1	0.1
	MAD ^b	0.8	0.7	0.7	1.3	1.5	1.5
	RMSD ^c	0.9	0.9	0.9	1.6	1.8	1.8

^a mean signed deviation. ^b mean absolute deviation. ^c root mean squared deviation.

Table 4. Comparison of experimental solvation free energies ($\Delta G^{\circ}_{\text{solv}}$) for a suite of 18 anionic, 15 cationic, and 25 neutral inorganic and organic compounds and corresponding theoretical estimates at the B3LYP/6-31+G** and G4 levels of theory using the IEFPCM-UFF and SMD solvent models in Gaussian 09 (G09). Experimental data taken from ref. [18, 28]. Values in brackets represent deviations from experimental $\Delta G^{\circ}_{\text{solv}}$. Values are in kcal/mol.

	IEFPCM-UFF			SMD	
	expt.	B3LYP/6-31+G**	G4	B3LYP/6-31+G**	G4
anions					
Br ⁻	-66.7	-61.7 [5.0]	-63.1 [3.6]	-50.6 [16.1]	-51.3 [15.4]
CH ₃ COO ⁻	-75.4	-62.5 [12.9]	-62.1 [13.3]	-69.2 [6.2]	-68.1 [7.3]
CH ₃ O ⁻	-93.3	-69.2 [24.1]	-68.6 [24.7]	-77.7 [15.6]	-77.9 [15.4]
CN ⁻	-65.7	-63.8 [1.9]	-64.3 [1.4]	-60.8 [4.9]	-61.4 [4.3]
Cl ⁻	-72.7	-67.1 [5.6]	-68.0 [4.7]	-62.6 [10.1]	-63.2 [9.5]
F ⁻	-103.1	-83.0 [20.1]	-82.4 [20.7]	-85.5 [17.6]	-84.8 [18.3]
HCC ⁻	-74.2	-64.7 [9.5]	-66.2 [8.0]	-64.8 [9.4]	-66.6 [7.6]
HCOO ⁻	-74.3	-63.6 [10.7]	-63.3 [11.0]	-68.1 [6.2]	-67.3 [7.0]
N ₃ ⁻	-68.8	-59.4 [9.4]	-59.3 [9.5]	-61.3 [7.5]	-61.1 [7.7]
OBr ⁻	-74.0	-63.5 [10.5]	-63.1 [10.9]	-61.6 [12.4]	-60.9 [13.1]
OCl ⁻	-78.8	-65.9 [12.9]	-65.4 [13.4]	-71.2 [7.6]	-70.0 [8.8]
C ₂ H ₅ O ⁻	-89.2	-65.7 [23.5]	-65.4 [23.8]	-74.1 [15.1]	-74.3 [14.9]
OH ⁻	-103.1	-81.0 [22.1]	-78.9 [24.2]	-92.9 [10.2]	-90.0 [13.1]
OOH ⁻	-95.4	-74.8 [20.6]	-73.6 [21.8]	-87.1 [8.3]	-85.2 [10.2]
O ⁻	-98.0	-78.7 [19.3]	-77.8 [20.2]	-89.0 [9.0]	-88.2 [9.8]
PH ₂ ⁻	-57.5	-59.8 [-2.3]	-60.4 [-2.9]	-59.2 [-1.7]	-59.9 [-2.4]
HS ⁻	-69.7	-64.7 [5.0]	-65.1 [4.6]	-60.0 [9.7]	-60.3 [9.4]
CH ₃ S ⁻	-71.8	-63.4 [8.4]	-63.4 [8.4]	-57.5 [14.3]	-57.4 [14.4]
	MSD ^a	12.2	12.3	9.9	10.2
	MAD ^b	12.4	12.6	10.1	10.5
	RMSD ^c	14.4	14.8	10.9	11.3
cations					
CH ₃ CH ₂ OH ₂ ⁺	-86.5	-60.7 [25.8]	-60.8 [25.7]	-73.7 [12.8]	-73.8 [12.7]
(CH ₃) ₂ C=OH ⁺	-74.9	-56.0 [18.9]	-55.9 [19.0]	-65.8 [9.1]	-64.8 [10.1]
CH ₃ CONH ₃ ⁺	-71.9	-63.0 [8.9]	-62.6 [9.3]	-73.9 [-2.0]	-73.0 [-1.1]
(CH ₃) ₂ NH ₂ ⁺	-66.7	-57.9 [8.8]	-57.8 [8.9]	-65.6 [1.1]	-65.4 [1.3]
CH ₃ NH ₃ ⁺	-74.6	-63.4 [11.2]	-63.6 [11.0]	-72.3 [2.3]	-72.2 [2.4]
CH ₃ OH ₂ ⁺	-91.2	-65.1 [26.1]	-65.2 [26.0]	-78.9 [12.3]	-78.6 [12.6]
(CH ₃) ₂ OH ⁺	-77.9	-58.4 [19.5]	-58.0 [19.9]	-67.0 [10.9]	-66.8 [11.1]
(CH ₃) ₂ SH ⁺	-62.6	-55.5 [7.1]	-55.3 [7.3]	-57.6 [5.0]	-57.3 [5.3]
C ₂ H ₅ NH ₃ ⁺	-71.1	-60.7 [10.4]	-60.7 [10.4]	-69.7 [1.4]	-69.6 [1.5]
H ₃ O ⁺	-108.3	-75.5 [32.8]	-75.5 [32.8]	-94.7 [13.6]	-94.3 [14.0]
HCONH ₃ ⁺	-80.6	-69.3 [11.3]	-68.8 [11.8]	-79.8 [0.8]	-78.8 [1.8]
NH ₄ ⁺	-83.3	-70.5 [12.8]	-70.6 [12.7]	-80.4 [2.9]	-80.5 [2.8]
C ₆ H ₅ NH ₃ ⁺	-70.9	-56.1 [14.8]	-55.8 [15.1]	-65.6 [5.3]	-65.1 [5.8]
CH ₃ CH ₂ CH ₂ NH ₃ ⁺	-69.6	-58.9 [10.7]	-58.8 [10.8]	-68.5 [1.1]	-68.3 [1.3]
(pyridine)H ⁺	-59.2	-51.1 [8.1]	-51.0 [8.2]	-58.0 [1.2]	-57.6 [1.6]
	MSD	15.1	15.3	5.2	5.5

	MAD	15.1	15.3	5.5	5.7
	RMSD	16.9	17.0	7.2	7.4
neutrals					
C ₂ H ₅ OH	-3.2	-1.9 [1.3]	-1.3 [1.8]	-3.7 [-0.5]	-2.7 [0.4]
(CH ₃) ₂ NH	-2.4	-0.5 [1.9]	-0.4 [2.0]	-2.0 [0.4]	-1.8 [0.6]
CH ₃ CHO	-1.6	-2.4 [-0.8]	-1.7 [-0.1]	-2.3 [-0.7]	-1.0 [0.6]
CH ₃ CN	-2.0	-3.7 [-1.7]	-3.2 [-1.2]	-2.6 [-0.6]	-1.8 [0.2]
CH ₃ COCH ₃	-2.0	-3.5 [-1.5]	-1.3 [0.6]	-4.3 [-2.3]	-1.3 [0.6]
CH ₃ COOH	-4.8	-3.4 [1.4]	-2.6 [2.2]	-4.6 [0.2]	-4.6 [0.2]
CH ₃ NH ₂	-2.7	-1.3 [1.4]	-0.9 [1.8]	-2.4 [0.3]	-1.8 [0.9]
CH ₃ OCH ₃	0.0	-0.6 [-0.6]	-0.2 [-0.2]	0.1 [0.1]	1.0 [1.0]
CH ₃ OH	-3.2	-2.0 [1.2]	-1.3 [1.9]	-3.4 [-0.2]	-2.3 [0.9]
CH ₃ SH	0.7	-0.6 [-1.3]	-0.1 [-0.8]	0.6 [0.0]	1.1 [0.5]
H ₂ O	-4.4	-3.8 [0.7]	-1.9 [2.5]	-7.5 [-3.1]	-5.0 [-0.6]
HBr	-1.5	0.1 [1.6]	0.5 [2.0]	1.2 [2.7]	1.4 [2.9]
HCCH	1.9	-0.3 [-2.2]	0.0 [-1.9]	1.7 [-0.2]	2.1 [0.3]
HCN	-1.3	-2.7 [-1.4]	-2.4 [-1.0]	-0.3 [1.0]	0.2 [1.5]
HCOOH	-5.1	-3.1 [2.0]	-2.4 [2.7]	-4.5 [0.6]	-3.1 [2.0]
HCl	-0.3	-0.4 [-0.1]	0.2 [0.5]	-0.1 [0.2]	0.5 [0.8]
HF	-5.7	-2.1 [3.5]	-1.5 [4.2]	-3.2 [2.5]	-2.3 [3.4]
HOBr	-2.7	-1.6 [1.1]	-0.9 [1.7]	-2.4 [0.3]	-1.7 [1.0]
HOCl	-3.8	-1.6 [2.3]	-0.9 [2.9]	-3.3 [0.6]	-2.3 [1.6]
HOOH	-6.7	-2.8 [3.9]	-2.2 [4.5]	-7.7 [-1.0]	-6.4 [0.3]
NH ₃	-2.4	-2.1 [0.3]	-1.3 [1.1]	-2.6 [-0.2]	-1.4 [1.0]
PH ₃	2.9	0.9 [-1.9]	1.3 [-1.6]	2.4 [-0.5]	2.9 [0.0]
C ₆ H ₅ OH	-4.7	-2.5 [2.2]	-2.0 [2.7]	-4.2 [0.5]	-3.3 [1.4]
C ₆ H ₅ SH	-0.7	-1.0 [-0.3]	-0.6 [0.1]	-0.2 [0.4]	0.3 [1.0]
SH ₂	1.5	-0.5 [-2.0]	0.2 [-1.3]	0.0 [-1.4]	0.7 [-0.8]
	MSD	0.4	1.1	0.0	0.9
	MAD	1.5	1.7	0.8	1.0
	RMSD	1.8	2.1	1.2	1.3

^a mean signed deviation. ^b mean absolute deviation. ^c root mean squared deviation.

Table 5. Comparison of experimental solvation free energies ($\Delta G^{\circ}_{\text{soln}}$) for a suite of 25 potentially “difficult” neutral organic compounds and corresponding theoretical estimates at the B3LYP/6-31+G** and G4 levels of theory using the IEFPCM-UFF and SMD solvent models in Gaussian 09 (G09). Experimental data taken from ref. ^[21, 47]. Values in brackets represent deviations from experimental $\Delta G^{\circ}_{\text{soln}}$. Values are in kcal/mol.

	expt.	IEFPCM-UFF		SMD	
		B3LYP/6-31+G**	G4	B3LYP/6-31+G**	G4
halocarbons					
fluoromethane	1.7	-0.7 [-2.3]	-0.2 [-1.9]	0.7 [-1.0]	1.4 [-0.2]
tetrafluoromethane	5.1	1.2 [-3.8]	1.3 [-3.7]	4.7 [-0.4]	4.9 [-0.2]
chloromethane	1.3	-0.3 [-1.7]	0.1 [-1.2]	0.6 [-0.7]	1.2 [-0.2]
dichloromethane	0.5	-0.8 [-1.4]	-0.3 [-0.8]	-0.6 [-1.2]	0.1 [-0.4]
chloroform	0.8	-0.2 [-1.1]	0.2 [-0.7]	0.3 [-0.5]	0.8 [0.0]
chlorofluoromethane	1.1	-1.1 [-2.2]	-0.5 [-1.6]	-0.5 [-1.6]	0.4 [-0.7]
chlorodifluoromethane	1.4	-0.5 [-1.9]	0.0 [-1.4]	0.7 [-0.7]	1.4 [0.0]
fluorotrichloromethane	2.7	1.0 [-1.8]	1.1 [-1.6]	3.4 [0.6]	3.5 [0.7]
dibromomethane	-0.2	-0.7 [-0.5]	-0.3 [-0.1]	-0.7 [-0.5]	-0.3 [-0.1]
tribromomethane	-0.1	-0.4 [-0.3]	0.0 [0.1]	-0.5 [-0.4]	-0.2 [-0.2]
1-chloro-2,2,2-trifluoroethane	2.0	-1.2 [-3.2]	-0.7 [-2.7]	-0.1 [-2.0]	0.7 [-1.2]
1,1,1-trichloroethane	1.7	-0.2 [-1.8]	0.2 [-1.4]	0.5 [-1.2]	1.1 [-0.6]
1,1,1,2-tetrachloroethane	0.8	-0.8 [-1.5]	-0.2 [-1.0]	-0.4 [-1.1]	0.4 [-0.4]
trichloroethene	1.5	0.4 [-1.1]	0.7 [-0.9]	2.2 [0.7]	2.5 [1.0]
tetrachloroethene	2.0	0.9 [-1.0]	1.1 [-0.9]	3.6 [1.7]	3.6 [1.7]
1,1,1-trifluoropropan-2-ol	-2.3	-2.6 [-0.4]	-2.2 [0.1]	-3.6 [-1.3]	-2.6 [-0.3]
	MSD ^a	-1.6	-1.2	-0.6	-0.1
	MAD ^b	1.6	1.2	1.0	0.5
	RMSD ^c	1.9	1.6	1.1	0.7
linear and branched alkanes					
n-hexane	4.4	1.4 [-3.0]	1.4 [-3.0]	3.4 [-1.0]	3.3 [-1.1]
2-methylpentane	4.4	1.4 [-3.0]	1.4 [-3.1]	4.0 [-0.4]	4.0 [-0.5]
cyclic					
cyclopropane	2.7	1.1 [-1.6]	1.2 [-1.5]	1.7 [-1.0]	1.8 [-0.9]
cyclopentane	3.1	1.8 [-1.3]	-0.4 [-3.5]	3.3 [0.2]	1.7 [-1.4]
cyclohexane	3.1	1.5 [-1.7]	1.6 [-1.5]	2.9 [-0.2]	2.9 [-0.2]

aziridine	-3.5	-1.8 [1.7]	-1.2 [2.3]	-6.9 [-3.4]	-6.3 [-2.8]
azetidine	-3.7	-0.8 [2.8]	-0.6 [3.1]	-4.0 [-0.4]	-3.7 [0.0]
tetrahydrofuran	-1.6	-1.0 [0.6]	-0.7 [0.8]	-1.2 [0.4]	-0.4 [1.2]
tetrahydropyran	-1.2	-0.7 [0.5]	-0.4 [0.8]	-1.3[-0.1]	-0.8 [0.5]
<hr/>					
nitro					
nitroethane	-1.8	-3.6 [-1.7]	-2.9 [-1.1]	-2.8 [-1.0]	-1.4 [0.4]
1-nitropropane	-1.4	-4.1 [-2.7]	-2.3 [-0.9]	-2.3 [-0.9]	-0.4 [1.0]
2-nitropropane	-1.2	-3.6 [-2.4]	-3.2 [-2.0]	-2.0 [-0.7]	-0.4 [0.8]

^a mean signed deviation. ^b mean absolute deviation. ^c root mean squared deviation.