Comparison of the Efficiency of the LIE and MM/GBSA Methods to Calculate Ligand-Binding Energies

Genheden, Samuel; Ryde, Ulf

Published in:
Journal of Chemical Theory and Computation

DOI:
10.1021/ct200163c

2011

Comparison of the efficiency of the LIE and MM/GBSA methods to calculate ligand-binding energies

Samuel Genheden and Ulf Ryde*

Department of Theoretical Chemistry, Lund University, Chemical Centre,
P. O. Box 124, SE-221 00 Lund, Sweden

*Correspondence to Ulf Ryde, E-mail: Ulf.Ryde@teokem.lu.se,
Tel: +46 – 46 2224502, Fax: +46 – 46 2224543

2011-09-20

We have evaluated the efficiency of two popular end-point methods to calculate ligand-binding free energies, LIE (linear interaction energy) and MM/GBSA (molecular mechanics with generalised Born surface-area), i.e. the computational effort needed to obtain estimates of a similar precision. As a test case, we use the binding of seven biotin analogous to avidin. The energy terms used by MM/GBSA and LIE exhibit a similar correlation time (~5 ps) and the equilibration time seems also to be similar, although it varies much between the various ligands. The results show that the LIE method is more effective than MM/GBSA, by a factor of 2–7 for a truncated spherical system with a radius of 26 Å and by a factor of 1.0–2.4 for the full avidin tetramer (radius 47 Å). The reason for this is the cost for the MM/GBSA entropy calculations, which more than compensates for the extra simulation of the free ligand in LIE. On the other hand, LIE requires that the protein is neutralised, whereas MM/GBSA has no such requirements. Our results indicate that both the truncation and neutralisation of the proteins may slow the convergence and emphasize small differences in the calculations, e.g. differences between the four subunits in avidin. Moreover, LIE cannot take advantage of the fact that avidin is a tetramer. For this test case, LIE gives poor result with the standard parametrisation, but after optimising the scaling factor of the van der Waals terms, reasonable binding affinities can be obtained, although MM/GBSA still gives significantly better predictive index and correlation to the experimental affinities.

Introduction

One of the most important challenges of computational chemistry is to accurately estimate the free-energy change of a biochemical reaction. For instance, in drug design, one is interested in the binding of small ligands to a biomolecular target, usually a protein. If accurate free energies could be estimated for this reaction by computational methods, billions of dollars could be saved because it would be necessary to synthesize fewer molecules.[1,2]

Many methods are available to estimate free energies, ranging from simple scoring functions that are fast, but not very accurate, to rigorous free energy perturbation (FEP), which are accurate, but time-consuming.[34] The reason for the latter is that FEP requires extensive sampling using molecular dynamics (MD) or Monte Carlo methods on a series of intermediate, unphysical states. A class of methods that is intermediate in efficiency is the so-called end-point methods, which still are based on physical laws and require sampling, but only of the reactants and the products, not of any intermediate states.[5] However, even with perfect sampling these method will not give the exact result, because they are based on several approximations. Therefore, such methods need to evaluated carefully to identify their strengths and weaknesses.

Two such methods are LIE[6,7,8] (linear interaction energy) and MM/GBSA[9,10] (molecular mechanics with generalised Born and surface-area solvation). LIE estimates the free energy for the binding of a ligand (L) to its target macromolecule (P) by simulating the free ligand in solution and the ligand–macromolecule complex (PL), using the relation[7]

$$\Delta G = \beta\left(\langle E_{\mathrm{ele}}^{\mathrm{L-S}}\rangle_{PL} - \langle E_{\mathrm{ele}}^{\mathrm{L-S}}\rangle_{L}\right) + \alpha\left(\langle E_{\mathrm{vdW}}^{\mathrm{L-S}}\rangle_{PL} - \langle E_{\mathrm{vdW}}^{\mathrm{L-S}}\rangle_{L}\right) \quad (1)$$

where $E_{\mathrm{ele}}^{\mathrm{L-S}}$ and $E_{\mathrm{vdW}}^{\mathrm{L-S}}$ are the electrostatic and van der Waals intermolecular interaction energies between the ligand and the surroundings (S; i.e. protein and solvent), α and β are two parameters, and the angle brackets indicate ensemble averages from the simulations of either the free ligand or the complex, as indicated by the subscripts. β was originally set to 0.5,[6] because LIE was derived from the linear-response approximation. However, this value has later been refined to reflect the chemical nature of the ligand,[7,11,12] based on FEP calculations. α is usually set to 0.18,[13,14] but this value has been much debated and may be system dependent.[5,8] In several studies, this parameter has been fitted to experimental data for each protein target and ligand type.[5] A third constant term has also been suggested.[15] but it is important only when estimating absolute free energies.[13]

MM/GBSA, on the other hand, estimates the free energy as[9,10]

$$\Delta G = G(\mathrm{PL}) - G(\mathrm{P}) - G(\mathrm{L}) \quad (2)$$

where each free energy is calculated from a sum of six terms

$$G = \langle E_{\mathrm{int}} + E_{\mathrm{ele}} + E_{\mathrm{vdW}} + G_{\mathrm{solv}} + G_{\mathrm{np}} - TS_{\mathrm{MM}}\rangle \quad (3)$$

The three first terms are the molecular-mechanics (MM) internal, electrostatics, and van der Waals energies, G_{solv} is the polar solvation energy, G_{np} is the non-polar solvation free energy, T is the absolute entropy, and S_{MM} is an entropy estimate from harmonic frequencies calculated at the MM level. The average in Eqn. 3 should in principle be calculated from three separate simulations PL, P, and L, but for stability reasons.[16] it is more common to simulate only the complex. In that case, E_{int} cancels. In the MM/GBSA approach, the polar solvation free energy is calculated by a generalised Born (GB) approach, but it could be calculated by any continuum-solvation method.[17] A common choice is the Poisson–Boltzmann method, giving the MM/PBSA approach. The non-polar solvation free energy is usually estimated by a

relation to the solvent-accessible surface area (SASA).[5]

Because LIE and MM/GBSA are two popular end-point methods, it is interesting to compare them. This has been done a few times in the past.[18,19,20,21] For the binding of biotin analogous to avidin, MM/PBSA was found to reproduce experimental results more accurately than LIE.[19] although both methods were less accurate than FEP.[22] However, for acetylcholinesterase huprine inhibitors, LIE gave better results than MM/PBSA.[18] For the binding of eight hydroxamate inhibitors to gelatinase A, the two methods showed a similar performance.[20] In all these studies, the LIE parameters were adjusted to an optimal fit. For the binding of fragment B of protein A to the Fc domain of immunoglobin G, LIE gave similar results to both MM/PBSA and MM/GBSA, but only one complex was examined, a protein for which it is not clear what α and β parameters should be used.[21] Apparently, the accuracy of the two methods (i.e. how well they reproduce experimental results) depends strongly on the systems studied and much larger test sets are needed before any general conclusion can be reached.

In this paper, we will instead focus on the precision (i.e. the statistical uncertainty of the results) and efficiency (i.e. the computer time required to reach a given precision) of the two methods. The statistical precision is important when comparing ligand-affinity methods[23]: Congeneric ligands often have quite similar affinities and an order of magnitude difference in the binding constant corresponds to only 6 kJ/mol in the free energy of binding. If statistically significant differences should be discerned, a precision of 1–2 kJ/mol is therefore needed. Such a precision is also needed to make results obtained by different groups comparable[24] and to avoid the temptation to rerun simulations that gave poor agreement with experiments. On the other hand, we have shown that once such a precision is reached, results obtained by MM/GBSA are reproducible and not sensitive to the setup of the simulations, except for the protonation of residues very close to the ligand[24]

In a previous paper, we developed a simulation protocol for MM/GBSA that gave a precision of 1 kJ/mol.[23] In particular, we showed that it was more favourable to run several rather short simulations instead of a single long one, as has been concluded also in other studies.[25,26,27] By running a proper number of independent simulations, any precision can be reached. In this paper, we develop a similar protocol for LIE. This also allows us to discuss the efficiency of the two methods, i.e. to compare the computational effort needed to obtain results of the same statistical precision. If the methods give similar accuracy, of course the more efficient method is preferred. To facilitate the comparison, we use the same test case as for MM/PBSA, viz. the binding of seven biotin analogues to avidin. This test system has been studied before with FEP,[28] MM/PB(GB)SA,[17,19,23,24,29,30,31,32,33] LIE ,[22] and experimentally structures[34] as well as affinities are available.[35,36,37]

Methods

System preparation

We have studied the binding of the seven biotin analogues in Figure 1 to avidin. Btn1–Btn3 have a net charge of –1, whereas the other four ligands are neutral. The structure of avidin was taken from the 1avd crystal structure,[34] which contains a co-crystallized biotin molecule in each subunit of the tetrameric protein. However, in this study we consider the binding of only a single ligand to the tetrameric protein. The six biotin analogues were built into the active site to mimic the binding mode of biotin, as been described previously.[30] In LIE, it is essential that the protein is neutral.[8] Therefore, all titratable residues were neutralised (all these residues are solvent exposed). This has shown to be the optimal approach to ensure that the complex and free ligand simulations have identical total charge, which is required if we

want to ignore long-range effects beyond the simulation sphere.[38] The single histidine residue in each subunit was modelled to be protonated on the NE2 atom.[30] The protein atoms were described by the Amber99SB force field[39] and parameters for the ligands were taken from the Amber99 force field.[30,40] Ligand charges were calculated with the RESP procedure,[41] using ESP points calculated at the Hartree–Fock 6-31G* level and sampled with the Merz–Kollman scheme,[42] as has been described before.[30]

Two sets of systems were prepared for each protein–ligand complex, a full system and a truncated system. The full system was prepared by solvating the entire (tetrameric) protein–ligand complex in a sphere with radius 47 Å (i.e. extending at least 10 Å outside the protein; in total ~43 925 atoms). The truncated system was prepared by solvating the complex in a 26 Å sphere, centred on the ligand and thereafter, removing all residues more than 26 Å from the ligand (~8 325 atoms). Atoms between 26 and 24 Å were restrained in the simulations, by a harmonic restraint of 41.84 kJ/mol/Å^2. The truncated system represents a more typical use of LIE.[8] Likewise, two sets of free-ligand systems were created by solvating the ligand in a sphere with a radius of either 47 or 26 Å, because LIE requires that the simulations of the complex and the free ligand have the same size, so that the ignored interactions outside the simulated systems cancel.[8] In these simulations, the geometrical centre of the ligand was restrained to the origin using a harmonic restraint of 41.84 kJ/mol/Å^2. In all simulations, the water model was TIP3P.[43] The systems were prepared with a combination of the Qprep program of the Q simulation package, the Leap module of the Amber package, and in-house programs.

Simulations

All MD simulations were run by the Q simulation package.[44] All bonds involving hydrogen atoms were constrained with the SHAKE approach[45] and a 2 fs time step was employed. The temperature was kept at 300 K using a Berendsen thermostat.[46] The non-bonded cut-off was 10 Å, except for interactions with the ligand, for which an infinite cut-off were applied. Long-range electrostatic interactions were treated with the local reaction-field approximation.[47] Water molecules at the surface of the simulated sphere were subjected to radial and polarisation restraints.[44,48]

Prior to the MD simulation, the systems were minimised using the sander module of Amber 10[49] using 100 steps of steepest descent and with a harmonic restraint of 104.6 kJ/mol/Å^2 on all atoms except hydrogen and water atoms. This was followed by starting a number of independent 20 ps MD simulations with the same restraint as in the minimisation. Thereafter, an unrestrained MD simulation was carried out for 800 ps (full systems) or 1600 ps (truncated systems) for each of the independent simulations. Snapshots were sampled each ps. Twenty independent simulations were employed for the free ligand and for each of the subunits of avidin (i.e. 20 + 80 in total). These independent simulations were initiated by assigning different initial random velocities to all atoms, i.e. the velocity-induced independent-trajectory approach.[24]

Free energy estimates

The LIE interaction energies in Eqn. 1 (with an infinite cut-off, but without any long-range corrections) were sampled by Q[44] during the simulation and were processed by in-house scripts. β was set to 0.5 for the charged ligands and 0.43 for the other ligands,[7] whereas α was set to 0.18 as default for all ligands[13] although it was also optimised (see below). A new parametrisation of β to include more chemical groups has been suggested,[12] but it does not involve thioether and other groups in our ligand set. For the charged ligands, a correction to the neutralisation of the charged residues, ΔG_{cc}, was estimated by placing a single charge at the position of the CG, CD, NZ, and CZ atoms of Asp, Glu, Lys, and Arg, respectively, and

calculating the Coulomb interaction between this charge and all atoms in the ligand for each snapshot, assuming a dielectric constant of 80, as suggested previously.[38,50,51] We tested this correction also for the neutral ligands, but it was found to be negligible, ~0.1 kJ/mol.

The MM/GBSA calculations were performed by the Amber 10 package on snapshots from the Q MD simulations.[49] The E_{ele} and E_{vdW} energies in Eqn. 3 were calculated with the same force field as the simulation and with an infinite cut-off. The polar solvation free energy was estimated by the GB method of Onufriev, Bashford and Case,[52] model I (OBCI, i.e. with $\alpha = 0.8$, $\beta = 0$, and $\gamma = 2.91$). The non-polar solvation energy was estimated from the SASA according to $\Delta G_{np} = \gamma$ SASA + b, where $\gamma = 0.0227$ kJ/mol/Å^2 and $b = 3.85$ kJ/mol.[53] The entropy was estimated by calculating harmonic frequencies at the MM level on a truncated and buffered system (8 + 4 Å from the ligand), as described previously, to improve the statistical precision of the estimate.[32]

Estimation of the correlation time

The correlation time of the LIE interaction energies was estimated with the statistical inefficiency method.[54,55] In this procedure, the following measure is calculated

$$\Phi = \frac{\tau \sigma^2(Y)_\tau}{\sigma^2(X)} \qquad (4)$$

where $\sigma^2(X)$ is the variance of the distribution $\{X\}$, i.e. the variance of the time series of a particular energy, e.g., E_{ele}^{L-S} in Eqn. 1, and $\sigma^2(Y)_\tau$ is the variance of the block average of $\{X\}$, where the block length is τ. This block average is calculated from

$$Y_i = \frac{1}{\tau} \sum_{j=n-i\tau+1}^{n-(i-1)\tau} X_j \qquad (5)$$

Put in another way, $\{X\}$ is divided into a number of non-overlapping segments, each with length τ. Once τ is so large that the successive values of Y_i are statistically independent, Φ will become a constant and an estimate of the correlation time of $\{X\}$. This method is sensitive to equilibration, long-time trends, and bumps in the data (increasing the apparent correlation time). Therefore, we divided the data into segments of 200 ps and calculated the correlation time within each segment separately. The correlation time of the whole simulation was then taken as the median of the calculated correlation time for the segments.

Error analysis

To measure the quality of the free-energy estimates, we use four different estimates: the correlation coefficient between the predicted and experimental data (r^2), Pearlman's predictive index (PI),[56] the mean absolute deviation (MAD), and the mean absolute deviation from the best correlation line through the origin (MADtr; i.e. MAD after subtraction of the mean signed deviation). These measures are rather meaningless without an estimate of their statistical uncertainty. They were obtained by a simple parametric bootstrap simulation[23]: Each ligand was assigned a random normal-distributed affinity, with a mean and standard deviation obtained from the free-energy estimates. Then, r^2, PI, MAD, and MADtr were evaluated and this procedure was repeated 10 000 times. The standard deviations of these re-sampled sets are reported as the standard errors of the quality measures. Throughout this paper, all reported statistical uncertainties are standard errors of the mean, i.e. the standard deviation divided by the square root of the number of estimates.

Results

We have estimated the binding free energy of seven biotin analogues to avidin using the LIE and MM/GBSA approaches. Binding affinities obtained with MM/GBSA have already been published for these ligands.[23] However, LIE requires calculations performed on neutralised spherical systems with only one ligand, and typically also for truncated systems. Therefore, the MM/GBSA calculations were rerun with the same settings as the LIE calculations to make the results completely comparable and also to allow for a comparison of the results obtained with the two setups. Calculations were performed both on the full avidin tetramer and a system truncated to 26 Å around the ligand of interest.

Before the comparison, we must decide how we best can obtain reliable and efficient free energies with a well-defined precision. This has already been done for MM/GBSA[23] and here we perform a similar analysis for LIE. Following our previous study,[23] as well as several other investigations,[25,26] we will not employ one single long simulation, but instead many shorter independent simulations, generated by using different starting velocities. This gives a more reliable estimate of statistical precision and we can obtain any desired precision by simply employing a proper number of independent simulations, because the standard deviation of the mean decreases with the square root of the number of independent simulations included in the average. Therefore, we only need to determine the sampling frequency of the energy terms in Eqns. 1, as well as the length of the equilibration and production parts of the individual simulations.

Correlation time

For all methods that average energies over a series of MD snapshots, it is essential to ensure that the consecutive estimates are independent, i.e. that sampling is not too frequent – otherwise the statistical error will be underestimated. The correlation time of the four time series that are the basis of the LIE estimates (Eqn. 1) were calculated using the method of statistical inefficiency.[54,55] This was done for the whole 800 ps (full system) or 1600 ps (truncated system) simulations (including equilibration) and for all independent simulations of each ligand.

We soon discovered that the original method is very sensitive to the drift during the equilibration period and also to occasional bumps in the data, which often are seen in long simulations, giving strongly increased correlation times (an example is shown in Figure S1). As a consequence, the estimated correlation time always increased when the simulation time was increased. Strictly speaking, this shows that there are ns time-scale motions in the structure that may indicate that much longer equilibration and simulation times are needed to obtain truly uncorrelated data that are independent to the starting structure.[57] However, in standard LIE and MM/GBSA applications, it is assumed that simulations on a ns time-scale started from the crystal structure provide representative structures for binding-energy calculations.[8,10] Therefore, we decided to divide all simulations into segments of 200 ps, for which a separate correlation time was estimated with the original method. This gave correction times of 1–4 ps for most segments, but segments during the equilibration period and segments with bumps gave much higher values. The latter were ignored by taking the median of the segment estimates (cf. Figure S1).

With this approach, we obtained stable results for all systems, which are presented in Figure 2 as the cumulative frequency of the number of simulations (among the 80 or 20 independent simulations) that have a particular correlation time. It can be seen that the correlation time for the $\langle E_{\mathrm{vdW}}^{\mathrm{L-S}} \rangle_L$ and $\langle E_{\mathrm{ele}}^{\mathrm{L-S}} \rangle_L$ terms are always 3 ps or less. The $\langle E_{\mathrm{vdW}}^{\mathrm{L-S}} \rangle_{PL}$ term has a slightly longer correlation time, up to 5 ps, whereas the $\langle E_{\mathrm{ele}}^{\mathrm{L-S}} \rangle_{PL}$ term shows the longest correlation time, up to 18 ps for Btn5, but up to 7 ps for the other systems. The reason for the long correlation time for Btn5 is that this energy term shows a long-term oscillation

(see Figure S2). If segments of 160 or 100 ps are instead used, we obtain correlation times of 4 and 2 ps, respectively.

In our previous study, we showed that the MM/GBSA energies have a correlation time of about 5 ps at which point 90% of the data were uncorrelated without discarding any data and 98% of the data if the first 100 ps of the simulations were discarded.[23] However, this conclusion was based on only two ligands (Btn1 and Btn2) and with a somewhat different setup (e.g. octahedral systems treated with particle-mesh Ewald summation and no neutralisation). Therefore, we repeated the analysis also for MM/GBSA for all ligands. The results in Figure 2e show that the correlation time is up to 7 ps, but 90% of the data is uncorrelated already at 4 ps.

Altogether, these data indicate that the correlation time is shortest for the free-ligand simulations (2–3 ps), slightly longer for the MM/GBSA results (~4 ps), and still somewhat longer for the $\langle E_{\mathrm{ele}}^{\mathrm{L-S}} \rangle_{PL}$ term (~6 ps). However, these differences are small and somewhat dependent on the details of the method to calculate the correlation time. Therefore, we decided to use the same correlation time for both methods and also for the two types of LIE simulations, 5 ps.

Equilibration time

The next step is to determine the length of the equilibration period of the simulations, i.e. the part of the simulation that is excluded in the averages. Many methods are available to determine the equilibration time (t_{eq}).[8,23,55,58] We have tested several different variants, e.g. including block averaging or reverse cumulative averaging with two different tests for normal distribution. Unfortunately, all methods to determine t_{eq} are sensitive to details and parameters of the algorithms. At the end, we decided to use the following scheme: For each ligand, we calculated block averages of either the MM/GBSA binding energy or the LIE $\beta \langle E_{\mathrm{ele}}^{\mathrm{L-S}} \rangle + \alpha \langle E_{\mathrm{vdW}}^{\mathrm{L-S}} \rangle$ energy terms for the complex or free-ligand simulations for each 100 ps of the simulations. These averages were compared to the average over the last 400 ps (full system) or 500 ps (truncated system) of the simulation and if the difference was over 2 kJ/mol, that block was rejected. The equilibration time was taken as the end of the last set of at least two consecutive rejected block, but including also isolated rejected blocks if they are one or two blocks away from a set of consecutive rejected blocks. By such a rule, we disregard occasional isolated rejected blocks late in the simulation, because the aim of the equilibration period is to remove data with a drift at the beginning of the simulation, but not bumps later in the simulation. A minimum equilibration time of 100 ps was assumed for all simulations. Several examples of typical equilibration curves and our selection of t_{eq} are shown in Figure S3.

For the present comparison between LIE and MM/GBSA, the exact length of the equilibration period is not of prime interest, but rather whether one of the two methods has a longer equilibration time than the other. However, for the present test case, we do not see any clear tendency: The two methods show similar equilibration time for (the complex simulation) of most ligands (eight out of the 7 + 7 simulations with full and truncated protein). When this is not the case, MM/GBSA gives the shorter equilibration time for two of the simulations and LIE a shorter time for four of the simulations. Therefore, we decided to use the same equilibration time for both MM/GBSA and LIE (viz. the largest of the two individual values) to avoid that the comparison is biased by differences in the equilibration.

For the free-ligand simulations (which are relevant only for LIE), the equilibration time is normally shorter than the complex. However, for the three charged ligands, it is notable that the free-ligand simulations give a quite large variation in the LIE energies, which quite often give rise to occasional isolated large deviations in block averages (cf. Figure S3e). However, since the simulations do not show any pronounced trends, only rejected blocks at

the beginning of the simulation were omitted (in accordance to the rule given above).

The selected equilibration times are collected in Table S1. In variance to our previous investigation of MM/GBSA[23] we allow for different equilibration times for different ligands, which is more realistic and economic. It can be seen that the equilibration times vary from the requested minimum of 100 ps up to 1000 ps for Btn3. During this process, we decided to increase the simulation time of the truncated systems from 800 to 1600 ps. For the full systems, this could not be afforded (remember that 80 + 20 independent simulations were run for each ligand, giving a total simulation time of 1.12 μs). It is notable that the equilibration times are longer than in our previous MM/GBSA investigation, in which all simulations were judged to be converged after 100 ps.[23] A typical example is shown in Figure S4.

Length of production simulation and efficiency

We have now determined the correlation and equilibration times for our simulations of the seven ligands. What remains is to determine the length of the production simulation. This is somewhat involved, because we run also a number of independent simulations for each ligand. Therefore, we can improve the precision either by elongating the production time of each independent simulation or by increasing the number of independent simulations. The latter is more effective because the standard error of the final estimate (average over the independent simulations) will decrease with the square root of the number of independent simulations, whereas the dependence on the production time is less clear, since the results are not fully independent (this is the reason why we use several independent simulations[23]). On the other hand, the independent simulations cost more, because an initial equilibration has to be run. Therefore, to reach an optimum distribution, we need to consider also the computational cost of the simulations and energy calculations (which of course depend on the simulated system and the computer equipment).

We will follow our previous suggestion to optimise the CPU time required to reach a certain precision, e.g. s_{av} = 1 kJ/mol[23] (but the comparison of the two methods will not depend on this limit). We can estimate the standard error of each independent simulation, s_{simu}, with a certain number of production snapshots, n_{prod}, from our available data. The number of independent simulations (n_{av}) needed to reach the desired precision is then simply

$$n_{av} = \frac{s_{simu}^2}{s_{av}^2} \tag{6}$$

With these data, we can then calculate the required total CPU time: A 50 ps MD simulation takes 8 and 0.6 CPU hours on a single 3.0 GHz Intel Xenon processor for the full and truncated systems, respectively. The MM/GBSA post-processing energies take ~0.25 CPU hours, irrespectively whether full or truncated systems are used (because the time is dominated by the entropy calculations, which are performed on truncated systems in both cases), whereas the LIE energies can be calculated without any overhead. Therefore, the time consumption for LIE is

$$\mathrm{CPU}_{\mathrm{LIE}} = n_{av}\left(t_{eq}\, c_{MD} + (n_{prod}-1) f\, c_{MD}\right) \tag{7}$$

and for MM/GBSA

$$\mathrm{CPU}_{\mathrm{MM/GBSA}} = n_{av}\left(t_{eq}\, c_{MD} + (n_{prod}-1) f\, c_{MD} + n_{prod}\, c_{ene}\right) \tag{8}$$

where f is the sampling frequency (so that (n_{prod} – 1)*f is the length of the production

simulation), c_{MD} is the cost of running the MD simulation and c_{ene} is the cost of doing a single MM/GBSA energy calculation. We can now calculate the CPU consumption as a function of n_{prod} using the equilibration times and the sampling frequency determined in the previous section, as well as s_{simu} obtained from the simulations. As can be seen in Figure S5, the CPU shows a minimum when n_{prod} is varied, because the first term in Eqns. 7 and 8 depends on n_{av}, which decreases as n_{prod} is increased, whereas the other terms depend on n_{av} * n_{prod}, which increases with n_{prod}. This is the optimum value of n_{prod}. Strictly speaking, the results depend on the equilibration time (ligands with long equilibration times prefer somewhat longer production times and therefore fewer independent simulations), but the dependence is rather week. Moreover, for some ligands it is also favourable to increase the sampling frequency. Therefore, we have optimised n_{prod} and f separately for each ligand (Table S2). However, the results are quite similar if we average s_{simu} and the CPU time over all seven ligand (and then f = 5 ps is optimal; cf. Figure S5b). These averaged results are given in Table 1.

It can be seen that for the complex simulation of LIE, it is most efficient to run short production simulations, ~50 ps, and instead run many independent simulations (91–135). For the free ligand, it is more efficient to run longer simulations (~300 ps) and fewer independent simulations (7–10). The reason for this is that the standard error of the free-ligand simulations is appreciably smaller for the complex simulations and that it decreases more with the number of snapshots. However, in both cases, the simulation time is shorter than typically is used with LIE, illustrating that long simulations underestimate the statistical uncertainty. On the other hand, the total simulation time, 7.4 + 2.3 ns for the truncated systems and 4.6 + 3.0 ns for the full systems, plus 10–140 ns equilibration, is much longer than normally used with LIE. The total LIE CPU times are 1050 and 3900 CPU hours for the truncated and full systems, respectively.

Looking at MM/GBSA instead, we see that also for this method, rather short production simulations are most efficient, 25 and 45 ps for the truncated and full systems, respectively. This amounts to a total time of 3135 and 5983 CPU hours, respectively. Thus, this analysis indicates that LIE is more efficient than MM/GBSA, by a factor 3 for the truncated systems and 1.5 for the full systems. Looking at the more detailed data in Table S2, it can be seen that for the truncated system, MM/GBSA requires between 1.7 and 6.7 times more CPU than LIE for the truncated system, whereas for the full system the ratios are 1.0–2.4.

From Eqns. 7 and 8, it can be seen that the CPU consumption depends on three terms, two of which are common to both methods, whereas the last one, the cost of the energy calculations, only applies to MM/GBSA. This inherent difference between the two methods will always favour LIE and also lead to that MM/GBSA typically prefers a slightly lower n_{prod} than LIE. However, the importance of this difference decreases with the size of the system, because for the truncated system the third term is 4–8 times larger than the second term, whereas for the full system, it is only 30–60% of the second term.

On the other hand, this effect is counteracted by the fact that LIE requires simulations of both the free ligand and the complex, whereas MM/GBSA is based only on the simulations of the complex. If everything else were equal, this would compensate for the extra cost of the energy calculations and MM/GBSA would always be preferable.

However, a third factor is also important, viz. the standard errors (s_{simu}) of the various energies and their dependence on the number of snapshots, which will affect n_{av} and the optimum n_{prod}. As we will see below, there is little difference in the standard error of the MM/GBSA and LIE estimates of the binding energy from the simulations of the complex (although the various ligands show a rather large variation). However, the standard error for the free-ligand simulations are appreciably smaller than for the complex simulations and also shows a larger decrease with n_{prod} as can be seen in Figure S6. This leads to a lower total

number of required snapshots (n_{av} * n_{prod},) for the free ligand than for the complex (469 compared to 1620 for the truncated system and 600 compared to 1001 for the full system). The combination of these three factors give the net efficiency illustrated in Table 1.

Affinity estimates

Next, we estimated the binding free energy for the seven biotin analogous using the optimum equilibration time and sampling frequency, but using all available snapshots for production (because these are the best estimates we can obtain with available data). For LIE, we used β = 0.5 for the charged ligands (Btn1–Btn3) and 0.43 for the neutral ligands (Btn4–Btn7), and α = 0.18 for all ligands. These values are usually denoted the standard parametrisation.[13,14] The binding free energies for the truncated and full systems are shown in Tables 2 and 3, respectively. It can be seen that the results with the standard parametrisation are poor with negative predictive indices and negative correlation coefficients (although r^2 is positive by definition). This is because LIE predicts a higher affinity for the neutral ligands than for the charged ones.

Kollman and co-workers have studied a similar set of biotin analogues with the LIE method, but they obtained a reasonable, positive correlation.[19] The reason for this is that they used a special value for α (1.0), fitted to the experimental data. Using the standard parametrisation for their data (available for all our biotin analogues, except Btn3), we obtain a negative correlation and r^2 = 0.01. Reported electrostatic and van der Waals energies are rather similar to ours with correlation coefficients (r^2) of 0.57 and 0.96 (Table S3). The rather large difference in the electrostatic energy is probably caused by differences in the simulation setup: They employed smaller systems and neutralised only a minimum amount of titrable residues.[19]

Therefore, we also tried to optimise the α value, keeping β at the default values. Since we use four different quality estimates (MAD, MADtr, r^2, and PI), we fitted α to optimise each of these measures by varying α from –5 to +5 with increments of 0.05. The results are shown in Table 4 for the truncated system (the full system gave similar results). Optimising α against r^2 gave unrealistic binding affinities because the correlation coefficient benefits from large energy differences, which are obtained when the energies are scaled up. r^2 converges asymptotically at $\alpha > 20$, but as can be seen in Table 4, both MAD and MADtr are poor already at $\alpha = 5$.

Optimising α against PI gave a non-smooth dependence on α, although it gave similar results as when optimising against MADtr. On the other hand, fitting α against MADtr and MAD gave a smooth, parabolic dependence on α. The fits gave optimal values of α = 0.70 and 1.15, respectively. For these two α values, PI and MADtr differ by only 0.05 and 0.3, respectively, which probably are not statistically significant. On the other hand, MAD and r^2 differ significantly. We consider it more important to obtain good relative estimates (high r^2) than absolute estimates (low MAD) and we therefore prefer the α value obtained by fitting to MADtr, 1.15.

The LIE results using α = 1.15 are included in Tables 2 and 3 (column LIE (opt)), and are plotted in Figure 3. It can be seen that the largest error is found for Btn4. Kollman et al. also had problem with this ligand. They argued that the error comes from the fact that the protein needs to be reorganised to accommodate the ester group of Btn4 and that such a term is missing in the LIE approach.[19]

Finally, we also considered the correction to the binding affinities of the charged ligands from the omitted surface charges, ΔG_{cc}. For the truncated systems this amounts to ~2 kJ/mol, irrespectively of ligand and for the full systems it amounts to ~7 kJ/mol. The largest contribution from a residue is ~2 kJ/mol, showing that neutralisation do not affect

individual interactions so much. However, since avidin contains almost 100 charged surface residues, the sum is significant for the full systems. The effect of adding ΔG_{cc} is shown in Table 2 for the truncated system, and as the correction is small, it hardly affects the result at all. However, for the full system the correlation coefficient increases to 0.50 and MADtr decreases to 14 kJ/mol (see Table 3).

The MM/GBSA estimates of binding free energies are also included in Tables 2 and 3, and are plotted in Figure 3. For the truncated system, MADtr is 16 kJ/mol, $r^2 = 0.76$, and PI = 0.95. The full-system estimates are slightly better (see Table 3). Comparing with LIE, the difference in MADtr is not statistically significant, but the better results for the correlation coefficient and the PI are statistically significant. This is because MM/GBSA do not have any problems with Btn4. It is notable that the LIE and MM/GBSA results are quite similar for the three charged ligands (differences less than 6 kJ/mol), whereas for the neutral systems, the difference is appreciably larger (27–35 kJ/mol), MM/GBSA always giving a less favourable binding.

The binding affinities obtained with the truncated and full systems are quite similar. For LIE, they differ by up to 12 kJ/mol, the results of the full system almost always being more positive. For MM/GBSA, the results differ by 1–10 kJ/mol in a less systematic way.

Compared to our previous MM/GBSA binding affinities,[23] the difference is 1–19 kJ/mol (for the full system), i.e. much larger than the statistical uncertainty. The difference is systematic in that the negative ligands give more negative affinities (by 5–12 kJ/mol) in the present calculations, whereas the neutral ligands give more positive affinities (by 1–19 kJ/mol). As a consequence, the present simulations give a similar MADtr (16 compared to 15 kJ/mol), but improved r^2 (0.79 compared to 0.59) and PI (1.00 compared to 0.85). This shows that the MM/GBSA results depend much stronger on the water model (TIP3P or TIP4P-Ewald), the treatment of long-range electrostatics (a spherical system with reaction-field corrections or a octahedral system with Ewald summation), and the treatment of surface charges (neutralisation or not) than on the placement of the explicit water molecules, the initial velocities, and the protonation and rotation of residues, which gave variations of less than 1 kJ/mol for Btn1 in a previous investigation.[24] We have previously compared the results of spherical vs. periodic simulations and neutralised systems with MM/PBSA, but the precision was too low to discern differences of relevant sizes.[30]

Precision

The statistical precision of the free-energy estimates is also shown in Tables 2 and 3 (standard deviation of the mean). It can be seen that LIE and MM/GBSA give similar uncertainties, 1–3 kJ/mol. For the truncated system, LIE with fitted α and charge corrections gives a smaller uncertainty than MM/GBSA for five ligands. In general, the charged ligands show a slightly larger uncertainty than the neutral ligands. For the full systems, the precision is often slightly better (1–2 kJ/mol) and in most cases, MM/GBSA has a lower uncertainty than LIE (not for Btn3 and Btn5). This better precision of the full system is unexpected considering that the simulations are only half as long (0.8 ns compared to 1.6 ns). This shows that the intrinsic standard deviation of the data is larger for the truncated system than for the full system.

It should be noted that the LIE terms are scaled by the parameters α and β. Without this scaling the uncertainty of the LIE energies are larger than that of the MM/PBSA energies (and the two methods would become more equal in efficiency). The fact that LIE with α =1.15 gives only a slightly larger uncertainty than with α = 0.18 (by 0.1–0.3 kJ/mol), although the van der Waals term is scaled up by a factor of $1.15/0.18 \approx 6$, shows that the precision of LIE is strongly dominated by the electrostatic term. This is also the reason why the optimisation of the LIE procedure does not need to be redone with the optimised α value.

It is somewhat disappointing that even with 80 + 20 independent simulations of 1.6 ns

length, we have not been able to reach a precision of 1 kJ/mol for four of the ligands. In fact, for Btn1, which has the poorest precision, this would require 390 simulations or 620 ns simulation time. It is also notable that the precision of the MM/GBSA results are worse than in our previous investigation,[23] for which a precision of 1 kJ/mol was reached for all seven ligands with 20–50 300 ps simulations of the complex. This emphasizes a specific shortcoming of LIE for this tetrameric protein: With LIE, we can only simulate one ligand at the time, whereas for MM/GBSA, we could obtain four affinity estimates from the simulation of the complex with four ligands. We employed that opportunity in the previous study, but in this study, we used the same simulations for both LIE and MM/GBSA. Moreover, it seems that the convergence of the MM/GBSA energy terms is slower for a neutralised and truncated protein.

Affinities of individual subunits

We have previously shown that MM/GBSA estimates employing several independent simulations gave identical affinities for the four subunits in avidin within statistical uncertainty.[23] It is of interest to see if the same holds for also LIE. The binding free energies for the four subunits are shown in Tables 5 and 6, for the truncated and full systems respectively, using α =1.15. It is evident that the four subunits do not give the same binding affinities for the truncated systems: The four subunits give results that differ by up to 21–36 kJ/mol for the charged ligands and by 5–8 kJ/mol for the neutral ones. This is much more than expected from the standard errors of the estimates, 2–4 and 1–2 kJ/mol, respectively. On the other hand, subunits B and D give binding affinities that are the same within statistical uncertainty (the difference is less than 3 kJ/mol) and the same applies to subunits A and C, although the difference is up to 5 kJ/mol. This indicates that the differences are caused by differences in the subunits, probably the fact that subunits A and C have one less amino-terminal residue than subunits B and D in the crystal structure.

Surprisingly, for the full systems, the differences between the subunits are smaller, 2–9 kJ/mol, with no difference between the charged and neutral ligands. In this case, the differences are statistically significant only for Btn5–Btn7. Subunits B and D still give very similar results (within 2 kJ/mol), whereas the differences for subunits A and C are larger, up to 6 kJ/mol for Btn2 (but they are not statistical significant at the 95% level).

These large differences between the subunits for LIE led us to check also the MM/GBSA results. From Tables 6 and 7, it can be seen that MM/GBSA actually give similar differences between the subunits as LIE, 2–28 kJ/mol differences for the truncated systems and 2–18 kJ/mol for the full systems. This is a surprising difference compared to our previous results,[23] which most likely is caused by differences in the setup of the two sets of calculations, in particular the neutralisation of surface charges, which may emphasize the one-residue difference between the subunits. Moreover, it is clear that the differences are amplified by the truncation of the protein. This is an important issue that will be the subject of future investigations of other proteins.

Conclusions

In this study, we have designed a simulation protocol for the LIE method to reach a certain level of statistical precision in the predicted affinities, in the same way as in a previous study with MM/GBSA.[23] Our results indicate that for this biotin–avidin system, a sampling frequency of ~5 ps and equilibration times of 0.1–1.0 ns are appropriate. By optimising the CPU time, we also suggest that rather short simulations should be used for the complex (50 ps after equilibration), but longer for the free ligand, ~300 ps. Then, a proper number of independent simulations should be run until the desired precision is obtained. The sampling frequency and equilibration times are probably similar for most systems, whereas the length

of the production simulation may depend on the simulated system and the computational equipment. Considering the long equilibration times observed for some systems, it would probably have been better to first run one long equilibration of each avidin–ligand complex (1–5 ps), before the independent simulations were started by using different sets of starting velocities. Then, the equilibration for the individual simulations could most likely have been reduced, as was observed for galectin-3 with MM/GBSA.[23]

In parallel, similar calculations have been performed with the MM/GBSA approach, based on the same simulations. This allows us to compare the efficiency of these two popular end-point methods. We have reached several interesting conclusions:

- The correlation time of the LIE and MM/GBSA energies is similar.
- The equilibration time varies heavily with the ligand, but the two methods seem to require similar equilibration times.
- In general, LIE seems to be more efficient than MM/PBSA by a factor of 2–7 for the truncated systems, but by a factor of 1.0–2.4 for the full system (i.e. it gives the same statistical precision with a computational effort that is lower by these factors). The lower efficiency of MM/GBSA comes from the extra time required for the entropy calculation, which more than compensates for the fact that LIE requires an extra simulation (of the free ligand).
- On the other hand, in variance to MM/GBSA, LIE contains one empirical parameter, α. If the standard value (α = 0.18) is used, LIE gives very poor results for this test case, with negative correlation and PI. However, if α is fitted, LIE and MM/GBSA give similar MADtr, ~16 kJ/mol, although MM/GBSA still outperforms LIE for r^2 and PI. This is mainly due to LIE problems with a single ligand, Btn4.
- LIE is more restrictive in the setup of the simulation: It requires that the size of the simulated systems is the same for the complex and the free ligand and also that the protein is neutralised in the simulations. Our results indicate that this neutralisation may slow the convergence and make the result different for the four subunits in avidin.
- Moreover, LIE simulations are typically performed on truncated systems with a radius of ~25 Å.[8] Our results indicate that such a truncation may also slow the convergence and emphasize differences between the subunits.
- The change of the water model, the treatment of long-range electrostatics, and the neutralisation of the protein have a quite large effect on the MM/GBSA binding energies (1–19 kJ/mol), much larger than the initial solvation, the starting velocities, as well as the protonation and rotation of residues.[23] It remains to show with larger test sets which of these setups is preferred, but for the present systems, the current setup gives a somewhat better correlation and PI compared to the experimental results.

Thus, we can conclude that LIE is inherently more effective than MM/GBSA (giving a certain precision at a smaller expense in computation time), at least for the present test case. Considering that the avidin tetramer is rather large and that LIE is typically run on truncated proteins, it is likely that this conclusion is valid also for other proteins, although more tests are required to confirm this. However, if the entropy term in MM/GBSA is ignored, as have been done in many studies,[5,59,60]MM/GBSA is expected to become the more effective method. This might be an interesting alternative for MM/GBSA, especially as the entropy term has been criticized[61] and it limits both the precision and the CPU consumption.

On the other hand, we have seen that LIE depends on an empirical parameter and that it has more restrictions on the setup of the calculations. Clearly, MM/GBSA is disfavoured by the LIE setup used in this study, giving a slower convergence, and it may be more effective with a more typical MM/GBSA setup. In particular, MM/GBSA may obtain four energy estimates from each snapshot for this tetrameric protein. Even more seriously, it is clear that the setup of the calculations quite strongly affects the results. It remains to be shown on much

larger and more diverse test sets which of the setups are more realistic and which of the MM/GBSA and the LIE methods give the more accurate results.

Acknowledgements
We thank Johan Åqvist and coworkers for help with LIE and the Q software package. This investigation has been supported by grants from the Swedish research council (project 2010-5025) and from the Research school in pharmaceutical science. It has also been supported by computer resources of Lunarc at Lund University, C3SE at Chalmers University of Technology and HPC2N at Umeå University.

Supporting Information Available: Equilibration times and optimum f, n_{prod}, and n_{av} estimates for the various simulations, comparison with the results in ref. 19, times series for two representative simulations, examples of the selection of the equilibration time, as well as the dependence of the CPU time and s_{simu} on n_{prod}. This material is available free of charge via the Internet at http://pubs.acs.org.

References

1 Michel, J.; Essex, J. W. *J. Comput. Aided Mol. Des.* **2010**, *24*, 639-658
2 Shirts, M.; Mobley, D. L.; Chodera, J. D. *Annu. Rep. Comput. Chem.* **2007**, *3*, 41-59
3 Gohlke, H.; Klebe, G. *Angew. Chem. Int.* **2002**, *41*, 2644-2676.
4 Steinbrecher, T.; Labahn, A. *Curr. Med. Chem.* **2010**, *17*, 767-785
5 Foloppe N.; Hubbard, R. *Curr. Med. Chem.* **2006**, *13*, 3583-3608
6 Åqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Eng.* **1994**, *7*, 385-391
7 Hansson, T.; Marelius, J.; Åqvist, J. *J. Comput. Aided Mol. Design* **1998**, *12*, 27-35
8 Brandsdal B. O.; Östberg F. M.; Almlöf M.; Feierberg I.; Luzhkov V. B.; Åqvist, J. *Adv. Prot. Chem.* **2003**, *66*, 123-158
9 Srinivasan, J.; Cheatham III, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *37*, 9401-9409
10 Kollman, P. A.; 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, D. A.; Cheatham III, T. E. *Acc. Chem. Res.* **2000**, *33*, 889-897
11 Åqvist, J.; Hansson, T. *J. Phys. Chem.* **1996**, *100*, 9512-9521
12 Almlöf, M.; Carlsson, J.; Åqvist, J. *J. Chem. Theory Comput.* **2007**, *3*, 2162–2175
13 Almlöf, M.; Brandsdal, B. O.; Åqvist, J. *J. Comput. Chem.* **2004**, *25*, 1242-1254
14 Carlsson, J.; Boukharta, L.; Åqvist, J. *J. Med. Chem.* **2008**, *51*, 2648-2656
15 Carlson, H. A.; Jorgensen, W. L. *J. Phys. Chem.* **1995**, *99*, 10667–10673
16 Swanson, J. M. J.; Henchman, R. H.; McCammon, J. A. *Biophys J.* **2004**, *86*, 67-74
17 Genheden, S.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Ryde U. *J. Phys. Chem. B* **2010**, *114*, 8505–8516
18 Barril, X.; Gelpi, J. L.; Lopez, J. M.; Orozco, M.; Luque, F. *J. Theor. Chem. Acc.* **2001**, *106*, 2-9
19 Kollman, P. A.; Kuhn, B. *J. Med. Chem.* **2002**, *43*, 3786-3791
20 Hou T.; Guo S.; Xu, X. *J. Phys. Chem. B* **2002**, *106*, 5527-5535
21 Salvalagli M, Zamolo L.; Busini, V.; Moscatelli, D.; Cavallotti, C. *J. Chrom. A* **2009**, *1216*, 8678-8686
22 Wang, J.; Dixon, R.; Kollman, P. A. *Prot. Struct. Funct. Genet.* **1999**, *34*, 69-81
23 Genheden, S.; Ryde, U. *J. Comput. Chem.* **2010**, *31*, 837–846
24 Genheden, S.; Ryde, U. *J. Comput. Chem.* **2011**, *32*, 187–195
25 Lawrenz, M.; Baron, R.; McCammon, J. A. *J. Chem. Theory Comput.* **2009**, *5*, 1106-1116

26 Fujitani, H.; Tanidal, Y.; Ito, M.; Jayachandran, G.; Snow, C. D.; Shirts, M. R.; Sorin, E. J.; Pande, V. S. *J. Chem. Phys.* **2005**, *123*, 084108-084112
27 Genheden, S.; Diehl, C.; Akke, M.; Ryde U. *J. Chem. Theory Comput.* **2010**, *6*, 2176-2190
28 Miyamoto, S.; Kollman, P. A. *Proteins: Struct., Func., Bioinf.* **1993**, *16*, 226–245
29 Kuhn, B.; Gerber, P.; Schulz-Gasch, T.; Stahl, M. *J. Med. Chem.* **2005**, *48*, 4040–4048
30 Weis, A.; Katebzadeh, K.; Söderhjelm, P.; Nilsson, I.; Ryde, U. *J. Med. Chem.* **2006**, *49*, 6596–6606
31 Brown, S. P.; Muchmore, S. W. *J. Chem. Inf. Model.* **2006**, *46*, 1493
32 Kongsted, J.; Ryde, U. .*; Comput. Aided Mol. Des.* **2009**, *23*, 63-71
33 Söderhjelm, P.; Kongsted, J.; Ryde, U. *J. Chem. Theory Comput.* **2010**, *6*, 1726-1737
34 Pugliese, L.; Coda, A.; Malcovati, M.; Bolognesi, M. *J. Mol. Biol.* **1993**, *231*, 698-710
35 Green, N. M. *Biochem. J.* **1966**, *101*, 774
36 Green, N. M. *Adv. Prot. Chem.* **1975**, *29*, 85-133
37 Green, N. M. *Methods Enzymol.* **1999**, *184*, 51-67
38 Åqvist, J. *J. Comput. Chem.* **1996**, *17*, 1587-1597
39 Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct. Funct., Bioinf.* **2006**, *65*, 712-725
40 Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197
41 Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269-10280
42 Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431-439
43 Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impley, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926-935
44 Marelius, J.; Kolmodin, K.; Feierberg, I.; Åqvist, J. *J. Mol. Graph. Model.* **1998**, *16*, 213-225
45 Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327-341
46 Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690
47 Lee, F. S.; Warshel, A. *J. Chem. Phys.* **1992**, *97*, 3100–3107
48 King, G.; Warshel, A. A. *J. Chem. Phys.* **1989**, *91*, 3647–3661
49 Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R.; C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K.; F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *Amber 10*, University of California, San Francisco, **2008**.
50 Brandsdal, B. O.; Smalås, A. O.; Åqvist, J. *FEBS Lett.* **2001**, *49*, 171-175
51 Schutz, C. N.; Warshel, A *Proteins* **2001**, *44*, 400-417.
52 Onufriev, A.; Bashford, D.; Case, D. A. *Proteins* **2004**, *55*, 383-394
53 Kuhn, B.; Kollman, P. A. *J. Med. Chem.* **2000**, *43*, 3786-3791
54 Bishop, M.; Frinks, S. *J. Chem. Phys.* **1987**, *87*, 3675-3678
55 Yang, W.; Bitetti-Putzer, R.; Karplus, M. *J. Chem. Phys.* **2004**, *120*, 2618-2629
56 Pearlman, D. A.; Charifson, P. S. *J. Med. Chem.* **2001**, *44*, 3417-3423
57 Zuckerman, D. M. *Annu. Rev. Biophys.* **2011**, *40*, 41-62.
58 Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids.* Clarendon Press, Oxford, **1991**.
59 Hayes, J. M.; Skamnaki, V. T.; Archontis, G.; Lamprakis, C.; Sarrou, J.; Bischler, N.; Skaltsounis, A-L.; Zographos, S. E.; Oikonomakos, N. H. *Proteins* **2011**, *79*, 704-719.

60 Hou, T.; Wang, J.; Li, Y.; Wang, W. *J. Chem. Inf. Model.* **2011**, *51*, 69-82
61 Singh N.; Warshel, A. *Proteins* **2010**, *78*, 1705-1723

Table 1. Optimum estimates of n_{prod} and n_{av}, together with the corresponding s_{av} (kJ/mol) and CPU (h) estimates for LIE and MM/GBSA, following the procedure described in the text (Eqns. 6–8), using f = 5 ps and the t_{eq} values listed in Table S1, and averaging s_{simu} and CPU over all seven ligands. n_{av} and CPU are calculated for a desired precision of (s_{av} = 1 kJ/mol).

System	Method	n_{prod}	n_{av}	s_{simu}	CPU
Truncated	MM/GBSA	6	366	19.1	3135
	LIE PL	12	135	11.6	1000
	LIE L	67	7	2.6	54
	LIE total		142		1053
Full	MM/GBSA	10	168	13.0	5983
	LIE PL	11	91	9.6	3218
	LIE L	60	10	3.1	715
	LIE total		101		3933

Table 2. Binding free energies for the various methods on the truncated systems in kJ/mol. A negative r^2 indicates that r is negative.

	LIE		LIE (ΔG_{cc})		LIE (opt)		LIE (ΔG_{cc},opt)		MM/GBSA	
α	0.18		0.18		1.15		1.15			
With ΔG_{cc}	no		yes		no		yes			
Btn1	-9.7	±2.0	-11.7	±2.0	-117.1	±2.2	-119.0	±2.2	-125.2	±2.3
Btn2	2.4	±2.0	0.3	±2.0	-107.9	±2.2	-109.9	±2.2	-105.3	±2.7
Btn3	-5.7	±1.6	-7.7	±1.6	-105.8	±1.7	-107.8	±1.7	-111.4	±1.8
Btn4	-19.7	±1.0	-19.7	±1.0	-133.2	±1.3	-133.2	±1.3	-98.1	±1.1
Btn5	-10.9	±0.7	-10.9	±0.7	-86.6	±0.9	-86.6	±0.9	-54.2	±2.3
Btn6	-16.4	±0.5	-16.4	±0.5	-86.0	±0.7	-86.0	±0.7	-58.6	±1.3
Btn7	-11.7	±0.7	-11.7	±0.7	-45.0	±0.7	-45.0	±0.7	-14.2	±0.7
MAD	34.7	±0.5	34.7	±0.5	52.4	±0.6	53.3	±0.6	37.4	±0.7
MADtr	24.8	±0.6	24.8	±0.6	16.2	±0.5	15.7	±0.4	16.2	±0.6
r^2	0.27	±0.08	0.27	±0.08	0.38	±0.02	0.41	±0.02	0.76	±0.01
PI	-0.70	±0.15	-0.70	±0.15	0.69	±0.02	0.69	±0.02	0.95	±0.01

Table 3. Binding free energies using various methods on the full systems in kJ/mol. A negative r^2 indicates that r is negative.

	LIE		LIE (ΔG_{cc})		LIE (opt)		LIE (ΔG_{cc},opt)		MM/GBSA	
α	0.18		0.18		1.15		1.15			
With ΔG_{cc}	no		yes		no		yes			
Btn1	-2.5	±1.5	-9.5	±1.5	-109.3	±1.7	-116.4	±1.7	-123.5	±1.4
Btn2	9.3	±1.4	2.1	±1.4	-103.6	±1.6	-110.7	±1.6	-114.4	±1.4
Btn3	2.5	±1.1	-4.2	±1.1	-94.2	±1.3	-100.9	±1.3	-106.3	±1.5
Btn4	-14.2	±1.3	-14.2	±1.3	-121.0	±1.6	-121.0	±1.6	-93.3	±1.1
Btn5	-9.7	±0.9	-9.7	±0.9	-82.7	±1.1	-82.7	±1.1	-56.2	±1.2
Btn6	-15.0	±0.7	-15.0	±0.7	-81.3	±0.9	-81.3	±0.9	-53.6	±0.7
Btn7	-13.0	±0.8	-13.0	±0.8	-43.0	±0.9	-43.0	±0.9	-13.5	±0.8
MAD	38.9	±0.4	35.9	±0.4	45.8	±0.5	48.8	±0.6	36.7	±0.5
MADtr	27.6	±0.4	24.2	±0.4	15.9	±0.5	14.1	±0.5	16.1	±0.4
r^2	0.53	±0.06	0.33	±0.08	0.39	±0.02	0.50	±0.03	0.79	±0.01
PI	-0.75	±0.04	-0.75	±0.11	0.69	±0.02	0.69	±0.04	1.00	±0.01

Table 4. Results for LIE obtained after optimising the α parameter with respect to the four quality measures MAD, MADtr, r^2, and PI.

Optimised measure	α	MAD	MADtr	r^2	PI
MAD	0.70	17.9	16.5	0.23	0.64
MADtr	1.15	52.4	16.2	0.38	0.69
r^2	5.00	377.0	91.9	0.50	0.61
PI	1.10	35.3	16.2	0.34	0.69
Kollman et al.[24]	1.00	35.3	16.3	0.34	0.65

Table 5. Binding free energies (kJ/mol) for each subunit of avidin in the truncated simulations. For LIE, α = 1.20 was used.

	LIE				MM/GBSA			
	A	B	C	D	A	B	C	D
Btn1	-132.1 ±2.3	-99.6 ±3.0	-135.8 ±3.1	-100.8 ±2.7	-128.9 ±3.2	-112.8 ±3.5	-136.2 ±4.0	-120.6 ±2.8
Btn2	-121.9 ±3.4	-92.5 ±3.8	-124.6 ±3.2	-92.5 ±2.6	-107.0 ±2.0	-86.0 ±2.7	-104.4 ±1.1	-96.9 ±2.3
Btn3	-116.1 ±2.7	-95.6 ±2.4	-115.6 ±3.0	-96.0 ±2.3	-89.8 ±4.7	-115.3 ±2.4	-97.3 ±5.7	-110.6 ±2.0
Btn4	-137.0 ±2.5	-133.5 ±1.9	-131.9 ±2.4	-130.3 ±2.1	-99.8 ±1.4	-98.1 ±1.6	-99.5 ±1.6	-100.3 ±1.5
Btn5	-83.5 ±1.8	-89.3 ±1.4	-83.0 ±1.5	-90.6 ±1.4	-45.5 ±2.7	-66.6 ±2.7	-39.1 ±3.7	-63.4 ±1.7
Btn6	-86.3 ±1.1	-86.7 ±1.2	-82.9 ±1.6	-88.1 ±1.1	-58.6 ±3.0	-58.9 ±1.3	-52.7 ±2.0	-56.6 ±1.6
Btn7	-41.5 ±0.8	-48.1 ±0.6	-42.2 ±1.0	-48.0 ±0.6	-9.1 ±1.0	-21.3 ±0.7	-10.2 ±0.8	-20.5 ±0.8

Table 6. Binding free energies (kJ/mol) for each subunit of avidin in the simulations of the full system. For LIE, α = 1.15 was used.

	LIE				MM/GBSA			
	A	B	C	D	A	B	C	D
Btn1	-108.8 ±2.9	-109.2 ±1.8	-109.1 ±2.9	-110.3 ±2.4	-116.8 ±3.4	-122.3 ±1.2	-121.5 ±2.7	-118.2 ±2.8
Btn2	-106.8 ±3.0	-103.5 ±1.9	-101.2 ±2.9	-102.7 ±2.6	-119.0 ±1.6	-112.4 ±0.7	-105.1 ±1.6	-101.5 ±1.6
Btn3	-95.0 ±3.1	-95.9 ±1.4	-91.2 ±2.6	-94.7 ±1.4	-100.4 ±4.2	-102.6 ±1.1	-99.5 ±3.2	-103.2 ±0.7
Btn4	-117.9 ±2.6	-124.4 ±1.9	-118.1 ±2.6	-124.2 ±1.9	-92.3 ±2.2	-91.9 ±2.0	-93.7 ±2.0	-93.7 ±2.1
Btn5	-77.9 ±1.6	-87.0 ±1.4	-80.0 ±1.8	-86.3 ±1.3	-51.0 ±2.6	-61.2 ±1.4	-48.3 ±1.9	-58.6 ±1.8
Btn6	-78.9 ±1.4	-83.1 ±0.9	-77.6 ±1.3	-85.1 ±1.2	-49.9 ±1.1	-56.0 ±0.8	-49.4 ±1.4	-58.2 ±0.9
Btn7	-39.6 ±1.8	-46.8 ±0.5	-37.9 ±1.5	-47.0 ±0.4	-9.4 ±1.2	-17.4 ±0.8	-7.7 ±1.3	-19.2 ±0.4

Figure 1. The seven biotin analogous studied, a) biotin (Btn1), b) to g) Btn2–Btn7. The numbers shown are experimental affinities in kJ/mol.[36]

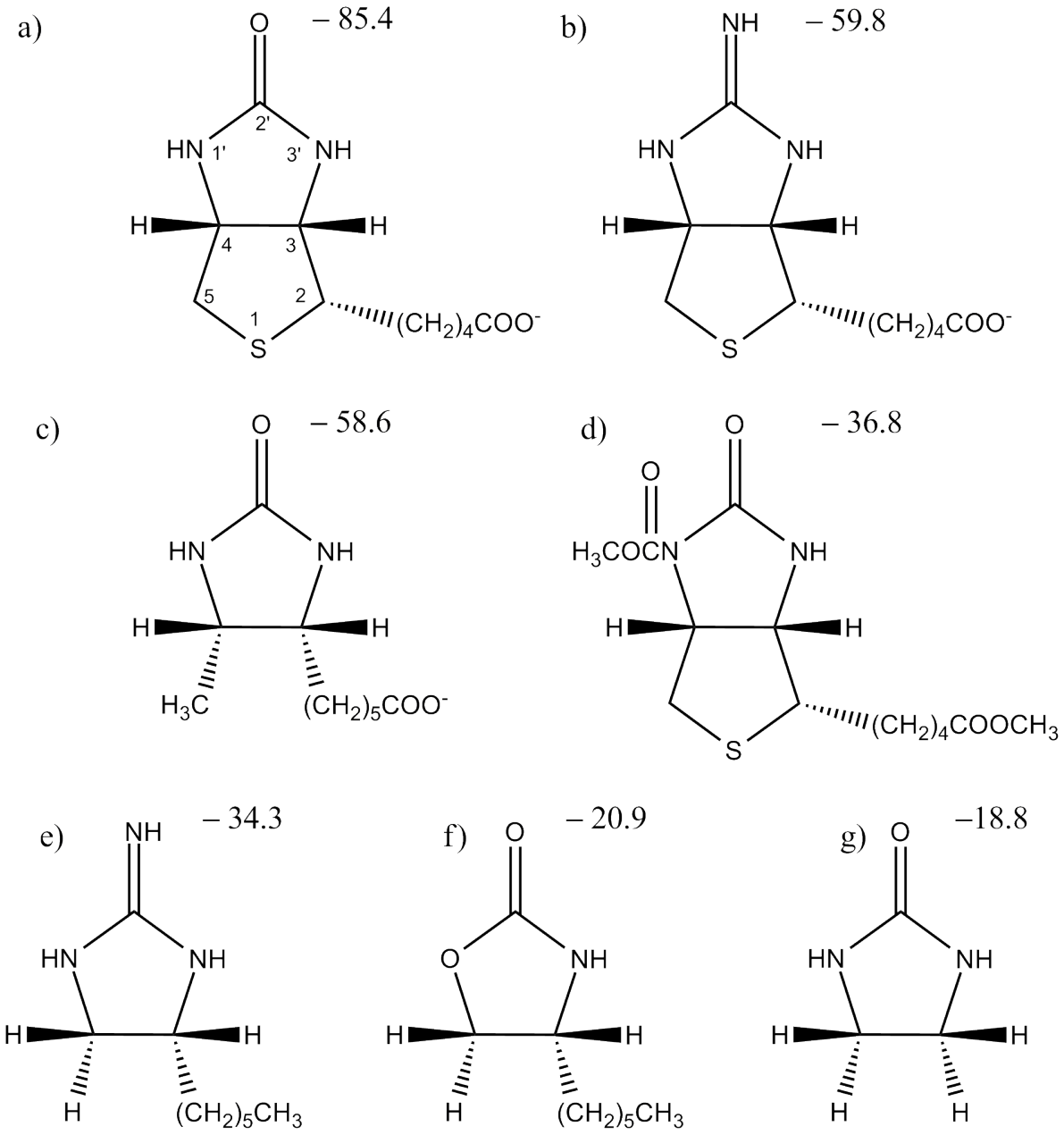

Figure 2. Correlation time of interaction energies of the truncated systems (the results for the full systems are similar), a) $\langle E_{\mathrm{vdW}}^{\mathrm{L\text{-}S}} \rangle_{L}$, b) $\langle E_{\mathrm{vdW}}^{\mathrm{L\text{-}S}} \rangle_{PL}$, c) $\langle E_{\mathrm{ele}}^{\mathrm{L\text{-}S}} \rangle_{L}$, d) $\langle E_{\mathrm{ele}}^{\mathrm{L\text{-}S}} \rangle_{PL}$, and e) $\langle \Delta G_{\mathrm{MM/GBSA}} \rangle$

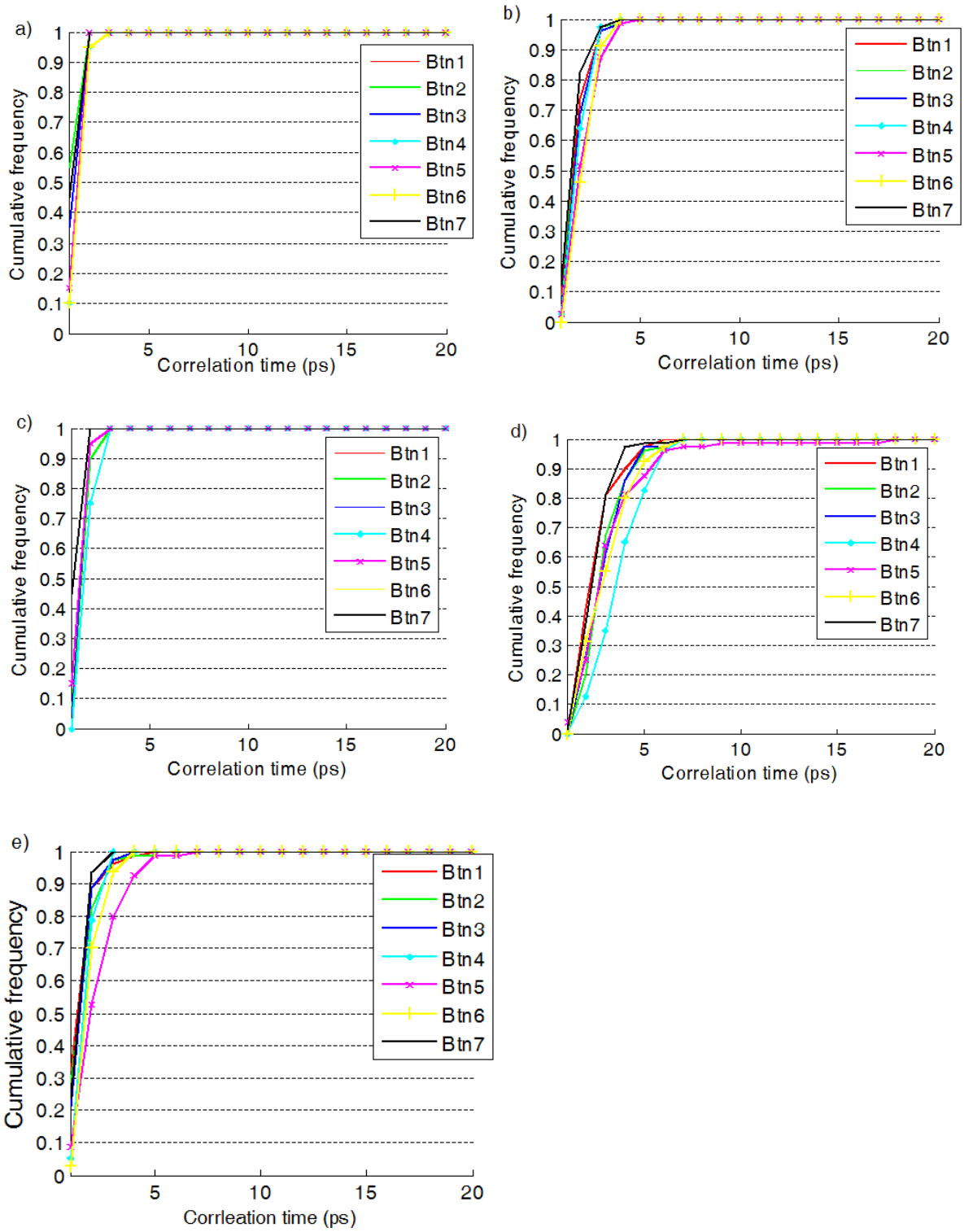

Figure 3. Correlation between predicted and experimental free energies of the seven biotin analogues for the truncated systems. The LIE results are obtained with α = 1.15 and with the ΔG_{cc} corrections.