J Comput Aided Mol Des (2012) 26:489–496
DOI 10.1007/s10822-012-9568-8

The SAMPL3 blind prediction challenge: transfer energy overview

Matthew T. Geballe · J. Peter Guthrie

Received: 9 December 2011 / Accepted: 21 March 2012 / Published online: 3 April 2012

Abstract Prediction of the free energy of solvation of a small molecule, or its transfer energy, is a necessary step along the path towards calculating the interactions between molecules that occur in an aqueous environment. A set of these transfer energies were gathered from the literature for series of chlorinated molecules with varying numbers of chlorines based on ethane, biphenyl, and dibenzo-p-dioxin. This focused set of molecules were then provided as a blinded challenge to assess the ability of current computational solvation methods to accurately model the interactions between water and increasingly chlorinated compounds. This was presented as part of the SAMPL3 challenge, which represented the fourth iterative blind prediction challenge involving transfer energies. The results of this exercise demonstrate that the field in general has difficulty predicting the transfer energies of more highly chlorinated compounds, and that methods seem to be erring in the same direction.

Keywords Solvation · Transfer energy · Blind prediction · Validation

Electronic supplementary material The online version of this article (doi:10.1007/s10822-012-9568-8) contains supplementary material, which is available to authorized users.

M. T. Geballe (✉)
OpenEye Scientific Software, Santa Fe, NM 87508, USA
e-mail: geballem@gmail.com

J. P. Guthrie
Department of Chemistry, University of Western Ontario, London, ON, Canada

Introduction

Solvation free energies are needed for many kinds of chemical investigations, yet there are comparatively few experimental values; only somewhat over 3,000 values are known out of the millions of organic compounds prepared to date. This puts a premium on computational methods, provided that these can supply reliable estimates. In particular there is wide interest in computational methods for calculating binding constants of organic molecules to proteins [1–3], which requires the difference in free energy between the molecule in solution solvated by water and the molecule in the binding site of the protein. Although calculating the free energy of solvation for the organic molecule may be the easier part of the overall computational problem, it is nonetheless essential that this part be reliable if the final result is to be reliable.

A blind challenge is the ideal method through which to compare how well current methods can perform in predicting experimental quantities (e.g. solvation free energies) for drug-like molecules. In such a challenge developers of computational solvation method are given the structures of some compounds for which solvation energies are known but either unpublished or relatively inaccessible. They then submit predictions before the answers are revealed. This approach has been used for over 10 years in the protein folding community [4, 5]. The appeal of blind prediction lies in its power to removes or minimizes many sources of intentional or unintentional bias on the parts of participants, as well as limiting the ability for training of methods or tweaking parameters to improve performance on a specific set.

The SAMPL3 challenge is the fourth iteration of this blind prediction experiment, and although the areas covered have varied in each instance of SAMPL, prediction of

solvation free energies (also known as transfer energies) has been a part of every challenge [6–8]. While previous data sets have attempted to provide participants with a broad set that is representative of the molecules involved in drug discovery, the organizers of SAMPL3 chose a different focus for the transfer set in the current challenge: exploring how methods handled increasing chlorination of some simple compounds. There were several reasons for this shift, the foremost being the difficulty in assembling a large, quality data set of "drug-like" molecules for which transfer energies were known, well-determined, and blinded given the dearth of new experimental work. Additionally, the availability of many transfer energies for poly-chlorinated compounds due to their importance in the fields of atmospheric and environmental chemistry, and the surprisingly erroneous results obtained by many participants on two heavily chlorinated compounds in the SAMPL2 challenge made this a particularly interesting challenge to present to the modeling field.

Methods

Challenge design

All participants were welcomed to take part by registering at the SAMPL website (sampl.eyesopen.com). Information about the challenge was posted on the website, sent to many commercial, industry, and academic parties, and posted on the computational chemistry list (CCL). The data sets were provided for download from the SAMPL website in either isomeric SMILES, SD format, or PDB format as appropriate, although it was noted to participants that the conformation, tautomerization, and protonation states provided were not intended to be representative. The data was released in a staggered fashion as it was curated and became available between October, 2010 and early April 2011. The final submission deadline was June 30th, 2011. As in previous SAMPLs, the experimental values shown in Table 1 were not released to participants until after the submission deadline. The initial analysis of each submission included the calculation of a variety of metrics to describe the performance of the prediction, as well as bootstrapping to estimate the effects of experimental error on the performance.

Transfer energies

The challenge sets consisted of three series of chlorinated compounds where increasing numbers of chlorines were introduced to a common carbon skeleton. The frameworks were ethane, biphenyl and dibenzo-p-dioxin. For ethane all 10 compounds from ethane to hexachloroethane were included; for biphenyl and dibenzo-p-dioxin the number of isomers becomes quite large so only a selection, biased to those for which papers with experimental data had been found, was included; 14 for biphenyl and 12 for dibenzo-p-dioxin.

The uncertainties in the average ΔG_s values are based on the following somewhat arbitrary and pessimistic guidelines. The reported uncertainties (where given) were used as a basis for weighted averaging (weighted mean calculation) [9]. The uncertainty in the weighted mean was adjusted as follows: (1) No value was less than 0.10; this represents approximately 10 % errors in each of the two measurements required for a solvation energy (either two concentrations for a partition constant or solubility and vapor pressure), which seems realistic. Examination of data in Table 2 of the supplemental data shows several examples where there are multiple values differing by more than the reported individual standard deviations. (2) Where the best literature value for a compound had an assigned uncertainty of 1.37 or 1.93 this was taken as the uncertainty of the average. These values are based on the convention that if there is no literature reference to a paper with experimental details then the value can only be trusted to be good to within an order of magnitude; this will sometimes be unjustifiably pessimistic and sometimes unjustifiably optimistic, but at least serves as a warning that these values (though based on critical reviews of the literature) are less well founded than would be desirable. These values come from a database of solvation energies now being constructed [11], but still incomplete. The ΔG_s values with uncertainties of 1.37 or 1.93 are based in whole or in part on cited values for which no experimental source has as yet been found. Where values are derived from solubility and vapor pressure values, and both are uncertain to within an order of magnitude (i.e. log p or log s is ± 1.0) then the uncertainty in ΔG_s will be 1.93 from the rules for propagation of error [10]; if one is judged more precise and the other is uncertain within an order of magnitude, then the uncertainty in ΔG_s will be 1.37. When the citation is to a Henry's Law constant or ΔG_s value the uncertainty in ΔG_s will be 1.93 because two experimental values had to be determined: either two concentrations or a concentration and a partial pressure. Where some values of ΔG_s for a compound are based on reported experimental results ("measured") and others are only cited with no experimental data ("cited") only the measured values are used in the final averaging. This is necessary because often numerous cited values can be traced back to the same original value, which may be experimental, or may be a source giving no experimental detail.

For compounds of low solubility in water, freeing the solution of suspended droplets or crystallites becomes important and difficult. The amount of such suspended

Table 1 Experimental solvation energies and errors for the SAMPL3 set

Molecule	Name	ΔG_s	Error
Ethane	xfer3.01	1.87	0.1
Chloroethane	xfer3.02	−0.39	0.1
1,1-dichloroethane	xfer3.03	−0.88	0.1
1,2-dichloroethane	xfer3.04	−1.8	0.1
1,1,1-trichloroethane	xfer3.05	−0.26	0.1
1,1,2-trichloroethane	xfer3.06	−1.97	0.1
1,1,1,2-tetrachloroethane	xfer3.07	−1.43	0.1
1,1,2,2-tetrachloroethane	xfer3.08	−2.37	0.1
Pentachloroethane	xfer3.09	−1.23	0.1
Hexachloroethane	xfer3.10	−0.64	0.1
Biphenyl	xfer3.11	−2.23	0.1
2-chlorobiphenyl	xfer3.12	−2.69	0.1
2,5-dichlorobiphenyl	xfer3.13	−2.46	0.1
2,4,6-trichlorobiphenyl	xfer3.14	−2.16	0.1
2,3,4,5-tetrachlorobiphenyl	xfer3.15	−3.48	1
2,2′,6,6′-tetrachlorobiphenyl	xfer3.16	−2.28	0.12
2′,3,4,5,5′-pentachlorobiphenyl	xfer3.17	−3.61	0.13
2,2′,4,6,6′-pentachlorobiphenyl	xfer3.18	−1.96	0.1
2,3,3′,4′,5,6-hexachlorobiphenyl	xfer3.19	−4.38	0.22
2,3,3′,4,4′,5-hexachlorobiphenyl	xfer3.20	−3.04	0.1
2,2′,3,3′,4,4′,5heptachlorobiphenyl	xfer3.21	−4.4	0.1
2,3,3′,4,4′,5,5′-heptachlorobiphenyl	xfer3.22	−3.17	0.1
2,2′,3,3′,4,4′,5,6′-octachlorobiphenyl	xfer3.23	−4.61	0.25
Decachlorobiphenyl	xfer3.24	−2.98	1
Dibenzo-p-dioxin	xfer3.25	−3.15	0.1
1-chlorodibenzo-p-dioxin	xfer3.26	−3.52	0.1
2-chlorodibenzo-p-dioxin	xfer3.27	−3.1	0.1
2,3-dichlorodibenzo-p-dioxin	xfer3.28	−3.56	1
2,7-dichlorodibenzo-p-dioxin	xfer3.29	−3.67	0.12
1,2,4-trichlorodibenzo-p-dioxin	xfer3.30	−4.05	0.1
1,2,3,4-tetrachlorodibenzo-p-dioxin	xfer3.31	−3.81	0.14
1,2,3,7-tetrachlorodibenzo-p-dioxin	xfer3.32	−3.84	1
2,3,7,8-tetrachlorodibenzo-p-dioxin	xfer3.33	−3.37	1
1,2,3,4,7-pentachlorodibenzo-p-dioxin	xfer3.34	−4.15	1
1,2,3,4,7,8-hexachlorodibenzo-p-dioxin	xfer3.35	−3.71	1
Octachlorodibenzo-p-dioxin	xfer3.36	−4.53	1

See the Supplemental material for further experimental detail and references

material will be small but if the true solubility is low the amount of suspended material can be greater than the amount in true solution. Two studies have shown that prolonged ultracentrifugation of "solutions" of DDT led to marked reductions in the measured concentration of DDT: from 5.3 to 1.2 ppb (84,150 G; 1.2 ppb is 3.4×10^{-9} M) [12] and from 12.5 to 1.7 ppb (39,100 G; 1.7 ppb is 4.8×10^{-9} M) [13].

Reproducibility of solvation measurements is no guarantee of freedom from systematic errors when the same experimental technique is used. For instance, two laboratories [14, 15] have studied chlorinated biphenyls using variations on the generator column technique, and reported results in reasonable agreement. Weil et al. [15] also measured the solubility of DDT, obtaining a value of 5.5 ppb, which is close to several literature reports [16, 17] but is considerably higher than what seems to be the true value of <1.7 ppb [12, 13].

There are two concerns about solubility measurements based on the generator column technique using compounds

deposited on a support by evaporation of a solution. (1) There is no guarantee that the material deposited on the inert packing of the column is crystalline; it is more likely to be glassy and thus a sub-cooled liquid. This means that the solubility is probably for the sub-cooled liquid rather than the solid, yet the vapor pressure will be measured for the crystalline solid. (2) There is no guarantee that the "solution" coming out of the generator column has all solute in solution and not as a suspension, whether of very small liquid droplets or very small crystallites. Aggregation of hydrophobic compounds can occur at very low concentrations [18, 19].

There is a basis for concern about reports of very low solubilities unless the issue of aggregation has been addressed. This has been shown to be a problem even at solubilities of 3×10^{-9} M for DDT, and is likely to be more severe at lower solubilities. This is distressing because even careful and reproducible work may be subject to systematic errors. The problem goes on because some direct measurements of Henry's law constants start with "saturated solutions" prepared using a generator column and may be subject to the same systematic error problem. For those polychlorinated biphenyls in the test set where direct determination of Henry's law constants has been reported, two papers were based on solutions prepared using generator columns [20, 21], one was based on solutions prepared by stirring with excess organic compound and then let stand [22], and one was based on solutions made with enough organic solute to give a solution less than half saturated based on reported solubilities [23]. This latter paper clearly comes closest to addressing the concerns over possible aggregation, but does not meet the strict criterion of demonstrating concentration independence of the Henry's law constants. However the two values of ΔG_s for compounds in the test set reported by Dunnivant et al. [23], at concentrations less than half "saturated" were 0.61 kcal/mol less negative than a value based on the generator column technique [20], and within 0.1 kcal/mol of a value [24] based on solutions well below the solubility limit and never exposed to excess compound. If the solubility is spuriously high, then the solvation free energy will be spuriously negative.

The full tables of data are provided in the supplemental data. Data are given for both indirect (solubility and vapor pressure) and direct determinations of ΔG_s. Where there is a concern that solubility reported or used may be spuriously high the letter Q (for question) follows the value. Where both cited and measured values are present only the measured values are used in the averaging. "Cited" values were quoted in the reference given, generally with a literature citation, but with no experimental data; these were often from earlier compilations. "Measured" values were from papers that gave experimental details.

It is quite possible that some of the experimental values are in error: this is a place where theory can help direct efforts to improve experiment. The likeliest source of systematic error in experimental solvation energies comes from the problems discussed above in determining very low solubilities. If computed solvation energies for a compound are less negative than experimental values, the computational method works well for more soluble analogs of the compound in question, and the experimental solubility for the compound in question is less than 10^{-7} M, then there is a strong basis for suspicion about the experimental values.

Metrics

As in previous SAMPL challenges, many common metrics were calculated for each submission, such as mean and median signed error, mean and median absolute error, and root mean square error (RMSE). While those metrics provide a measure of "absolute" performance at predicting experimental transfer energies, a Kendall's Tau rank correlation furnishes a measure of rank performance, and the gradient, intercept, and R-squared from a linear regression provide further information about the relative performance on the data set. Bootstrapping was utilized as a simple method to provide confidence intervals for these calculated metrics over the provided set of transfer energies. This involves repeated calculations of the metrics on trial sets determined by choosing random samples with replacement from the original set of predictions. Given the known and varying experimental error for experimental transfer energies in this set, smooth bootstrapping was employed by using the experimental error as the sigma value for Gaussian noise added to experimental values in each resample. This procedure was repeated until the standard deviation of all the RMSE values converged to a given tolerance of 0.0001, which resulted in convergence in less than 3,000 iterations in all cases. Then the median value of each metric over the set of samples, as well as 95 % confidence intervals derived from the calculated set of samples. To compare performance of one submission against another, the dependent sample *t*-test was used to compare unsigned errors from each pair of submissions.

Results and discussion

Challenges of the data

Several aspects of the transfer energy set presented different challenges to participants than previous SAMPLs, which focused on predicting transfer energies of a diverse set of 'drug-like' compounds. The focus on compounds of

increasing chlorination was a direct attempt to explore chemistry that piqued the interest of SAMPL2 participants. Within the SAMPL2 explanatory set, where the experimental value was revealed to participants, were hexachloroethane (present in SAMPL3) and hexachlorobenzene. Participants on average under-predicted the interaction between these compounds and water by about 2 kcals/mol, and most predictions were for positive transfer energies while the experimental values indicated the opposite. This result was particularly intriguing given the non-polar nature of these molecules. The SAMPL3 set was then designed to provide a strong challenge to the chlorine models of participating methods. There is great potential for additive or compounding errors to be revealed as the set includes in compounds with large numbers of chlorines. In addition to the range in overall number of chlorines, the data set contains several pairs of compounds with the same number of chlorines that differ only in the distribution of the chlorines around the scaffold, as shown in Fig. 1. Some of these pairs differ in transfer energy by up to 2 kcals/mol, resulting in a strong evaluation of the methods ability to reproduce and evaluate the dipoles and/or conformational differences between these structural isomers.

The narrow dynamic range of the set results in limited statistical power with which to differentiate methods performance from each other and from a null model. While the range of transfer energies were all within the range of values where methods were found to perform well in SAMPL2, the dynamic range of the entire data set was less than 7 kcal/mol, with a standard deviation of experimental values of 1.42 about the mean. This indicates that a null model that predicts the average transfer energy of the data set for all molecules would have an RMSE of 1.42 kcal/mol. No method was able to perform statistically significantly better than this null model. It should be recognized that such a null model requires knowledge of the experimental values to calculate the average value, and thus is not directly comparable to the blind predictions made by participants. However, noting the standard deviation of the experimental values in any prediction set provide a warning about the range of this set and the inherent ability of this set to distinguish models from each other and from null models.

Participation and overall performance

Participation in the transfer energy challenge of SAMPL3 was distributed differently among methods than in previous years, with an increase in the use of explicit solvent models with molecular dynamics simulations. Just over half of the 21 submissions (from 8 different participants) employed MD simulation, while the remaining 10 submissions split evenly between different implicit solvent models and one semi-explicit method (see Fig. 2).

As seen in previous SAMPLs, ascertaining that there is a statistically significant difference in the performance of two methods is challenging. In SAMPL3, the dependent *t*-test for paired sample was used to test the null hypothesis that the errors in two predictions are equivalent. The top-performing submission of each overall type of method (implicit, semi-explicit, and MD), as well as 7 of the top 8 performing methods, when ranked by RMSE, do not have a statistically significant difference in the distribution of their errors. Stated another way, there is not sufficient evidence that there are significant differences in the performance of these methods at the 95 % confidence level. Additionally, the 95 % confidence intervals of RMSE of these methods also overlapped with the 1.42 kcal/mol RMSE value that represents the performance of the null model described earlier. This provides an additional indicator of the overall

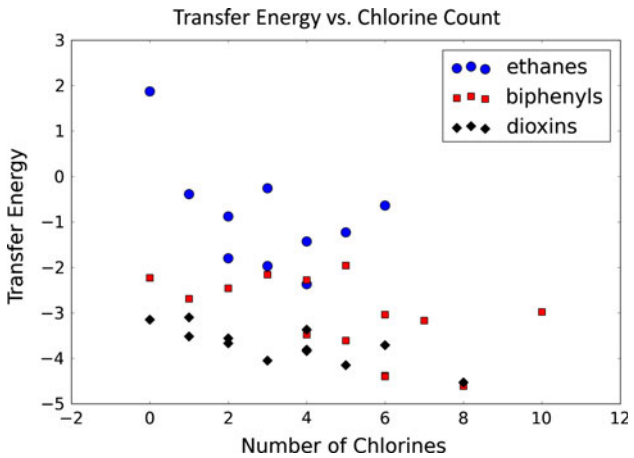

Fig. 1 The relationship between number of chlorine atoms and transfer energy for the three series in the data set

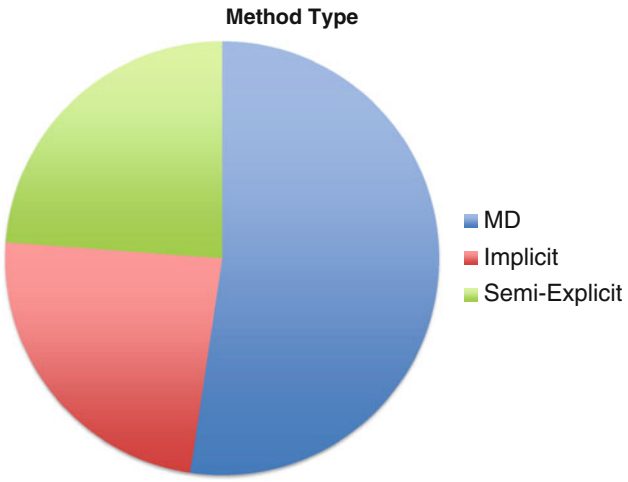

Fig. 2 The distribution of different solvent models among the prediction submissions

difficulty of this set, in that the best-performing submissions do not perform better than a simple, albeit retrospective, null model.

Performance across series

If we shift from comparing performance across different methods to examining the general performance on the three different series of molecule within the data set, it is apparent that all the methods performed much better on the ethane series than on either the dioxin or biphenyl series, as demonstrated by Fig. 3. This holds across many metrics such as RMSE, gradient of a best-fit line, and even a rank-correlation metric such as Kendall's Tau. Certainly one factor in this difference is the smaller dynamic ranges of the biphenyl and dioxin series, resulting in effectively a more difficult set, especially when measured by metrics such as Kendall's Tau or gradient of the best-fit line. However, further analysis elucidates that the 'narrowness' of these series is not the source of their difficulty.

Comparison of the mean absolute error (MAE) for each compound to the number of chlorines reveals a strong correlation between the average size of errors in the predictions and the chlorine count, as shown in Fig. 4. Importantly, it is clear that increases in difficulty are mostly independent of the actual transfer energy. That is, for compounds with less than five chlorines the average errors are quite similar despite differences in the transfer energies of up to 5 kcals/mol. No compounds with less than 5 chlorines had mean average errors of more than 2 kcal/mols, while 10 out of 13 molecules with 5 or more chlorines had MAEs above 2 kcal/mol.

Not only did the size of errors increase for compounds with higher numbers of chlorines, most methods also consistently under-predicted the affinity of heavily chlorinated for solvent. As seen clearly when examining the signed errors for the prediction of hexachloroethane shown

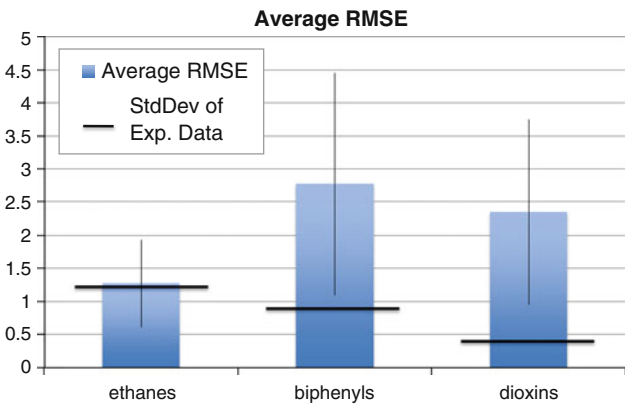

Fig. 3 The average RMSE of all predictions on the different molecule scaffolds that comprise the data set

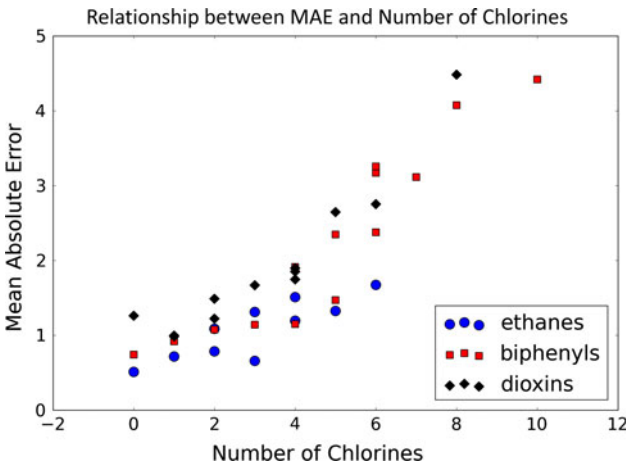

Fig. 4 The relationship between the MAE over all predictions of a given molecule and the number of chlorines in that molecule clearly indicate that difficulty in prediction increased with the number of chlorines

in Fig. 5, all but one prediction predicted a less favorable interaction with solvent than experiment, with an average of about 1.5 kcals/mol signed error for the field. Comparison to the same plot for biphenyl (Fig. 6) reveals a similar variance of errors but a mean much closer to zero. However plotting the signed errors for 2,3,3′,4,4′,5-hexachlorobiphenyl, which also has 6 chlorines like hexachloroethane, shows a both a positive bias in the distribution of signed errors as well as greater variation (Fig. 7). The comparison of mean signed error (MSE) to the chlorine count, seen in Fig. 8, reveals a very similar correlation to the plot of MAE (Fig. 4), which implies that almost all of the errors are due to predicting a less favorable interaction with water than measured by experiment. Similar to the trend seen with size of the errors, this relationship seems to be more strongly driven by the

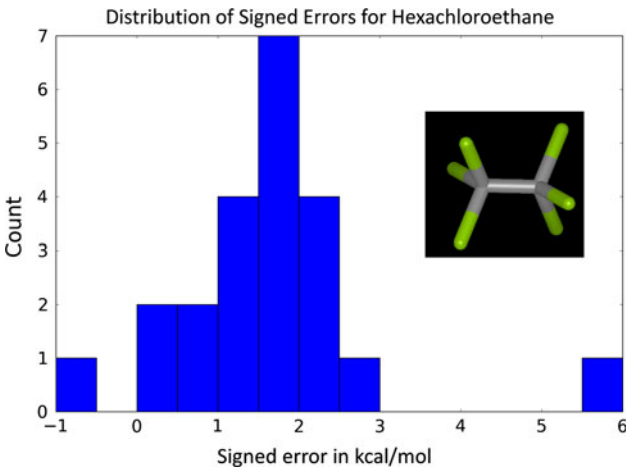

Fig. 5 The signed errors distribution for hexachloroethane indicates almost all predictions underestimated the favorability of the interaction between the molecule and water

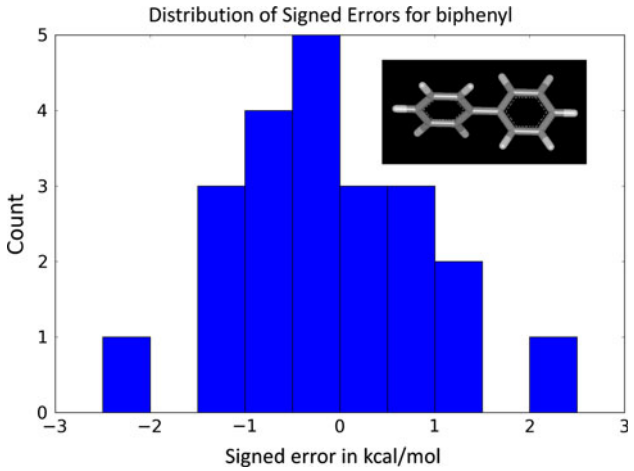

Fig. 6 The signed errors for predictions of biphenyl demonstrate a much more uniform distribution of predictions around the experimental value

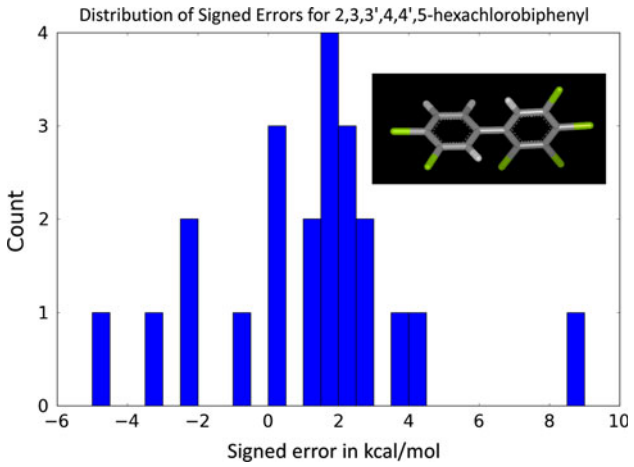

Fig. 7 The signed errors for predictions of 2,3,3′,4,4′,5-hexachlorobiphenyl has a shifted distribution of predictions like hexachloroethane, as well as an increased variance in predictions compared to biphenyl

number of chlorines rather than the transfer energy. This can be quickly deduced by examining the MSE for compounds with similar transfer energies but large differences in the number of chlorines, such as the mono- and per-chlorinated ethane compounds, which differ in transfer energy by only 0.25 kcals/mol yet the per-chlorinated compound's MSE increases by roughly 1 kcal/mol. Also important is the relationship between the variance in predictions and the number of chlorines. The same comparison between MSE and number of chlorines is shown in Fig. 9, but in this case the size of the point denotes the variance in predictions for each compound. As with the mean absolute error and mean signed error, the increases in variance correlate strongly with increases in number of chlorines, particularly for molecules with more than 5 chlorines.

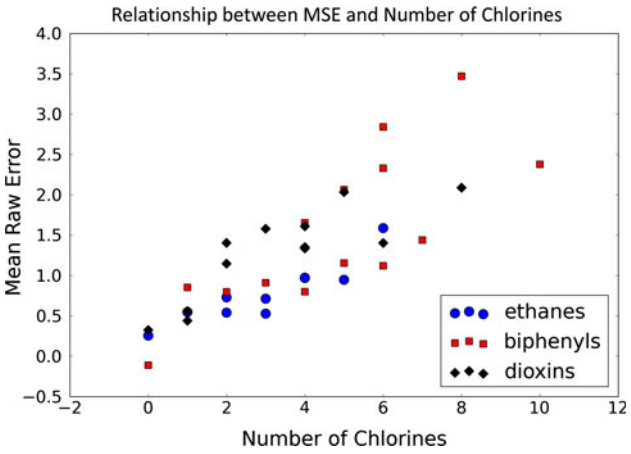

Fig. 8 The plot of MSE versus number of chlorines indicates that most of the errors in predictions were correlated in the same directions as well as increasing in magnitude with increasing number of chlorines

Conclusions

Implications for future predictions

Unlike previous SAMPL transfer energies, which attempted to present a set of "drug-like" molecules to assess the overall ability of prediction methods to address the types of small molecules important to pharmaceutical research, the transfer energy in SAMPL3 was designed to probe the capacity of methods to handle increasingly chlorinated molecules. This design was a response to intriguing results seen on a handful of chlorinated molecules in previous SAMPLs. The focused nature of this set indicates that the overall performance of a method in SAMPL3 should not be taken as an indicator of its ability to predict transfer energies of small molecules in general. Rather, the performance of the

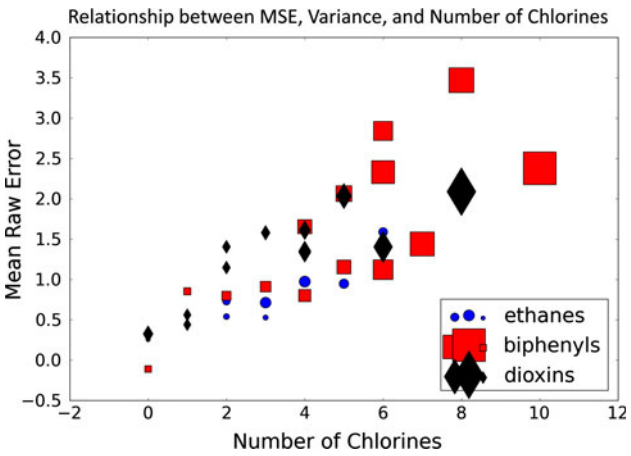

Fig. 9 The plot of MSE versus number of chlorines indicates that most of the errors in predictions were correlated in the same directions as well as increasing in magnitude with increasing number of chlorines

general field on the compounds in SAMPL3 indicate that there is an aspect of the interaction between chlorine and solvent that is erroneously predicted by the methods and accumulates as more chlorines are introduced. That is, one could argue that there is a modest systematic bias leading to positive errors for all chlorinated compounds, but a particular problem for compounds with more than 5 chlorines (which are also the ones most likely to be particularly difficult to measure) leading to more positive apparent errors.

If the experimental values are spuriously negative, then calculated values will be apparently too positive. One cannot say that the experimental values are wrong, but one can say that they might be wrong, and that it would be very helpful if some were re-determined with particular care to avoid possible problems. A case could be made for a positive error in calculated transfer energies linear in number of chlorines with about 1 kcal/mol scatter about the line, with a few (namely the hexa- and octa-chlorobiphenyls) well above the correlation line. Despite the fact that some of the errors in predictions for compounds may be due to deceptively negative transfer energy measurements (see Methods), the trend of increasing size of prediction errors, increasing under-estimation of the interaction with solvent, and increasing variance in predictions holds for molecules over the range of transfer energies, and is strongly correlated with the number of chlorines. While the overall data cannot distinguish if the error results from an over-estimation of a repulsive term, an under-estimation of an attractive term, or an entropy-driven issue, it presents a clear problem to modelers and method developers. One potential hypothesis could be under-estimation of polarizability, which is always an attractive effect and would be expected to increase in magnitude as number of chlorines increase as well as in the aromatic systems rather than the ethane series.

Experiments to challenge predictions

We as a field are rapidly approaching the exhaustion of blinded transfer energies for molecules that are relevant for pharmaceutical research. The last several SAMPL challenges have relied on extremely detailed and careful searching and curation of the literature to find previously measured values and often requires looking back several decades. Often the age of the publications and the measurements can then raise questions about the validity of the data and the methods. Moreover, the lack of new measurements hampers the ability to design data sets, find molecules to progressively test hypotheses, and iteratively challenge methods developers. The collection of quality experimental data is a difficult, time-consuming undertaking requiring meticulous care, yet it is on the back of such work that discoveries are made, methods are improved, and the field of molecular modeling is pushed forward. The time may soon arrive when the existing body of transfer energy data is no longer sufficient to continue to advance the field of solvation energy prediction. Any field such as molecular modeling, which is rooted in making predictions, is contingent on the ability to iteratively compare to and confront itself with new experimental data. It is therefore crucial to cultivate and preserve sources of experimental data that allow us in the modeling community to truly measure, challenge, and expand the performance our field.

References

1. Dominy BN (2008) Curr Pharm Biotechnol 9:87–95
2. Raha K, Merz K (2005) Annu Rep Comput Chem 1:113–130
3. Shirts MR, Mobley DL, Chodera JD (2007) Annu Rep Comput Chem 3:41–59
4. Moult J (2005) Curr Opin Struct Biol 15:285–289
5. Battey JND, Kop J, Bordoli L, Read RJ, Clarke ND, Schwede T (2007) Proteins 69(Suppl 8):68–82
6. Geballe MT, Skillman AG, Nicholls A, Guthrie JP, Taylor PJ (2010) The SAMPL2 blind prediction challenge: introduction and overview. J Comput Aided Mol Des 24(4):259–279
7. Guthrie JP (2009) A blind challenge for computational solvation free energies: introduction and overview. J Phys Chem B 113(14):4501–4507
8. Nicholls A, Mobley DL, Guthrie JP, Chodera JD, Bayly CI, Cooper MD, Pande VS (2008) Predicting small-molecule solvation free energies: an informal blind test for computational chemistry. J Med Chem 51(4):769–779
9. Bevington PR (1969) Data reduction and error analysis for the physical sciences. McGraw-Hill, New York, p 73
10. Bevington PR (1969) Data reduction and error analysis for the physical sciences. McGraw-Hill: New York (Chapter 4)
11. Guthrie JP (2011) In preparation
12. Bowman MC, Acree F, Corbett MK (1960) J Agric Food Chem 8:406–408
13. Biggar JW, Dutt GR, Riggs RL (1967) Bull Environ Contam Toxicol 2:90–100
14. Miller MM, Ghodbane S, Wasik SP, Tewari YB, Martire DE (1984) J Chem Eng Data 29:184–190
15. Weil L, Dure G, Quentin K (1974) Zeitschrift fuer Wasser und Abwasser Forschung 7:169–175
16. Webster GRB, Friesen KJ, Sarna LP, Muir DCG (1985) Chemosphere 14:609–622
17. Friesen KJ, Sarna LP, Webster GRB (1985) Chemosphere 14:1267–1274
18. Guthrie JP (1972) Chem. Commun 897–899
19. Guthrie JP (1973) Can J Chem 51:3494–3498
20. Brunner S, Hornung E, Santl H, Wolff E, Piringer OG (1990) Environ Sci Technol 24:1751–1754
21. Fang F, Chu S, Hong CS (2006) Anal Chem 78:5412–5418
22. Murphy TJ, Mullin MD, Meyer JA (1987) Environ Sci Technol 21:155–162
23. Dunnivant FM, Coates JT, Eizerman AW (1988) Environ Sci Technol 22:448–453
24. ten Hulscher TEM, van der Velde LE, Bruggeman WA (1992) Environ Toxicol Chem 11:1595–1603