J Comput Aided Mol Des (2014) 28:235–244
DOI 10.1007/s10822-014-9733-3

Prediction of hydration free energies for the SAMPL4 data set with the AMOEBA polarizable force field

Francesco Manzoni · Pär Söderhjelm

Received: 15 November 2013 / Accepted: 13 February 2014 / Published online: 1 March 2014

Abstract Hydration free energy calculations are often used to validate molecular simulation methodologies and molecular mechanics force fields. We use the free-energy perturbation method together with the AMOEBA polarizable force field and the Poltype parametrization protocol to predict the hydration free energies of 52 molecules as part of the SAMPL4 blind challenge. For comparison, similar calculations are performed using the non-polarizable General Amber force field. Against our expectations, the latter force field gives the better results compared to experiment. One possible explanation is the sensitivity of the AMOEBA results to the conformation used for parametrization.

keywords AMOEBA · Hydration free energies · GAFF · Polarizable force fields · SAMPL · Poltype

Introduction

Hydration free energies (HFEs) for small molecules are fairly easy to measure accurately and thus have been used extensively for parametrization and validation of computational methods. The number of compounds for which consistent measurements exist steadily increases. For example, in recent evaluations of solvation models [1–3] more than 500 compounds were used. HFEs are also important in their own respect, being one of the key components for determining binding affinities of ligands to proteins and for estimating the solubility of e.g. drug candidates. Furthermore, a thorough understanding of the water solvation process is beneficial in many other contexts as well.

Models for incorporating water effects in molecular modeling can be roughly classified into explicit and implicit models. When using explicit solvent, each water molecule is modeled explicitly and a substantial averaging of the water configurations is needed, commonly using free-energy perturbation (FEP) calculations. Implicit solvent models, on the other hand, treat water as a featureless continuum, typically characterized by the dielectric constant. Clearly, this offers a great speedup, as both the number of particles and the needed amount of sampling decrease. Explicit-solvent methods are in principle more accurate, because they include specific interactions (e.g. stable hydrogen bonds) and are able to model effects such as asymmetric solvation. However, they are limited by the quality of the force field, in particular the solute–water and water–water potentials.

Force fields can be roughly classified into pairwise additive and polarizable potentials. The pairwise additive models usually contain the polarization in an effective way, by using point charges that are enhanced to include the average polarization from the bulk. This limits their transferability between various systems. For example, when the GROMOS non-polarizable force field was parametrized to reproduce solvation free energies in cyclohexane and water, the authors had to publish two sets of charges for the same molecules, each set optimized for one of the two solvents [4]. In contrast, a polarizable force field handles this situation by explicitly including a mechanism for electronic polarization (e.g. polarizabilities, Drude models, or fluctuating charges), and thereby improves the transferability [5].

F. Manzoni · P. Söderhjelm (✉)
Department of Theoretical Chemistry, Chemical Center, Lund University, P.O. Box 124, 221 00 Lund, Sweden
e-mail: par.soderhjelm@teokem.lu.se

F. Manzoni
e-mail: francesco.manzoni@biochemistry.lu.se

The AMOEBA force field [6] is one of the most developed polarizability-based force fields for biomolecular systems, with ready-to-use multipole libraries for e.g. proteins. Recently, a rigorous AMOEBA parametrization protocol for new molecules was developed [7], which we refer to as *Poltype* after the name of the corresponding Python package. This tool, which handles multipole derivation, atom type recognition, and torsional parameter fitting in a consistent manner, offers a great possibility for testing and evaluating the AMOEBA force field for new systems. A similar protocol has been developed for other force fields [8].

Hydration free energies have been used several times to test and improve the AMOEBA force field. In fact, the first serious attempt to use a polarizable force field together with rigorous statistical-mechanical methods to compute HFEs was the AMOEBA submission in the SAMPL2 challenge [6]. With a root mean squared error of 10 kJ/mol, the result was among the best submissions using explicit solvent for this difficult test set, but actually worse than the best methods based on implicit water. Two problematic types of molecules were polyhalogenated compounds (due to underestimated polarizabilities of halogens) and nitro compounds. A further study optimized all aspects of the protocol for computing HFEs, including parametrization, and obtained significantly smaller errors (3 kJ/mol) but for simpler molecules [9]. Similar errors were obtained for a set of more complex compounds [7] extracted from a larger data set, for which corresponding results with the General Amber force field (GAFF) [10] were available [2]. For the compared molecules, the error was roughly a factor of 2 smaller with AMOEBA than with GAFF.

In this study, we test the AMOEBA force field and the Poltype parametrization by computing HFEs for the 52 molecules in the SAMPL4 data set, and comparing with GAFF results obtained with the same free-energy protocol. Quite surprisingly, the GAFF force field gives significantly better results. Possible reasons for the big discrepancy between the two models are discussed.

Methods

Choice of conformers

We studied the 52 molecules of the SAMPL4 data set, listed in Table 1. Some of the molecules can have many conformations. For parametrization, we selected one conformer of each molecule by the following procedure. First, a conformational search was performed using the Open Babel software [11] with an energy score, a genetic algorithm with 10 children, and generating maximally 40 conformers. Second, each conformer was geometry-optimized at the B3LYP/6-31G* level [12, 13] in implicit water modeled by IEPCM [14], as implemented in Gaussian 09 [15]. Finally, the conformer with the lowest energy was selected for parametrization and used as a starting structure in the simulations.

AMOEBA parametrization

The standard AMOEBA water model [16] was used as the solvent. For deriving AMOEBA parameters for the solute, the Poltype procedure [7] and the `poltype` Python package (version 1.1.2) was used without modifications, together with the Gaussian 09 [15], GDMA [17] and TINKER 5 software. The parametrization mainly involved the following automated steps:

- The symmetry of the molecule was analyzed to ensure identical parameters for symmetry-equivalent atoms.
- A distributed multipole analysis [18] was performed using the electron density determined at the MP2/6-311G** level.
- All multipoles except the charges were refined by fitting to the electrostatic potential determined at the MP2/6-311++G(2d,2p) level.
- Polarizabilities, van der Waals parameters and standard torsional parameters were assigned based on atom-type identification.
- For each rotable bond, a quantum-chemical scan was done, with energy calculations at the M06L/6-31G** level [19] performed at intervals of 30°. The torsional parameters were fitted to the quantum-chemical energies after the intramolecular electrostatic and polarization contributions were subtracted.

Only trivial manual interventions of this procedure were done, such as ensuring that C–Cl bonds were modeled as bonds. However, after submission of the results for the SAMPL4 blind challenge, we discovered that the parameters for the six anthraquinone derivatives were not enforcing planarity of the molecules strongly enough; the molecules were severly bent when simulated in water. This gave extremely negative HFEs and does not correspond to the real system. Due to the limited time, we were not able to do a thorough reparametrization of these molecules, but to avoid the problem we reran all the simulations with all torsional amplitudes for the ring system set to the ad hoc value of 24 kcal/mol (with minima at 0 and 180°), which we assume to be much higher than needed. No problem with distorsions was seen in these new calculations.

GAFF parametrization

For comparison, we also performed free-energy calculations with the General AMBER force field (GAFF) [10]

Table 1 Experimental and calculated hydration free energies (kJ/mol) obtained with the AMOEBA and GAFF force fields, respectively, and divided into electrostatic (polar) and van der Waals (non-polar) contributions

Mol	Name	Exp.	AMOEBA			GAFF			Corr.
		ΔG_{solv}	ΔG_{solv}	ΔG_{ele}	ΔG_{vdw}	ΔG_{solv}	ΔG_{ele}	ΔG_{vdw}	
1	Mannitol	−98.8	−112.2	−134.4	22.2	−77.0	−75.2	−1.8	−5.5
2	Linalyl acetate	−10.4	−2.9	−31.9	29.0	−11.0	−26.3	15.3	2.8
3	Nerol	−20.0	−6.6	−30.0	23.4	−10.1	−23.2	13.0	6.6
4	Geraniol	−18.6	−9.1	−33.0	23.9	−11.0	−23.4	12.4	5.0
5	1,2-Dimethoxybenzene	−22.3	−20.5	−34.9	14.4	−14.4	−21.5	7.2	−3.3
6	4-Propylguaiacol	−22.0	−34.1	−43.4	9.4	−16.0	−24.2	8.2	−4.1
7	4,5-Dichloroguaiacol	N/A	−24.9	−41.5	16.7	−18.6	−21.4	2.8	5.2
8	5-Chloroguaiacol	N/A	−25.6	−40.8	15.1	−18.1	−22.2	4.0	2.5
9	2,6-Dichlorosyringaldehyde	−34.5	−30.1	−53.0	22.9	−40.0	−42.2	2.2	13.8
10	3,5-Dichlorosyringol	−26.1	−29.2	−42.0	12.8	−23.8	−27.3	3.5	−3.2
11	2-Chlorosyringaldehyde	−32.6	−54.9	−75.9	21.0	−38.7	−40.7	2.1	−8.8
12	Dihydrocarvone	−15.7	−9.5	−30.1	20.7	−13.7	−22.0	8.3	−1.5
13	Carveol	−18.6	−14.2	−35.0	20.8	−14.2	−23.0	8.8	−1.1
14	l-Perillaldehyde	−17.1	−6.3	−26.4	20.1	−14.7	−23.0	8.3	−24.8
15	Piperitone	−18.9	−12.4	−30.7	18.3	−14.7	−23.3	8.6	−3.3
16	Menthol	−13.4	−14.0	−34.8	20.8	−13.3	−20.6	7.3	−2.4
17	Menthone	−10.6	−7.8	−27.7	19.9	−14.5	−20.8	6.3	3.6
18	1,2-Dihydroxy-9,10-AQ	N/A	−74.1	−86.4	12.3	−65.2	−60.7	−4.5	114.4
19	9,10-Dihydroanthracene	−15.8	−3.4	−18.5	15.1	−19.1	−23.1	4.0	−3.1
20	1,1-Diphenylethene	−11.6	2.9	−17.0	19.9	−15.4	−23.6	8.2	−2.3
21	1-Benzylimidazole	−31.9	−30.2	−44.3	14.1	−33.4	−37.9	4.5	−3.6
22	Mefenamic acid	−28.4	−68.6	−93.0	24.4	−32.2	−36.2	4.0	−18.8
23	Diphenhydramine	−39.1	−16.1	−41.9	25.8	−34.5	−45.0	10.5	−9.9
24	Amitriptyline	−31.1	−4.1	−29.0	24.9	−34.9	−40.8	5.9	−5.3
25	1-Butoxy-2-propanol	−24.0	−13.8	−34.4	20.6	−15.7	−25.1	9.4	5.3
26	2-Ethoxyethyl acetate	−22.2	−18.2	−35.8	17.6	−24.6	−33.4	8.8	−4.5
27	1,3-Bis-(nitrooxy)propane	−20.1	−13.6	−29.2	15.6	−23.8	−18.2	−5.6	−5.4
28	1,3-Bis-(nitrooxy)butane	−17.9	−10.5	−28.5	18.1	−20.6	−16.6	−4.0	1.2
29	Hexyl nitrate	−6.9	−1.0	−19.4	18.3	−6.3	−10.4	4.1	−3.7
30	Hexyl acetate	−9.6	−4.0	−25.2	21.2	−13.2	−23.3	10.1	−17.2
31	(e)-3-Hexenylacetate	N/A	−22.3	−42.4	20.1	−35.1	−44.3	9.2	1.7
32	3,4-Dichlorophenol	−30.5	−25.3	−38.6	13.3	−20.8	−25.7	4.9	−2.9
33	2,6-Dimethoxyphenol	−29.1	−29.2	−44.8	15.7	−27.3	−32.6	5.3	−4.3
34	4-Methyl-2-methoxyphenol	−24.3	−23.5	−39.0	15.5	−17.8	−23.8	6.1	−3.4
35	2-Hydroxybenzaldehyde	−19.6	−39.1	−50.8	11.7	−19.9	−23.9	3.9	−3.0
36	2-Ethylphenol	−23.7	−17.2	−31.8	14.6	−19.0	−26.8	7.8	−3.1
37	2-Methoxyphenol	−24.9	−25.3	−37.6	12.3	−18.6	−23.7	5.1	−3.1
38	2-Methylbenzaldehyde	−16.4	−15.5	−28.4	12.9	−20.2	−26.0	5.8	−2.5
39	1-Ethyl-2-methylbenzene	−3.6	4.9	−11.9	16.8	−3.2	−12.6	9.4	−2.2
40	1,2-Dioxane	N/A	−18.2	−28.5	10.4	−10.6	−13.6	3.0	−2.2
41	Piperidine	−21.1	−11.7	−22.6	10.9	−15.1	−18.7	3.6	−2.2
42	Tetrahydropyran	−13.1	−13.2	−24.4	11.2	−6.6	−11.1	4.5	−2.3
43	Cyclohexene	0.6	2.6	−7.0	9.6	4.9	−2.2	7.1	−1.6
44	1,4-Dioxane	−21.3	−30.8	−41.0	10.2	−17.9	−19.7	1.8	−2.7
45	2-Amino-9,10-AQ	−48.2	−38.2	−49.6	11.4	−57.9	−55.3	−2.6	71.8
46	1-Amino-9,10-AQ	−39.5	−22.0	−33.7	11.7	−46.6	−44.2	−2.4	59.2

Table 1 continued

Mol	Name	Exp.	AMOEBA			GAFF			Corr.
		ΔG_{solv}	ΔG_{solv}	ΔG_{ele}	ΔG_{vdw}	ΔG_{solv}	ΔG_{ele}	ΔG_{vdw}	
47	1-(2-Hydroxyethylamino)-9,10-AQ	−59.5	−41.0	−56.4	15.4	−56.7	−53.7	−2.9	58.3
48	1,4-Diamino-9,10-AQ	−49.6	−21.7	−36.6	14.8	−57.4	−52.4	−5.1	68.2
49	Dibenzo-p-dioxin	−13.2	−7.3	−18.6	11.3	−17.2	−17.7	0.4	−3.0
50	Anthracene	−17.3	−4.6	−17.2	12.6	−21.7	−25.4	3.7	−3.0
51	1-Amino-4-hydroxy-9,10-AQ	−39.9	−21.6	−34.2	12.6	−45.2	−40.5	−4.6	43.0
52	Diphenyl ether	−12.0	−2.4	−18.0	15.7	−11.2	−19.3	8.2	−2.9
MAE			9.8[a]			4.7			

The last column (corr.) shows the difference between the presented AMOEBA hydration free energy and the originally submitted value. AQ is short for anthraquinone. The mean absolute error (MAE) of each model is also shown

[a] The original submission had a MAE of 13.6 kJ/mol

together with the TIP3P water model [20]. This combination was also used to perform a preliminary equilibration before starting the AMOEBA equilibration. To derive parameters for the selected conformer, we used the `antechamber` program from the AmberTools 12 suite [21]. Atomic charges were determined by the AM1-BCC method [22], which has been shown to give HFEs of similar quality as more elaborate methods [23]. Moreover, the aim of this study was not to optimize the performance of the GAFF method.

Simulations

Each molecule was first inserted in a cubic box of water molecules extending at least 13 Å from the solute, using the GROMACS tools `editconf` and `genbox` [24]. The number of water molecules was 950–1,700 depending on the size of the solute. An initial constant pressure and temperature (NPT) equilibration was run with the GAFF/TIP3P force field until the volume and energy was stable (1–3 ns). The final configuration of this simulation was used as a starting point for the NPT equilibration with the AMOEBA force field, which was run for an additional 1 ns, after which the volume and energy were stable at their values for AMOEBA.

The AMOEBA simulations were run with the `pmemd.amoeba` module included in the Amber 12 software package [21]. All calculations were run in the NPT ensemble at 300 K, using Langevin dynamics with a damping coefficient of 2 ps^{-1} and isotropic position scaling with a pressure relaxation time of 1 ps. A time step of 1 fs was used and all bonds were modeled as flexible. Long-range electrostatics were treated using Particle Mesh Ewald (PME) summation with an ewald-direct-sum cutoff of 7 Å, fifth-order B-spline interpolation, and a PME grid size based on the Ewald coefficient 0.45 $Å^{-1}$. A cutoff for Thole damping interactions of 4.5 Å was used and the convergence threshold for the induced dipoles was set to 0.0001 atomic units, resulting in 12–15 iterations per timestep. The cutoff for van der Waals interactions was 12 Å and long-range van der Waals corrections were applied.

The GAFF simulations were run with the `pmemd` module and had mainly the same settings. However, for these simulations we constrained the bonds involving hydrogen atoms using the SHAKE algorithm [25] and used a longer timestep, 2 fs. A non-bonded cutoff of 12 Å was used for both electrostatics and van der Waals interactions and the automatic PME grid size was used.

The vacuum calculations for both force fields were run with the `sander` module. Non-periodic calculations with infinite cutoffs were used. The AMOEBA vacuum calculation used an Andersen temperature coupling scheme (due to technical problems with the Langevin dynamics) that randomized the velocities every 100 steps.

Free energy simulations

To compute the hydration free energy ΔG_{solv} we employed the thermodynamic cycle shown in Fig. 1, where M is the solute molecule. Thus, the calculations proceeded in three separate reaction steps:

- For determining ΔG_{ele}^{sol} the solute multipoles and polarizabilities were gradually turned on in a series of solute–water simulations, while all van der Waals interactions were modeled as normally.
- For determining ΔG_{vdw} the van der Waals interactions (repulsive and attractive) between the solute and solvent were gradually turned on in a series of solute–water simulations with an artificial solute without multipoles or polarizabilities. Note that the intramolecular van der Waals interactions were not affected by the scaling.

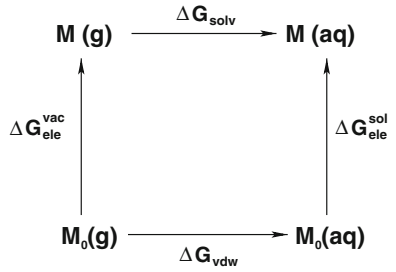

Fig. 1 Thermodynamic cycle used to calculate the hydration free energy ΔG_{solv} of molecule **M**. The *vertical* transformations amount to gradually turning on the solute multipoles and polarizabilities in vacuum and in water, respectively. The *lower horizontal* transformation amounts to gradually turning on the van der Waals interactions between the solute and solvent

- For determining ΔG_{ele}^{vac} the solute multipoles and polarizabilities were gradually turned on in a series of simulations of a single solute molecule in vacuum.

The free energy of each process was calculated using the multistate Bennett acceptance ratio (MBAR) method [26], as implemented in the `pymbar` Python package [26]. In each case, the gradual onset of the interactions was controlled by a scaling factor λ, varying from 0 (off) to 1 (on). For technical reasons, the electrostatic scaling with respect to a given λ was done slightly differently for GAFF and AMOEBA. For GAFF, a linear scaling of all electrostatic *energy terms* involving the solute was used. For AMOEBA, on the other hand, a linear scaling of all solute *multipoles and polarizabilities* was used. For the solute–solvent terms, these schemes are identical, because the interaction energy is linearly related to the values of the solute multipoles. However, for the solute intramolecular terms, the energy scales quadratically with the multipoles and thus the same λ gives different overall scaling in the two methods. This makes comparisons between the GAFF and AMOEBA charging profiles difficult. The polarizabilities cause further discrepancy between the two schemes. However, this difference is only relevant if the free energy is determined by thermodynamic integration; for MBAR the values of λ are not used in any case. For the van der Waals scaling, a softcore buffered 14-7 potential was used in AMOEBA, as described in Ref. [9]. A small modification of `pmemd.amoeba` was needed to avoid scaling the intramolecular van der Waals interactions. For GAFF, the default softcore Lennard-Jones potential with $\alpha = 0.5$ was employed and no modification was needed.

All simulations started from the same equilibrated structure. The same MD settings were used as for the equilibration. The simulation time of each window was 1.0 ns, from which the initial 0.2 ns was discarded. Snapshots were collected every 1 ps, and the 800 snapshots of each simulation were postprocessed to construct a matrix of potential-energy differences suitable for MBAR.

For all three transformations, we used 12 λ values, i.e. similar to what was done previously with AMOEBA [9], but we chose to distribute the λ's evenly between 0.01 and 0.99, as we did not experience any convergence problems (in fact, the results were converged after only 500 ps for most molecules). For one molecule (**25**), we tested to use 15 λ values for the van der Waals scaling, with a denser distribution in the $\lambda = 0.6$–0.8 range (as suggested in Ref. [9]), but the result differed by <0.4 kJ/mol. The reason for not including $\lambda = 0$ and $\lambda = 1$ is purely technical; it is not implemented in Amber 12. We added a correction term based on thermodynamic integration and the assumption that the *slope* of $\langle dH/d\lambda \rangle$ is constant in the intervals 0–0.01 and 0.99–1, respectively. For AMOEBA, we calculated $dH/d\lambda$ numerically from the MBAR matrix of potential-energy differences using a three-point quadratic interpolation formula; this approximation gave an error in the final ΔG_{solv} of <0.2 kJ/mol for all molecules when evaluated with GAFF, for which the exact $dH/d\lambda$ is available. The total contribution from the correction was always negative and in average 3.5 kJ/mol for AMOEBA and 1.0 kJ/mol for GAFF, with maximum values of 7 and 3 kJ/mol, respectively, for molecule **1**.

The total hydration free energy was finally computed as

$$\Delta G_{solv} = \left(\Delta G_{ele}^{sol} - \Delta G_{ele}^{vac}\right) + \Delta G_{vdw} = \Delta G_{ele} + \Delta G_{vdw} \quad (1)$$

whereby we define the net electrostatic contribution ΔG_{ele} to the hydration free energy. A combined uncertainty for the calculated ΔG_{solv} was estimated (on the safe side) as the sum of the three errors from the individual transformations.

Results

The AMOEBA and GAFF hydration free energies and their polar and non-polar contributions are presented in Table 1. The correlation between the calculated and experimental values are shown in Fig. 2. The average statistical error is 0.6 kJ/mol for AMOEBA and 0.5 kJ/mol for GAFF, with a maximum error of 1.0 kJ/mol in both cases (observed for molecule **23** with both force fields and additionally **1, 22**, and **24** with AMOEBA). Thus, all results are sufficiently converged for the differences discussed below to be significant.

We submitted our AMOEBA results to the SAMPL4 blind challenge [27] without knowledge of the experimental HFEs. After the submission, we discovered several errors in our original calculations, which we have corrected. The changes from the originally submitted values

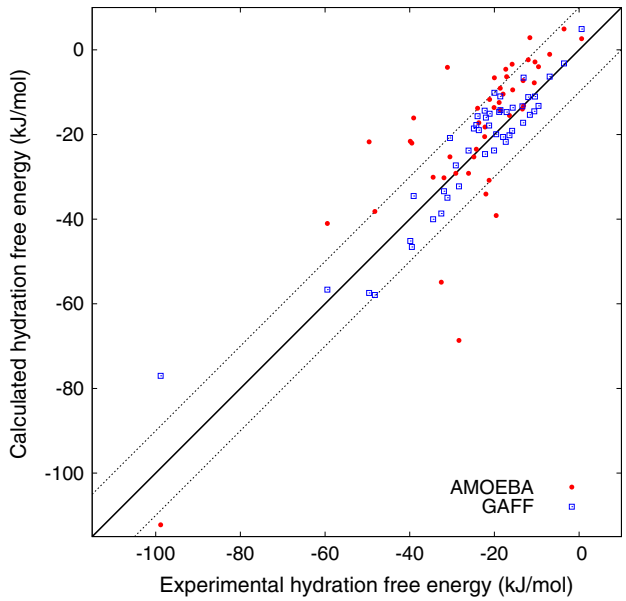

Fig. 2 Correlation between the experimental HFEs and the HFEs calculated using the AMOEBA and GAFF force fields, respectively. Only 47 of the 52 molecules are shown, as the remaining 5 were retracted from the SAMPL4 data set. The *solid line* indicates perfect agreement and the *dashed lines* indicate a 10 kJ/mol deviation

are listed in the column *Corr* in Table 1. The corrections were of four types: First, the parameters for the anthraquinones (**18, 45, 46, 47**, **48**, and **51**) obtained through Poltype gave a non-planar geometry in solution and completely meaningless results (see the Methods section). Second, ΔG_{ele}^{vac} were not calculated with infinite electrostatic cutoff. This affected about half of the molecules (the big ones), but only by <3 kJ/mol for all molecules except **14** and **27**, for which the error was 19 and 25 kJ/mol, respectively. In the original submission, we tried to solve the cutoff problem (being unaware that it was a cutoff problem) by rerunning the vacuum calculations for molecules **14** and **25–31** with a restraint that locked the molecule to a single conformation. This restraint, though in most cases acting in the right direction, sometimes caused a larger error than the cutoff problem itself. Third, for some molecules (**2–4**, **7–9**, **11**, **14**, **20**, **22**, and **23**), we noticed that we had not used the conformation with the lowest energy for the Poltype parametrization, so we repeated all calculations for these molecules to investigate whether this affected the results (see below). Finally, we had not included the extrapolation to $\lambda = 0$ and $\lambda = 1$ in the original submission.

The mean absolute error (MAE) for AMOEBA is 9.8 kJ/mol (the originally submitted values had a MAE of 13.6 kJ/mol). The MAE for GAFF is only 4.7 kJ/mol, which if submitted would have been among the 10 best submissions to the SAMPL4 challenge [27]. It is unexpected that a fixed-charge force field gives a substantially better prediction than a polarizable force field, in which each term is treated in a more physically correct way. Therefore, it is interesting to investigate the differences between the two force fields.

First, we note that neither of the force fields predicts sufficiently negative HFEs in average, but this systematic error is significantly larger for AMOEBA (4.6 kJ/mol) than for GAFF (1.0 kJ/mol). If we allow for a constant shift of the energies (i.e. by subtracting these average errors), the MAE for AMOEBA is reduced from 9.8 to 8.1 kJ/mol whereas the MAE for GAFF is unchanged.

The correlation between the GAFF and AMOEBA results are shown in Fig. 3 both for the total HFEs and for the ΔG_{ele} and ΔG_{vdw} contributions. Despite having entirely different electrostatic treatments in the two models, there is a reasonable correlation ($R^2 = 0.56$) between the electrostatic energies. Some molecules have a significantly more negative ΔG_{ele} when modeled by AMOEBA, in particular molecules **1**, **22**, and **11** with differences of 59, 57, and 35 kJ/mol, respectively.

Interestingly, there is much weaker correlation ($R^2 = 0.22$) between the non-polar contributions, despite that the van der Waals term is modeled by similar expressions in the two force fields. The van der Waals contribution is always more positive when modeled by the AMOEBA force field, in average by 12.0 kJ/mol. This is partly counterbalanced by a more negative electrostatic contribution (in average 9.1 kJ/mol), but the poor agreement with experiment suggests that this balance between two large terms might be more sensitive to the parametrization than the smaller-magnitude balance given by GAFF.

For a qualitative analysis, the non-polar contribution can be divided into two terms,

$$\Delta G_{vdw} = \langle E_{vdw} \rangle + \Delta G_{cav} \quad (2)$$

where $\langle E_{vdw} \rangle$ is the average van der Waals interaction energy between the solute and solvent, evaluated over the ensemble generated with the electrostatics of the solute turned off, and ΔG_{cav} is the remaining energy, loosely corresponding to the energy cost of creating a solute-sized cavity in the particular model of water. Fig. 4 shows the strong correlation between the GAFF and AMOEBA estimates of both these terms ($R^2 = 0.93$ and 0.97, respectively); thus the reason for the weak correlation in the total ΔG_{vdw} rather lies in the lack of cancellation between the terms than in any of the individual terms. However, there is a systematic shift in ΔG_{cav}, the AMOEBA values being 11 % larger in average. This effect could possibly be related to the general overestimation of the HFEs by AMOEBA, but further study would be needed to confirm this hypothesis. Finally, we note that the difference

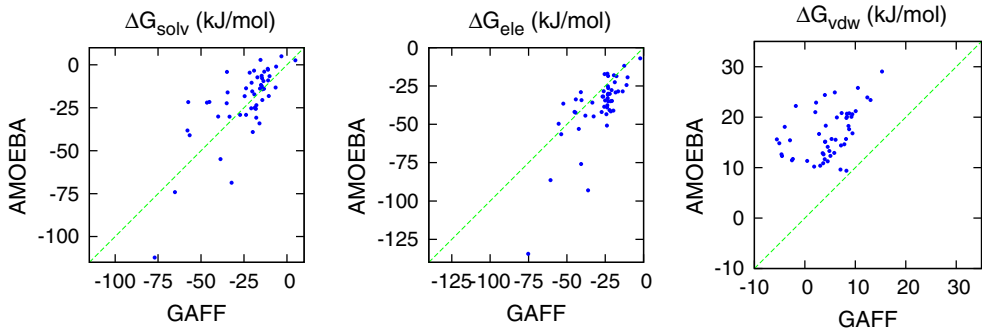

Fig. 3 Comparison between the HFEs calculated with the AMOEBA and GAFF force fields, respectively (*left*) and analogous comparisons for the electrostatic and van der Waals contributions to the HFE

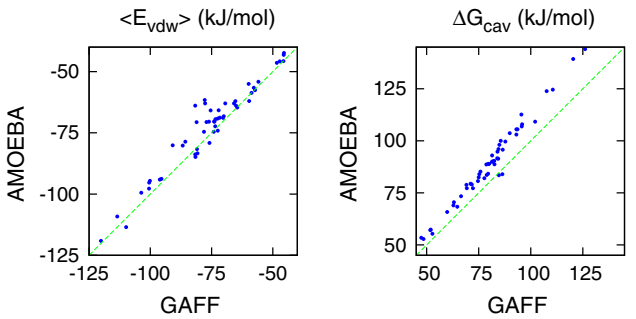

Fig. 4 Comparison between the GAFF and AMOEBA estimates of $\langle E_{vdw} \rangle$ and ΔG_{cav}, as defined in Eq. 2

between evaluating $\langle E_{vdw} \rangle$ with the electrostatics of the solute turned on or off is in average only 8 kJ/mol for GAFF but 43 kJ/mol for AMOEBA. Again, this indicates that AMOEBA relies on a more challenging balance between strongly negative electrostatics and a significant repulsive contribution, especially for hydrogen-bonded interactions.

A common complication in HFE calculations is the presence of multiple conformations, which can influence the results in at least three different ways. First, if there are significant free-energy barriers between the conformational minima, there is a risk that the calculation gets stuck in a single conformation and does not sample all the relevant conformations. Second, as the parameters (e.g. multipoles) are determined for a single conformation, they might not be suitable for other sampled conformations; this problem has been the subject of much research in the context of electrostatic potential-fitted charges [28–32]. Third, if the molecule switches conformation between the gas phase and the solvated state, the HFE will include the energy cost of this transformation, and this type of intramolecular energies might not be well described by the employed force field.

With the possible exception of molecule **1**, we see no clear signs of the barrier problem for this data set; moreover, we obtain good results with GAFF using the same FEP protocol. To investigate the problem of force field parameters, we analyzed the changes in ΔG_{ele}, $\langle E_{vdw} \rangle$, and ΔG_{cav} when the conformation used for parametrization by Poltype was altered (but both conformations being locally geometry-optimized). The results are shown in Table 2. Interestingly, for many molecules, the largest changes occur in ΔG_{cav}, despite the fact that the only geometry-dependent parameters are the multipoles (which are completely turned off during the calculation of $\langle E_{vdw} \rangle$ and ΔG_{cav}) and the bonded parameters, especially the torsions. Visual inspection of the trajectories for molecule **3** showed that there is indeed a change of preferred geometry between the two sets of parameters: with the original parameters, the C_2–C_3–C_4–C_5 dihedral angle stays around 90°, whereas for the best-conformation parameters, the same angle stays around −80°. Probably, this change makes the molecule effectively bigger, so that the cost of inserting it into water, ΔG_{cav}, increases by as much as

Table 2 Changes in AMOEBA hydration free energies (kJ/mol) upon altering the molecular conformation used in the parametrization

Mol	$\Delta\Delta G_{ele}$	$\Delta\langle E_{vdw} \rangle$	$\Delta\Delta G_{cav}$	$\Delta\Delta G_{solv}$
2	−1.8	0.0	9.0	7.2
3	3.1	−1.8	10.6	11.9
4	0.9	−3.6	12.5	9.8
7	−1.4	−1.6	11.5	8.5
8	−1.4	−1.1	8.2	5.7
9	5.1	−1.6	14.7	18.2
11	−7.0	0.7	3.1	−3.1
14	−6.0	−0.1	0.1	−6.0
20	1.1	−0.2	0.8	1.7
22	−18.5	−0.9	2.7	−16.7
23	−3.1	−2.9	1.2	−4.7

Besides the change in the total free energy ($\Delta\Delta G_{solv}$), three contributions are shown: the change in electrostatic free energy, the change in average van der Waals energy, and the change in cavity formation energy

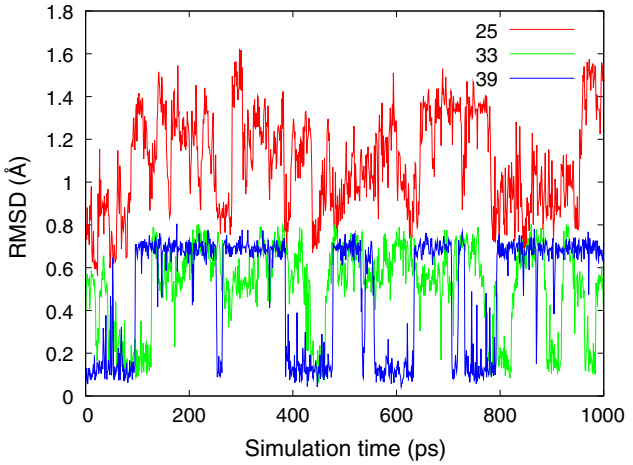

Fig. 5 Root mean square deviation (RMSD) of the non-hydrogen atoms of the solute relative to its starting conformation during the AMOEBA equilibration (in water) of molecules **25**, **33**, and **39**

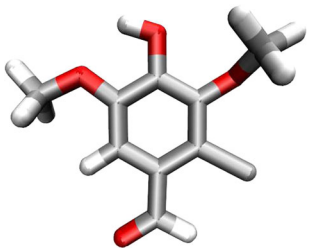

Fig. 6 Structure of 2-chlorosyringaldehyde (molecule **11**) used to examplify the importance of intramolecular hydrogen bonds

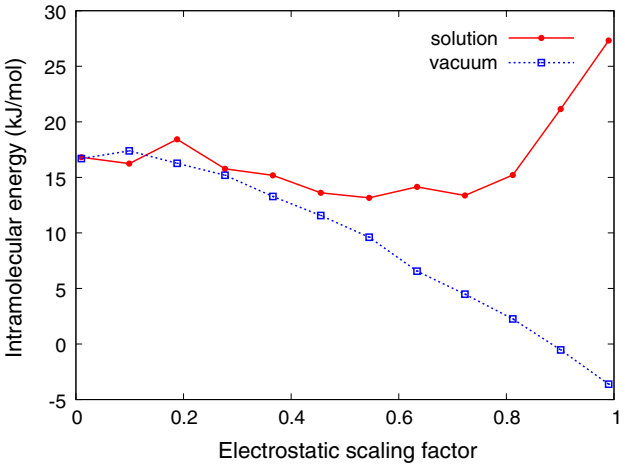

Fig. 7 Average intramolecular energy of molecule **11** over ensembles obtained from solvated and vacuum simulations with varying scaling factor λ for the multipoles and polarizabilities

11 kJ/mol. Thus, we conclude that the conformation used to determine the AMOEBA parameters can have a substantial effect on the HFEs; we are currently investigating this issue further.

The importance of accurately modeling the conformational energies can be tested by investigating the extent that different conformations are visited during the gas-phase and solvated trajectories. We first performed standard RMSD analyses of the trajectories, with some examples shown in Fig. 5. This showed which molecules remained in a single conformation and which visited a few well-defined conformations (e.g. **33** and **39**), and made it possible to confirm that the production simulation time of 0.8 ns was sufficient for these molecules. However, many molecules visit a large number of conformations that deviate significantly from the starting conformation, as exemplified by molecule **25**, but with a negligible effect on either the inter- or intramolecular energy. For example, they may represent symmetry-equivalent conformations or changes in non-polar parts of the molecule with no specific interactions with water.

To address the conformational sampling from an energetic perspective, we computed the intramolecular energy of the solute along the gas-phase and solvated trajectories. Although most molecules showed negligible variation along a single trajectory, there were several molecules that had significantly different average intramolecular energy in the two ensembles (with the solvated ensemble always giving the highest energy). In particular, the following molecules have larger differences than 5 kJ/mol: **1** (18 kJ/mol), **6** (6 kJ/mol), **10** (7 kJ/mol), **11** (31 kJ/mol), **18** (28 kJ/mol), **22** (11 kJ/mol), **28** (7 kJ/mol), **31** (11 kJ/mol), **35** (50 kJ/mol), and **47** (27 kJ/mol). We particularly analyzed molecule **11** (cf. Fig. 6), which has the second-largest difference and whose hydroxyl group has the possibility to form a hydrogen bond with either of the neighboring ether oxygens. The closest distance between the hydrogen and any of these oxygen is in average 2.68 Å in the solvated trajectory and 2.13 Å in the gas-phase trajectory, indicating that the discrepancy in the intramolecular energy is caused by the breakage of the intramolecular hydrogen bond when there are water molecules that the hydroxyl group can bind to instead.

The different preference can be studied by calculating the average intramolecular energy as a function of the electrostatic scaling factor (λ) used to generate the ensemble (but using the full multipoles when evaluating the energy). As can be seen in Fig. 7, the average intramolecular energy from the vacuum trajectory decreases monotonically upon turning on the charges. This corresponds to a gradually stronger alignment of the polar parts of the molecule, in particular the hydrogen bond. In contrast, the energy from the solvated trajectory is rather constant up to $\lambda = 0.8$ due to a fair competition against

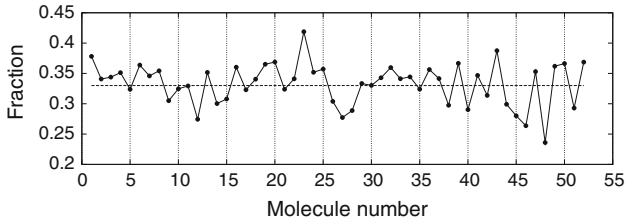

Fig. 8 Fraction of the total electrostatic interaction energy that comes from polarizabilities, for each of the 52 molecules in the data set. The *line* represents the average fraction (0.33)

alignment to the surrounding water molecules; after that the water "wins" and the large discrepancy of 31 kJ/mol is attained. The same explanation can be applied to molecules **35**, **18**, and **47**, as these also show large energy differences and all have the possibility to form internal hydrogen bonds. Clearly, the energy cost of breaking this hydrogen bond has a large and direct effect on the hydration free energy, and thus a more thorough parametrization of the intramolecular terms towards quantum-mechanical calculations should be beneficial to the modeling of such molecules.

Another interesting question is how large a part of the electrostatic hydration free energy comes from the polarization. In our FEP calculations, we did not separate the contributions from multipoles and polarization, but earlier tests have shown that doing the analysis for the average interaction energy between the molecule and the surrounding water tends to give similar results. The ratio between the polarization energy and the full electrostatic energy (including the polarization) for each molecule is shown in Fig. 8. Interestingly, this ratio is remarkably stable, fluctuating around the average value 0.33 with a standard deviation of 0.034. The highest ratios (0.42 and 0.39) are observed for molecules **23** and **43** and the lowest ratios (0.24 and 0.26) are observed for **48** and **46**.

Conclusions

We have calculated statistically well-converged hydration free energies with the AMOEBA and GAFF force fields using identical free-energy-perturbation methodologies and compared the results to the experimental values in a blind manner. The GAFF force field with AM1–BCC charges gave surprisingly good results (e.g. a MAE of 4.7 kJ/mol) considering its simplicity. In contrast, the AMOEBA polarizable force field parametrized using the Poltype procedure, gave disappointing results with a MAE of 9.8 kJ/mol. Partly this can be explained by a general underestimation of the free energies (i.e. not predicting sufficiently negative values), but a MAE of 8.1 kJ/mol remains even after subtracting this systematic shift.

A termwise comparison showed that there are substantial differences in the non-polar contribution predicted by the two force fields, in addition to the anticipated differences in the polar contribution. Moreover, both terms were found to be sensitive to the force field parameters: Simply performing the parametrization for another conformation of the molecule changed the cavity formation term by up to 15 kJ/mol. Even larger changes were observed in the electrostatic contribution for one of the reparametrized molecules.

Several of the molecules occur in different conformations in the gas phase and solvated phase, with differences in intramolecular energy as large as 50 kJ/mol. This presents a great challenge for force fields used to estimate hydration free energies, as they should not only model the intermolecular interactions accurately, but also the intramolecular energy. One solution would be to complement the force-field calculations with quantum-mechanical evaluations of these special cases. However, we see no reason why the much simpler GAFF force field would treat these intramolecular contributions in a better way. We are currently investigating all these issues in greater detail.

Acknowledgments This investigation has been supported by a grant from the Swedish research council (agreement C0020401). The computations were performed on computer resources provided by the Swedish National Infrastructure for Computing (SNIC) at Lunarc at Lund University and HPC2N at Umeå University.

References

1. Mobley DL, Dill KA, Chodera JD (2008) Treating entropy and conformational changes in implicit solvent simulations of small molecules. J Phys Chem B 112:938–946
2. Mobley DL, Bayly CI, Cooper MD, Shirts MR, Dill KA (2009) Small molecule hydration free energies in explicit solvent: an extensive test of fixed-charge atomistic simulations. J Chem Theory Comput 5(2):350–358
3. Sulea T, Corbeil CR, Purisima EO (2010) Rapid prediction of solvation free energy. 1. An extensive test of linear interaction energy (LIE). J Chem Theory Comput 6(5):1608–1621
4. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the gromos force field parameter sets 53A5 and 53A6. J Comput Chem 25(13):1656–1676
5. Cieplak P, Dupradeau FY, Duan Y, Wang JM (2009) Polarization effects in molecular mechanical force fields. J Phys Condens Matter 21(33):333102
6. Ponder JW, Wu C, Ren P, Pande VS, Chodera JD, Schnieders MJ, Haque I, Mobley DL, Lambrecht DS, DiStasio RA Jr (2010) Current status of the AMOEBA polarizable force field. J Phys Chem B 114(8):2549–2564
7. Wu JC, Chattree G, Ren PY (2012) Automation of AMOEBA polarizable force field parameterization for small molecules. Theor Chem Acc 131(3):1138

8. Huang L, Roux B (2013) Automated force field parameterization for non-polarizable and polarizable atomic models based on target data. J Chem Theory Comput 9(8):3543–3556
9. Shi Y, Wu CJ, Ponder JW, Ren PY (2011) Multipole electrostatics in hydration free energy calculations. J Comput Chem 32(5):967–977
10. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general AMBER force field. J Comput Chem 25(9):1157–1174
11. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR (2011) Open babel: an open chemical toolbox. J Cheminform 3:33
12. Becke AD (1993) J Chem Phys 98:1372
13. Hariharan PC, Pople JA (1973) Theor Chem Acc 28:213
14. Tomasi J, Mennucci B, Cammi R (2005) Quantum mechanical continuum solvation models. Chem Rev 105(8):2999–3093
15. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery Jr JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09 Revision A.1. Gaussian Inc. Wallingford
16. Ren P, Ponder JW (2003) Polarizable atomic multipole water model for molecular mechanics simulation. J Phys Chem B 107:5933
17. Stone AJ, Alderton M (2005) Distributed multipole analysis: stability for large basis sets. J Chem Theory Comput 1:1128
18. Stone AJ, Alderton M (1985) Distributed multipole analysis—methods and applications. Mol Phys 56:1047
19. Zhao Y, Truhlar DG (2006) A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. J Chem Phys 125:194101
20. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79:926
21. Case D, Darden T, TE Cheatham I, Simmerling C, Wang J, Duke R, Luo R, Walker R, Zhang W, Merz K, Roberts B, Hayik S, Roitberg A, Seabra G, Swails J, Goetz A, Kolossvary I, Wong K, Paesani F, Vanicek J, Wolf R, Liu J, Wu X, Brozell S, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh MJ, Cui G, Roe D, Mathews D, Seetin M, Salomon-Ferrer R, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman P (2012) Amber 12. University of California, San Francisco
22. Jakalian A, Bush BL, Jack DB, Bayly CI (2000) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. J Comput Chem 21(2):132–146
23. Mobley DL, Dumont É, Chodera JD, Dill KA (2007) Comparison of charge models for fixed-charge force fields: small-molecule hydration free energies in explicit solvent. J Phys Chem B 111(9):2242–2254
24. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4(3):435–447
25. Ryckaert JP, Ciccotti G, Berendsen HJC (2008) J Comput Phys 23:327
26. Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. J Chem Phys 129:124105
27. Mobley DL, Wymer K, Lim NM (2014) Blind prediction of solvation free energies from the SAMPL4 challenge. doi:10.1007/s10822-014-9718-2
28. Reynolds CA, Essex JW, Richards WG (1992) Atomic charges for variable molecular conformations. J Am Chem Soc 114:9075
29. Stouch TR, Williams DE (1992) Conformational dependence of electrostatic potential derived charges of a lipid headgroup: glycerylphosphorylcholine. J Comput Chem 13:622–632
30. Stouch TR, Williams DE (1993) Conformational dependence of electrostatic potential-derived charges: studies of the fitting procedure. J Comput Chem 14:858–866
31. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J Phys Chem 97:10269
32. Söderhjelm P, Ryde U (2009) Conformational dependence of charges in protein simulations. J Comput Chem 30:750