J Comput Aided Mol Des (2011) 25:1085–1093
DOI 10.1007/s10822-011-9486-1

MM/GBSA and LIE estimates of host–guest affinities: dependence on charges and solvation model

Samuel Genheden

Received: 19 July 2011 / Accepted: 7 November 2011 / Published online: 19 November 2011

Abstract The affinities of two sets of guest–host systems were estimated using the popular end-point methods MM/GBSA (molecular-mechanics with generalised Born and surface-area solvation) and LIE (linear interaction energy). A set of six primary alcohols that bind to α-cyclodextrin (α-CD) and a set of eight guest molecules to cucurbit[8]uril (CB8) were considered. Three different charge schemes were used to obtain charges for the host and guest molecules, viz., AM1-BCC, RESP, and the recently suggested xAvESP (which average ESP charges over a number of molecular dynamics snapshots). Furthermore, both the generalised Born and Poisson–Boltzmann solvation models were used in the MM/GBSA calculations. The two solvation models perform equally well in predicting relative affinities, and hence there is no point in using the more expensive Poisson–Boltzmann model for these systems. Both the LIE and MM/GBSA estimates are shown to be robust with respect to the charge model, and therefore it is recommended to use the cheapest AM1-BCC charges. Using AM1-BCC charges, the MM/GBSA method gave a MADtr (mean absolute deviation after removal of systematic error) of 17 kJ/mol and a correlation coefficient (r^2) of 0.67 for the CB8 complexes, and a MADtr of 10 kJ/mol and an r^2 of 0.96 for the α-CD complexes. The LIE method gave a MADtr of 20 kJ/mol and an r^2 of 0.10 for the CB8 complexes, after optimisation of the non-polar scaling parameter. For the α-CD complexes, no optimisation was necessary and the method gave a MADtr of 2 kJ/mol and a r^2 of 0.96. These results indicate that both MM/GBSA and LIE are able to estimate host–guest affinities accurately.

Keywords MM/GBSA · MM/PBSA · LIE · MD simulations · α-cyclodextrin · Cucurbit[8]uril

Introduction

An important goal in computational chemistry is to develop an accurate method to estimate the free energy difference of a chemical process [1, 2]. Such a method would be useful in a wide range of subjects such as enzymatic catalysis, folding and molecular recognition. The latter is particular interesting as an aid in drug development and for the understanding of for instance biochemical processes [3, 4].

Pertinent to the development of a computational method to predict free energy differences is a well-defined test system. Host–guest systems are appropriate model systems to study molecular recognition. Instead of testing the methodology on a large biomolecule such as a protein, one studies a much smaller system. Thereby, the phase space to traverse is more tractable, and one can hope to avoid large sampling problems, i.e., less noise in the results. At the same time, one expects that the binding reaction is dictated by the same intermolecular forces as in a protein–ligand complex although the binding affinity of the host–guest complex might be weaker [5]. There has been an increased interest in host–guest systems and predictions of host–guest systems were included in the SAMPL3 challenge organized by OpenEye Scientific Software in 2011 [6].

Methods to predict binding free energies ranges from simple scoring functions to formally exact alchemical free

Electronic supplementary material The online version of this article (doi:10.1007/s10822-011-9486-1) contains supplementary material, which is available to authorized users.

S. Genheden (✉)
Division of Theoretical Chemistry, Department of Chemistry, Lund University, P. O. Box 124, 221 00 Lund, Sweden
e-mail: Samuel.Genheden@teokem.lu.se

energy calculations [2]. Simulation-based methods have a clear theoretical and physical basis and are therefore the focus of this paper. Two popular end-point methods, viz., MM/GBSA [7, 8] (molecular mechanics with generalised Born and surface-area solvation) and LIE [9, 10] (linear interaction energy) were chosen for a detailed evaluation. These methods simulate only the end-points of the reaction (the reactants and the product) and are therefore relatively inexpensive, contrary to alchemical methods, which rely on the simulation of unphysical, intermediate states. It is worth mentioning that end-point methods are not exact by any means and there is only a hope that they will reproduce experimental results. Therefore, it is very important to study these methods carefully to identify their strengths and weaknesses.

The basis of the simulations used in end-point and alchemical methods is an empirical potential, a so-called force field. Empirical force fields have been extensively and systematically developed for various biomolecules, e.g. proteins and DNA [11–14]. The development of an analogous force field for smaller molecules, e.g. drugs and co-factors, have been less systematic and more ad-hoc. However, the last decade have seen the birth of general-purpose force fields for small molecules that should be compatible with the macromolecular ones [15, 16]. These are parametrised on small organic molecules and it is of interest to investigate how well they work for the larger host molecules and for the interactions between a host and a guest.

Herein, a test of the ability of the MM/GBSA and LIE methods to reproduce experimental binding affinities of two sets of host–guest systems is described. Because the electrostatics part is arguably the most important part of a force field, the effect of the charge model will be investigated. Three different charge models have been chosen, viz., the RESP (restrained electrostatic potential) procedure [17], which is the standard procedure to obtain charges in the Amber force field, the AM1-BCC (AM1 with bond charge corrections) charges [18] which has been parameterised on small organic molecules to reproduce RESP charges, and the eXAvESP [19] (extensively averaged electrostatic potential) procedure that has shown to be promising in studies of protein–ligand affinities [20]. With the MM/GBSA method, the ability of the generalised Born and Poisson–Boltzmann methods to calculate polar solvation free-energy will be compared. In passing, it should be mentioned that similar studies have been published previously, but only for protein–ligand systems [see for instance [21–24]]. However, rigorous studies of these issues for host–guest systems have not been performed.

As a test set the cucurbit[8]uril host with eight different guests and α-cyclodextrin with five alcohols have been chosen. These system have been experimentally studied by x-ray crystallography [25–27] and affinity measurements, [25, 28, 29] as well as by theoretical methods [30–32]. Therefore, they are suitable for this benchmark study.

Methods

Preparation of molecular systems

Cucurbit[8]uril (CB8) with the guest molecules shown in Fig. 1 and α-cyclodextrin (α-CD) with the first five primary alcohols have been studied. The CB8 systems were based on the crystal structure of CB8-L3 and CB8-L11. The other CB8 guests were placed in the host to mimic the binding mode of L3 and L11, or as suggested in [25]. The protonation state of the guests were based upon a pH 7 and confirmed with tools from ChemAxon [33]. The assignments are shown in Fig. 1. The α-CD guests were placed manually in the crystal structure [26] as suggested in [32].

Atom types of the hosts or the guests were assigned from the general Amber force field [15] using the antechamber tool [34]. This assigns all force field parameters except the charges. The latter were calculated in three different ways. First, AM1-BCC charges [18] were assigned using antechamber. Secondly, the molecule was optimized using AM1 and the electrostatic potential (ESP) was calculated for the molecule of interest at the HF/6-31G* level of theory, sampled using the Merz–Kollman scheme [35] but using a higher than default density of points (10 concentric layers with 17 points/Å^2). Thereafter, charges were fitted to the ESP using the RESP procedure [17] using antechamber. Third, eXAvESP charges were calculated in a procedure similar to what has been described previously [19, 20]. Twenty snapshots were extracted from 20 independent simulations of either the free host or the free guest (see below). Then, the ESP charges were obtained for each snapshot at the HF/6-31G* level using the Merz–Kollman sampling scheme. The final eXAvESP charges were then taken as the average over all $20 \times 20 = 400$ snapshot. AM1-BCC charges were used in the simulations. All QM calculations were done with Gaussian03 [36].

Each of the host–guest complexes were immersed in a truncated octahedral box of TIP3P [37] water molecules that extended at least 10 Å from the complex. For the LIE calculations, the free guests were also immersed in a similar box and care was taken so that the size of the box was similar to that of the corresponding complex box.

Simulations

All MD simulations were run using the sander module of Amber 10 [34]. The SHAKE algorithm [38] was used to

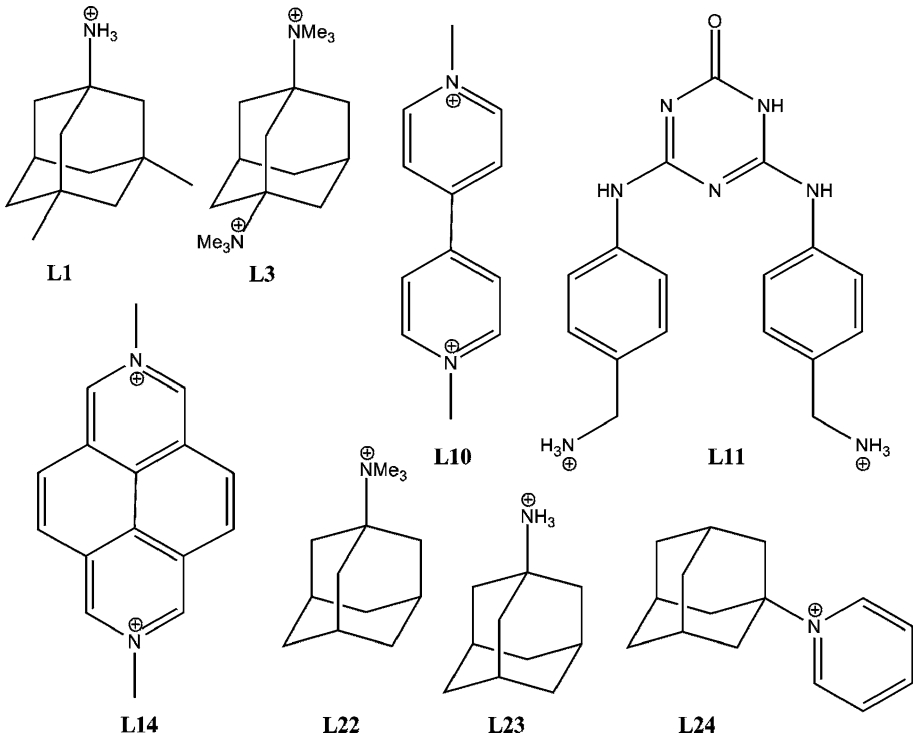

Fig. 1 The studied guest molecules for CB8

constrain bonds involving hydrogen atoms and a 2 fs time step was used in the integration of motion. The cut-off for the non-bonded interactions was set to 8 Å. Long-range electrostatics were handled using particle-mesh Ewald summation [39] and long-range van der Waals interactions were estimated by a continuum approach. The non-bonded pair-list was updated every 50 fs. The temperature was kept constant at 300 K using Langevin dynamics with a collision frequency of 2.0 ps^{-1} [40], and the pressure was kept constant at 1 atm using a weak-coupling isotropic algorithm with a relaxation time of 1 ps [41].

The systems were first minimised with 100 steps of steepest descent, using a harmonic restraint with a force constant of 418 kJ/mol/$Å^2$ on all atoms, except water molecules and hydrogen atoms. This was followed by a 20 ps simulation in the NPT ensemble and with the same restraint, and a 1 ns unrestrained simulation in the NPT ensemble. Thereafter, 20 independent simulations were initiated by using the last snapshot from the unrestrained equilibration but assigning different starting velocities, viz., the velocity-induced independent-trajectory approach [42]. Each of these simulations was further equilibrated for 50 ps in the NPT ensemble, followed by a 200 ps production run in the same ensemble. Coordinates were saved for MM/GBSA and LIE analysis every fifth ps. This follows a protocol optimized for protein–ligand systems [47], and is therefore most likely appropriate even for host–guest calculations.

Free energy calculations

The MM/GBSA method estimates the free energy using the following equations [7, 8]

$$\Delta G_{\text{bind}} = \langle G(\text{RL})\rangle - \langle G(\text{R})\rangle - \langle G(\text{L})\rangle \quad (1)$$

$$G = E_{\text{vdw}} + E_{\text{ele}} + \Delta G_{\text{solv}} + \Delta G_{\text{np}} - TS \quad (2)$$

where RL, R, and L are the host–guest complex, the receptor (the host) and the ligand (the guest), respectively. The energies in Eq. 2 are the van der Waals and electrostatic molecular mechanics energies, the polar solvation free energy, the non-polar solvation free energy, and the absolute temperature multiplied by an entropy estimate, respectively. E_{vdw} and E_{ele} were calculated with the same force field that was used in the MD simulations, but without any cut-off. ΔG_{solv} was estimated using either a generalised Born (GB) approach (giving the MM/GBSA method) or by solving the Poisson–Boltzmann (PB) equation (giving the MM/PBSA method). For the GB calculations, the method of Onufriev et al. [43], model I, was used, i.e., with $\alpha = 0.8$, $\beta = 0$, and $\gamma = 2.91$. For the PB calculations, the sander module was used with a spacing of 0.5 Å and a probe radius of 1.4 Å. ΔG_{np} was estimated from the solvent-accessible surface area according to $\Delta G_{\text{np}} = \gamma \text{SASA} + b$, where $\gamma = 0.0227$ kJ/mol/$Å^2$ and $b = 3.85$ kJ/mol [44]. The entropy was estimated as a sum of translational, rotational and vibrational entropies. The translational and rotational entropies were calculated from

statistical mechanical-formula for gas-phase molecules [44]. The vibrational entropies were calculated from harmonic frequencies calculated by the nmode module of Amber 10, after a minimisation with a distance-dependent dielectric constant model. The averages in Eq. 1 were all calculated for structures from the simulation of the complex (the host and guest coordinates were obtained by deleting the other species) and calculated from 40 snapshots for each of the 20 independent simulations. Hence, all the terms were evaluated on $40 \times 20 = 800$ snapshots.

The LIE method estimates the free energy using the following equation [9]

$$\Delta G_{\text{bind}} = \beta\left(\left\langle E_{\text{ele}}^{\text{L-S}}\right\rangle_{\text{RL}} - \left\langle E_{\text{ele}}^{\text{L-S}}\right\rangle_{\text{L}}\right) + \alpha\left(\left\langle E_{\text{vdw}}^{\text{L-S}}\right\rangle_{\text{RL}} - \left\langle E_{\text{vdw}}^{\text{L-S}}\right\rangle_{\text{L}}\right) \quad (3)$$

where $E^{\text{L-S}}$ is the interaction energy, either the electrostatic or van der Waals, between the guest and the surroundings (either solvent or host molecule). The angle brackets indicate ensemble averages and their subscripts indicate from which simulation they are computed (either a simulation of the complex or the free guest molecule). α and β are two parameters, derived from fitting and free energy calculations, respectively [22, 45]. The interaction energies in Eq. 3 were calculated by an in-house program, without any cut-off or long-range corrections. α was set to 0.18, and β was set according to the nature of the guests (0.5 for the CB8 guests and 0.37 for the α-CD guests).

Error estimates

All reported uncertainties are standard deviations (SD) of the mean (the SD divided by the square root of the number of samples). The reported standard error (SE) is the SD of the mean over the 20 independent simulations (ignoring the SD among the 40 snapshots in each simulation).

The performance of the free-energy estimates were quantified by the mean absolute deviation (MAD) compared to experimental data, the mean absolute deviation when the systematic error (as estimated by the mean signed deviation) has been removed (MADtr), the correlation coefficient (r^2), and Pearlman's predictive index (PI) [46]. MAD and r^2 are usually reported for affinity studies, whereas MADtr and PI measure the quality of the more important relative affinities. The SD of these quality measures was obtained by a simple parametric bootstrap simulation as has been described previously [47].

Results and discussion

The binding affinities of two sets of host–guest complexes using the popular end-point methods MM/GBSA and LIE have been estimated. To investigate the influence of the force-field on the accuracy of the results, three different charge models have been chosen. Each of these will be discussed in turn, starting with the AM1-BCC charges.

AM1-BCC charges

The MM/GBSA, MM/PBSA, and LIE results, using the AM1-BCC charges in both the simulations and energy calculations, are shown in Table 1 for the CB8 complexes and in Table 2 for the α-CD complexes. The MM/GBSA results reasonably reproduce the experimental order between the different guests, as shown by a correlation coefficient (r^2) of 0.7 and a predictive index (PI) of 0.8 (see also Fig. 2). However, the affinities are all too negative by $\sim$40 kJ/mol on average, indicated by the discrepancy between the MAD (61 kJ/mol) and MADtr (17 kJ/mol). This is a typical behaviour of this method, as has been shown previously for the estimation of protein–ligand affinities. In particular, it has been shown that the polar solvation energy to a large extent determines the absolute estimates [24]. This is supported by the results in Table 1 and Fig. 2, which show that MM/PBSA (using the PB solvation instead) gives absolute energies closer to the experimental results. The PI and r^2 are also significantly better with 0.9 and 0.9, respectively, and the difference between MAD and MADtr is small. Recently, it has been suggested that the MM/GBSA entropy is not very accurate [48] and because it is also the most time consuming part of the calculations it is sometimes omitted [49, 50]. The MM/GBSA and MM/PBSA results, excluding the entropy, are shown in Table S1, but the results are actually worse than the results in Table 1, although the PI of 0.8 shows that the ranking of the compounds is not affected by the entropy. The precision of the results are for most guests around or below 1 kJ/mol, with exception for the MM/PBSA result of L23, which have a SE of 3 kJ/mol. This is surprising, considering that the corresponding precision of the MM/GBSA result is 1 kJ/mol.

Looking at the LIE estimates instead, the results are much worse, with basically a zero correlation and a PI of only 0.3. In these calculations, the so-called standard model of LIE was used with a non-polar scaling, $\alpha = 0.18$, a value that has been optimised in many studies [22]. Therefore, α was freely optimised by minimising the MADtr of the predictions. The results of the optimisation are shown in Table 3, and show that the optimal scaling for this charge model is 0.34. The change in scaling, does not affect the MADtr particularly much (the difference is not statistical significant), but the PI and r^2 is improved to 0.5 and 0.1, respectively, changes that are statistically significant. This shows that the LIE might be useful, but require parametrisation. The precision is overall similar to the MM/GBSA and MM/PBSA results, although there are two guests with quite high SE, up to 2.1 kJ/mol.

Table 1 Free energy of binding for the CB8 complexes (kJ/mol)

	AM1-BCC			RESP			xAvESP			Experimental [25, 28]
	MM/GBSA	MM/PBSA	LIE	MM/GBSA	MM/PBSA	LIE	MM/GBSA	MM/PBSA	LIE	
L01	−114.9 ± 0.6	−77.3 ± 0.6	−17.9 ± 1.0	−117.8 ± 0.8	−75.3 ± 0.6	−10.5 ± 1.0	−115.8 ± 0.5	−64.9 ± 0.4	−9.3 ± 1.0	−66.8 ± 0.6
L03	−158.9 ± 0.6	−83.5 ± 0.4	15.9 ± 1.3	−121.2 ± 0.9	−65.7 ± 1.1	36.9 ± 1.5	−120.3 ± 0.4	−27.9 ± 0.4	56.4 ± 1.5	−69.2 ± 0.6
L10	−67.7 ± 0.8	−8.2 ± 0.7	6.9 ± 1.6	−33.0 ± 0.7	−4.8 ± 0.6	22.8 ± 1.5	14.8 ± 1.4	15.2 ± 1.8	42.6 ± 3.3	−29.0
L11	−138.6 ± 0.9	−89.3 ± 0.7	−91.2 ± 2.1	−71.0 ± 1.2	−47.7 ± 1.4	−98.7 ± 3.4	−92.4 ± 0.8	−70.6 ± 0.5	−86.7 ± 2.3	−61.8 ± 0.6
L14	−132.5 ± 1.5	−43.0 ± 1.1	5.7 ± 1.0	−68.8 ± 0.7	−38.6 ± 0.7	17.1 ± 1.6	−84.1 ± 1.9	−12.9 ± 1.0	37.2 ± 1.6	−50.6 ± 0.5
L22	−125.9 ± 0.9	−82.9 ± 0.4	−8.0 ± 0.7	−112.1 ± 0.8	−82.4 ± 0.5	−0.8 ± 0.7	−101.6 ± 0.5	−55.5 ± 0.6	−2.6 ± 0.9	−63.1 ± 0.6
L23	−90.3 ± 1.0	−38.8 ± 3.2	−14.7 ± 1.0	−81.1 ± 1.1	−48.2 ± 1.0	−5.5 ± 1.1	−79.6 ± 0.5	−34.9 ± 0.6	−3.9 ± 0.9	−51.2 ± 0.5
L24	−105.1 ± 0.4	−69.9 ± 0.5	−18.5 ± 0.8	−103.1 ± 0.4	−73.2 ± 0.5	−7.7 ± 0.9	−103.4 ± 0.4	−60.2 ± 0.4	−3.9 ± 0.9	−53.4 ± 0.6
MAD	61.1 ± 0.3	16.2 ± 0.5	47.7 ± 0.5	32.9 ± 0.3	13.0 ± 0.3	59.1 ± 0.6	40.6 ± 0.3	20.6 ± 0.3	65.6 ± 0.6	
MADtr	16.7 ± 0.3	14.7 ± 0.6	20.9 ± 0.5	17.5 ± 0.3	12.7 ± 0.3	23.7 ± 0.6	18.7 ± 0.3	18.3 ± 0.3	27.0 ± 0.6	
r^2	0.67 ± 0.01	0.87 ± 0.01	0.06 ± 0.01	0.77 ± 0.01	0.72 ± 0.01	0.06 ± 0.01	0.89 ± 0.01	0.56 ± 0.01	0.09 ± 0.01	
PI	0.79 ± 0.00	0.91 ± 0.01	0.26 ± 0.06	0.91 ± 0.01	0.77 ± 0.03	0.33 ± 0.01	0.91 ± 0.00	0.68 ± 0.00	0.32 ± 0.05	

Experimental uncertainty is missing for L10 because the affinity is from another earlier study [23]

Table 2 Free energy of binding for the α-CD complexes (kJ/mol)

	AM1-BCC			RESP			xAvESP			Experiment [29]
	MM/GBSA'	MM/PBSA'	LIE	MM/GBSA'	MM/PBSA'	LIE	MM/GBSA'	MM/PBSA'	LIE	
1-methanol	−3.5 ± 0.1	−3.7 ± 0.1	5.5 ± 0.5	−3.4 ± 0.1	−3.6 ± 0.1	4.7 ± 0.6	−3.2 ± 0.1	−3.3 ± 0.2	3.5 ± 0.8	0.2
1-ethanol	−17.0 ± 1.2	−19.4 ± 1.7	−0.1 ± 0.3	−3.6 ± 0.1	−4.0 ± 0.1	7.2 ± 0.6	−17.2 ± 1.5	−19.0 ± 1.8	−0.2 ± 0.4	−4.3
1-propanol	−16.9 ± 2.6	−18.5 ± 3.0	−0.6 ± 0.8	−26.7 ± 1.2	−27.4 ± 1.6	−1.8 ± 0.3	−15.5 ± 1.0	−13.4 ± 1.4	−2.3 ± 0.4	−7.8
1-butanol	−35.0 ± 0.3	−40.2 ± 0.5	−3.3 ± 0.3	−36.3 ± 0.6	−38.8 ± 0.6	−2.3 ± 0.3	−36.3 ± 0.4	−40.1 ± 0.4	−3.6 ± 0.3	−11.1
1-pentanol	−43.0 ± 0.5	−49.7 ± 0.6	−5.4 ± 0.4	−44.8 ± 0.3	−48.5 ± 0.3	−3.2 ± 0.4	−42.7 ± 0.5	−48.1 ± 0.6	−6.7 ± 0.5	−14.3
1-hexanol	−50.3 ± 0.5	−58.0 ± 0.6	−6.8 ± 0.3	−50.3 ± 0.5	−53.6 ± 0.6	−5.6 ± 0.3	−47.1 ± 0.4	−52.6 ± 0.4	−7.1 ± 0.3	−16.8
MAD	18.6 ± 0.5	22.5 ± 0.6	7.3 ± 0.2	18.7 ± 0.2	20.4 ± 0.3	8.9 ± 0.2	18.0 ± 0.3	20.4 ± 0.4	6.3 ± 0.2	
MADtr	10.1 ± 0.5	12.7 ± 0.6	1.7 ± 0.2	11.4 ± 0.2	12.6 ± 0.2	2.4 ± 0.2	10.0 ± 0.3	12.4 ± 0.4	2.0 ± 0.2	
r^2	0.96 ± 0.02	0.95 ± 0.02	0.96 ± 0.02	0.94 ± 0.00	0.95 ± 0.00	0.80 ± 0.04	0.94 ± 0.01	0.90 ± 0.02	0.98 ± 0.01	
PI	0.94 ± 0.03	0.94 ± 0.03	1.00 ± 0.03	1.00 ± 0.01	1.00 ± 0.01	0.93 ± 0.02	0.94 ± 0.02	0.94 ± 0.01	1.00 ± 0.02	

Entropy estimates are excluded from the MM/GBSA and MM/PBSA results as indicated by the '

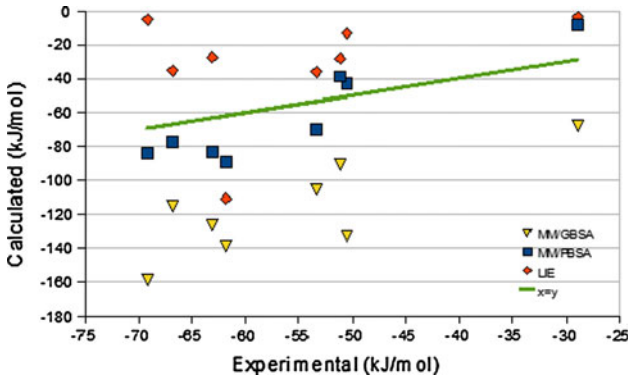

Fig. 2 Correlation between experimental and calculated binding affinities with the AM1-BCC charges for the CB8 test case

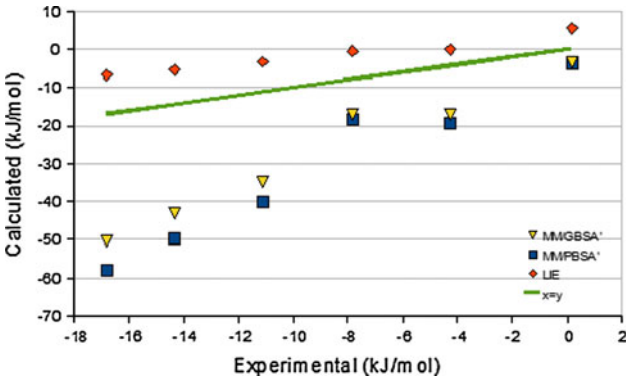

Fig. 3 Correlation between experimental and calculated binding affinities with the AM1-BCC charges for the α-CD test case

Table 3 Quality measures of optimised LIE for CB8 (kJ/mol)

	AM1-BCC	RESP	xAvESP
Optimised α	0.34	0.46	0.08
MAD	35.4	39.7	70.9
MADtr	19.5	23.6	25.6
r^2	0.10	0.12	0.04
PI	0.46	0.59	0.21

Even though the MM/GBSA and MM/PBSA results were not improved when the entropy was excluded from the CB8 complexes, such exclusion was necessary to obtain reasonable results for the α-CD complexes. The results without the entropy are shown in Table S2 and it can be seen that the correlation is weak and the PI is zero. If the entropy estimate is omitted, a good correlation is obtained, as can be seen in Table 2. The PI is 0.9 and $r^2 > 0.9$ for both MM/GBSA and MM/PBSA. Furthermore, it is clear that the systematic error is much smaller, ~10 kJ/mol. For this test case, the MM/GBSA and MM/PBSA results are much more similar, with a maximum difference of ~8 kJ/mol (see for instance Fig. 3). The precision of the results are similar to the CB8 complexes and most guests have a SE of around or below 1 kJ/mol. A notable exception is 1-propanol that gives a SE of ~3 kJ/mol for both MM/GBSA and MM/PBSA.

Even though MM/GBSA and MM/PBSA give good results for this test case, the LIE method is actually much better with a PI of 1.0 and $r^2 > 0.95$ (see Fig. 2). The estimates are systematically about 5 kJ/mol too positive, but the MADtr is less than 2 kJ/mol. It is noticeable that this was achieved with the standard LIE model, which shows that the standard α-value is system dependent. The precision is consistently less than 1 kJ/mol, and in most cases below 0.5 kJ/mol.

RESP charges

The MM/GBSA, MM/PBSA and LIE results using the RESP charges in both the simulations and energy calculations are also shown in Tables 1 and 2. Using this charge model on the CB8 complexes, the MM/GBSA results are slightly improved compared to those obtained with AM1-BCC charges, whereas the MM/PBSA results are slightly deteriorated. The PI and r^2 for the MM/GBSA method is 0.8 and 0.9, respectively. Using this charge model, the energies are only ~30 kJ/mol too negative on average, as indicated by the MAD. However, the MADtr of 18 kJ/mol is slightly worse than when using the AM1-BCC charges. The MM/PBSA gives a PI and r^2 of 0.7 and 0.8 kJ/mol, respectively, and as with AM1-BCC charge, the absolute energies are much closer to the experimental results than the MM/GBSA results. Excluding the entropy from the calculations, the correlation with experimental results improves for this charge model, as can be seen in Table S1. This indicates that excluding the entropy is not a universal fix that could be applied to improve MM/GBSA results, and that the estimates involve an intricate balance between different components. The precision of the result is similar to the AM1-BCC results, and all guests have a SE of <1.5 kJ/mol.

The LIE method using $\alpha = 0.18$, gives a poor PI and no correlation and therefore, an optimisation of α was necessary. The optimised scaling factor for this charge model is 0.46 (see Table 3), i.e. a slightly larger value than was obtained for the AM1-BCC charges. Using the optimised α, the correlation is improved to 0.1 and the PI to 0.6, changes that are statistical significant. Both the MAD and MADtr are slightly worse than those obtained with the AM1-BCC charges. The precision is slightly worse and more varying with the RESP charges, than with the AM1-BCC charges. The worst precision is seen for L11 with 3 kJ/mol.

The MM/GBSA and MM/PBSA results for the α-CD complexes, are very similar to the AM1-BCC results, if one

look at the quality measures, although the PI is somewhat improved to 1.0. However, two of guests have statistically significant difference compared to the AM1-BCC results. The largest differences, compared to AM1-BCC, occur for the 1-ethanol and 1-propanol. However, the LIE results are slightly worse with this charge model. This mainly stems from the fact that the affinity of 1-ethanol is predicted to be more positive than 1-methanol, which disrupts the perfect order that was obtained with the AM1-BCC charges. The other guests display a much smaller change, compared to the AM1-BCC results. Taking all these results into consideration, it seems that this test case is much less sensitive to the charge model than the CB8 test case. The precision of the MM/GBSA and MM/PBSA methods is overall slightly better than the AM1-BCC results, but again the precision of 1-propanol is an exception. The SE of the LIE estimates are all below 1 kJ/mol.

xAvESP charges

The MM/GBSA, MM/PBSA and LIE results, using the xAvESP charges in both the simulations and energy calculations, are shown in Tables 1 and 2. Similar to the RESP results for CB8, the MM/PBSA results are worse with this charge model, than with AM1-BCC. For the MM/GBSA method, the xAvESP charges gave the best MM/GBSA results for the CB8 test case, with a PI and r^2 of 0.9. However, the MADtr is slightly worse than with RESP or AM1-BCC charges. The correlation and PI for the MM/PBSA results is significantly lower, 0.6 and 0.7, respectively, whereas MADtr is not statistically different compared to MM/GBSA. Again, the absolute MM/PBSA results are closer to experimental affinities than MM/GBSA. The precision of the results are similar to the precision of the AM1-BCC and RESP calculations, with only one guest with a SE larger than 1.5 kJ/mol.

As with the other charge models, the LIE results show a zero r^2 and PI of 0.3. The optimised $\alpha = 0.08$, is quite different from the values obtained with the RESP and AM1-BCC models (see Table 3). The correlation and PI did not improve, r^2 becomes 0.4, and the PI is actually worse (0.2). This mainly stems from a problem of ordering L03 and L10 correctly. Experimentally L03 has a higher affinity than L10, but the xAvESP LIE results show the opposite trend. However, it should be noted that the LIE method with all charge models seem to have a problem in estimating the affinity of L03. Instead of using the scaling parameter optimised for xAvESP charges, the value of 0.46, as optimised for RESP charges, could be used. Using this scaling, the r^2 is 0.22, PI is 0.58, and MADtr is 26 kJ/mol. This indicates that optimising the LIE scaling parameters is a non-trivial problem.

Looking at the α-CD guests, the MM/GBSA and MM/PBSA quality measures are similar to the quality measures using AM1-BCC or RESP charges as expected, although deviations for individual guests are present. The PI and $r^2 > 0.9$ for both MM/GBSA and MM/PBSA. With LIE, there was no trouble with the 1-ethanol guest with this charge model, so that the correlation and PI is almost perfect. The precision of the results is similar to that of the other calculations with only a single guest with a SD larger than 1.5 kJ/mol for the MM/GBSA and MM/PBSA methods.

Instead of simulating another set of structures with the xAvESP charges, the structures generated with AM1-BCC charges could be used in free energy calculations with xAvESP charges. This is analogous to a protocol in quantum chemistry, where a cheap method is used to optimise structure, followed by a more expensive method for the energy calculation. Of course it is possible to make other permutations of the charge models, i.e., simulating with one set of charges and re-calculate with another. However, the combination of simulating with AM1-BCC and energy calculating with xAvESP charges is the only one that makes sense from an efficiency perspective. The quality measurements of these results are shown in Table 4. The MM/GBSA results for the CB8 test case is somewhat improved compared to both the pure AM1-BCC and pure xAvESP calculations. The MADtr is improved to 13 kJ/mol, but the correlation and PI is unchanged compared to the pure xAvESP results, but slightly improved compared to pure AM1-BCC results. For the MM/PBSA method, the MADtr is 20 kJ/mol, r^2 0.5, and PI 0.70. Therefore, the results are unchanged compared to pure xAvESP results, and slightly worse than the pure AM1-BCC results. The LIE energies require optimisation of α as with the pure AM1-BCC and pure xAvESP results. Using the optimised results, the LIE results are similar to the pure AM1-BCC results. The MADtr is 26 kJ/mol, r^2 0.2, and PI 0.5.

Table 4 Quality measures of calculated affinities

	MM/GBSA	MM/PBSA	LIE
	CB8		
MAD	42.8 ± 0.3	21.8 ± 0.2	82.3 ± 0.4
MADtr	12.6 ± 0.3	19.7 ± 0.2	29.5 ± 0.5
r^2	0.89 ± 0.01	0.47 ± 0.01	0.03 ± 0.01
PI	0.91 ± 0.01	0.70 ± 0.00	0.16 ± 0.05
	α-CD		
MAD	17.3 ± 0.5	19.6 ± 0.5	7.0 ± 0.2
MADtr	10.0 ± 0.5	11.4 ± 0.6	2.1 ± 0.2
r^2	0.96 ± 0.02	0.95 ± 0.02	0.95 ± 0.02
PI	1.00 ± 0.03	0.94 ± 0.03	1.00 ± 0.03

The AM1-BCC charges were used in the simulations, but xAvESP charges were used in the free energy calculations

For the α-CD test case, the correlation for the MM/GBSA and MM/PBSA methods are slightly improved too but the correlation is in any case >0.94. The MADtr is 10 and 11 kJ/mol, for MM/GBSA and MM/PBSA, and hence no improvement compared to pure AM1-BCC and pure xAvESP results. Looking at LIE, there is no improvement in the MADtr because it is still about 2 kJ/mol. The correlation coefficient is slightly worse, but still 0.95, and the PI is 1.0 with all methods.

Conclusions

Binding free energies of two sets of host–guest complexes with the MM/GB(PB)SA and LIE methods have been estimated. Several general conclusions have been found

- There is no consistent improvement when more rigorous charge models are used. AM1-BCC charges tend to give a large MAD, but because relative energies are most important this trend is of minor relevance. Using xAvESP charges in the free energy calculations on structures generated with AM1-BCC charges also does not guarantee an improvement. Therefore it is more economic to use the AM1-BCC charges because these are the fastest to obtain. The AM1-BCC charge calculation on the host takes up to an hour on a single processor, whereas the guest only takes a few minutes.
- The LIE method is highly system dependent. The standard model gives a zero correlation for the CB8 systems but an almost perfect correlation for the α-CD systems. An optimisation of the non-polar scaling parameter, α was necessary to obtain reasonable results for the CB8 systems although the correlation is still mediocre. Furthermore, different charge models give different optimised α-values.
- There is no consistent benefit of using either MM/GBSA or MM/PBSA. They give reasonably similar results for relative affinities and therefore it would be advisable to use the cheaper MM/GBSA method. This is in line with studies of protein–ligand systems [24], although PB has been shown to be the more accurate method for small molecule solvation free energies [51].
- It is more difficult to determine which is more accurate of the MM/GBSA and LIE methods. For the CB8 test system, the MM/GBSA method is clearly preferable because no parametrisation is necessary. However, for the α-CD system LIE seems to give the best results. The r^2 is larger than 0.9 for both methods, but LIE shows a much better MADtr.

Both MM/GBSA and LIE have been shown to give affinity estimates that well reproduce the experimental order between different guest molecules. This implies that the general Amber force field is sufficiently good to also simulate these kinds of systems, and therefore, it is not necessary to re-parameterise the force field. A larger test set is necessary to draw any final conclusions about the relative merits of the MM/GBSA and LIE methods. However, LIE seems to be highly system dependent and the method cannot readily be used to study charged host molecules without introducing explicit salt. Another issue is predictability, i.e., if the method is able to predict unknown affinities. Here, LIE would not be very useful because α cannot be known a priori. The problem with MM/GBSA and MM/PBSA in such a scenario is whether to include the entropy or not. Although, the results for CB8 were better with entropy, the results without entropy were not entirely bad. For instance, PI was 0.8 with AM1-BCC charges, indicating that the arguably most important quality, ranking, is preserved without entropy. Hence, it seems reasonable to omit the entropy at a first stage for any host–guest complex. Taking altogether, it seems that MM/GBSA is the method of choice.

Acknowledgments This investigation has been supported by grants from the Research school in pharmaceutical science. It has also been supported by computer resources of LUNARC at Lund University (project SNIC001-10-225), NSC at Linköping University and HPC2 N at Umeå University (project SNIC014-10-24). Prof. Ulf Ryde and Svante Hedström are acknowledged for discussions and critical proofreading.

References

1. Chipot C, Pohorille A (eds) (2007) Free Energy Calculations. Springer, New York
2. Shirts M, Mobley DL, Chodera JD (2007) Annu Rep Comput Chem 3:41–59
3. Michel J, Essex JW (2010) J Comput Aided Mol Des 24:639–658
4. Gohlke H, Klebe G (2002) Angew Chem Int 41:2644–2676
5. Moghaddam S, Inoue Y, Gilson MK (2009) J Am Chem Soc 131:4012–4021
6. Mikulskis P, Genheden S, Sandberg L, Olsen L, Ryde U (2011) J Comput Aided Mol Des, submitted
7. Srinivasan J, Cheatham TE III, Cieplak P, Kollman PA, Case DA (1998) J Am Chem Soc 37:9401–9409
8. Kollman PA, Massova I, Reyes I, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, CheathamIII TE (2000) Acc Chem Res 33:889–897
9. Åqvist J, Medina C, Samuelsson J-E (1994) Protein Eng 7:385–391
10. Hansson T, Marelius J, Åqvist J (1998) J Comput-Aided Mol Design 12:27–35
11. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) J Am Chem Soc 117:5179–5197
12. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe

M, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) J Phys Chem B 102:3586–3616
13. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) J Am Chem Soc 118:11225–11236
14. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF (2004) J Comput Chem 25:1656–1676
15. Wang JM, Wolf RM, Caldwell KW, Kollman PA, Case DA (2004) J Comput Chem 25:1157–1174
16. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, Mackerell AD Jr (2009) J Comput Chem 31:671–690
17. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1997) J Phys Chem 97:10269–10280
18. Jakalian A, Jack DB, Bayly CI (2002) J Comput Chem 23:1623–1641
19. Söderhjelm P, Ryde U (2009) J Comput Chem 30:750–760
20. Genheden S, Söderhjelm P, Ryde U (2011) Int J Quant Chem. doi:10.1002/qua.22967
21. Wall ID, Leach AR, Salt DW (1999) J Med Chem 42:5142–5152
22. Almlöf M, Brandsdal BO, Åqvist J (2004) J Comput Chem 25:1242–1254
23. Lee MC, Yang R, Duan Y (2005) J Mol Model 12:101–110
24. Genheden S, Luchko T, Gusarov S, Kovalenko A, Ryde U (2010) J Phys Chem B 114:8505–8516
25. Liu S, Ruspic C, Mukhopadhyay P, Chakrabarti S, Zavalij PY, Isaacs L (2005) J Am Chem Soc 127:15959–15967
26. Puliti R, Mattia CA, Paduano L (1998) Carbohydr Res 310:1–8
27. Linder K, Saenger W (1982) Carbohydr Res 99:103–115
28. Jeon WS, Kim H-J, Lee C, Kim K (2002) Chem Commun 1828–1929
29. Matsui Y, Mochida K (1979) Bull Chem Soc Jpn 52:2808–2814
30. Choi Y, Jung S (2004) Carbohydr Res 339:1961–1966
31. El-Barghouthi MI, Assaf KI, Rawashdeh AMM (2010) J Chem Theory Comput 6:984–992
32. El-Barghouthi MI, Jaime C, Al-Sakhen NA, Issa AA, Abdoh AA, Al Omari MM, Badwan AA, Zughul MB (2008) J Mol Struct: Theochem 853:45–52
33. ChemAxon, Marvin version 5.6 (2011)
34. Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Crowley M, Walker RC, Zhang W, Merz KM, Wang B, Hayik S, Roitberg A, Seabra G, Kolossvary I, Wong KF, Paesani F, Vanicek J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Mathews DH, Seetin MG, Sagui C, Babin V (2008) Kollman, P A Amber 10. University of California, San Francisco
35. Besler BH, Merz KM, Kollman PA (1990) J Comput Chem 11:431–439
36. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian, 03 Revision. Gaussian Inc, Wallingford
37. Jorgensen WL, Chandrasekhar J, Madura JD, Impley RW, Klein ML (1983) J Chem Phys 79:926–935
38. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) J Comput Phys 23:327–341
39. Darden T, York D, Pedersen L (1993) J Chem Phys 98:10089–10092
40. Wu X, Brooks BR (2003) Chem Phys Lett 381:512–518
41. Berendsen HJC, Postma JPM, Van Gunsteren WF, Dinola A, Haak JR (1984) J Chem Phys 81:3684–3690
42. Genheden S, Ryde U (2011) J Comput Chem 32:187–195
43. Onufriev A, Bashford D, Case DA (2004) Proteins 55:383–394
44. Kuhn B, Kollman PA (2000) J Med Chem 43:3786–3791
45. Åqvist J, Hansson T (1996) J Phys Chem 100:9512–9521
46. Pearlman DA, Chareifson PS (2001) J Med Chem 44:3417–3423
47. Genheden S, Ryde U (2010) J Comput Chem 31:837–846
48. Singh N, Warshel A (2010) Proteins 78:1705–1723
49. Hayes JM, Skamnaki VT, Archontis G, Lamprakis C, Sarrou J, Bischler N, Skaltsounis A-L, Zographos SE, Oikonomakos NH (2011) Proteins 79:704–719
50. Hou T, Wang J, Li Y, Wang W (2011) J Chem Inf Model 51:69–82
51. Shivakumar D, Deng Y, Roux B (2009) J Chem Theory Comput 5:919–930