of the study, it is difficult to be sure whether the high accuracy here is fortuitous or because of the use of a polarizable force field itself.

Dixit and Chipot also applied absolute free-energy calculations to compute the binding free energy of biotin to streptavidin; they obtained -16.6 ± 1.9 kcal/mol compared to an experimental value of -18.3 kcal/mol.[210]

Solvation free energies

Alchemical free-energy methods have often been tested by or applied to computing hydration free energies of small molecules like amino acid side-chain analogs or other small neutral compounds,[81,205,211] occasionally in a predictive context.[212] These tests can provide insight into the fundamental level of accuracy that can be expected of current force fields and also provide guidance for improvements to force fields. Also, several studies have highlighted the fact that solvation of small molecules can play an important role in determining binding free energies. One recent study found that the affinity of two trypsin inhibitors for water was very different, but these differences nearly canceled with differences in the binding site.[172] Another study suggested that solvation free energies played a substantial role in determining the change in binding affinity when optimizing fructose 1,6-bisphosphatase inhibitors.[168]

Predictive tests

Predictive tests of alchemical free-energy methods have been relatively rare, but at the same time are especially valuable for two reasons. First, to apply these methods in the context of drug discovery, they need to be predictive, and so testing them in a predictive context is a more realistic. Second, when doing retrospective studies, it is easy to be unintentionally influenced by the existing experimental results. For example, one might perform several sets of binding free-energy calculations with altered parameters and conclude that the "correct" set is the set that agrees best with experiment. Was any variation with parameters due to (a) the parameters themselves or (b) random errors due to poor convergence? And how would one proceed in a predictive setting?

However, alchemical free-energy methods have been applied predictively (together with or in advance of experiment) in a few studies. Here we focus especially on cases where experimental results are known and pass over those where we are not aware of any experiment that tests the computational results.

A number of studies have applied alchemical methods in the context of lead optimization. Some of the work from the Jorgensen lab on HIV-1 NNRTI has been predictive,[157,160,161] where free-energy calculations were used to help identify a binding mode and guide lead optimization. Similarly, the work on fructose 1,6-bisphosphatase has been predictive and applied in drug discovery.[168] Alchemical methods have been used to successfully predict an optimization of a neutrophil elastase inhibitor that was subsequently synthesized and tested.[213] Another application in lead optimization used grand canonical Monte Carlo techniques to guide lead optimization.[192]

Alchemical free-energy calculations were also used predictively, or at least in a joint theory-experimental study, in examining binding of benzamidium derivatives and their binding to trypsin.[214] The computational results correctly captured experimental trends, though falling short of quantitative accuracy. A joint theory-experiment study that examined relative binding free energies of inhibitors to a GPCR was used to help validate a homology model; the computational results proved accurate to less than 1 kcal/mol.[164] Alchemical free-energy calculations were also used to predict the tight binding site of ATP in F1-ATPase,[165] a prediction subsequently confirmed experimentally.[215]

The only predictive absolute binding free-energy calculation we are aware of to date is the work on the T4 lysozyme system by Mobley, Graves, and collaborators, where absolute free energies were used to predict binding affinities and binding modes of several new ligands.[113]

Pitfalls and negative results

Negative results and failures can in some cases be extremely informative, especially when it is possible to identify failures with specific issues, because in such cases the failures highlight the importance of certain factors. Unfortunately, negative results and failures are not always published, so it can be difficult to gather information in this area, and it is often even more difficult to trace failures back to specific issues. Nevertheless, there are several articles that highlight issues in this area – either by tracing failure to a specific cause or by identifying and avoiding a potential pitfall.

One major, common pitfall is a dependence of computed free energies on starting structure. Because a binding free energy is a ratio of partition functions, it involves integrals over all of the relevant configurations of several systems and thus must be independent of the starting configurations of the system. Unfortunately, computed results often depend on the starting configuration of the system – for example, different starting ligand orientations or different starting protein structures may give different results, as noted in a number of studies. This kinetic trapping is inevitable whenever energy barriers are sufficiently large.[217] Mobley et al. found that results could depend significantly (by more than 1 kcal/mol, in some cases) on starting ligand orientation.[179] In FKBP, Shirts[206] and Wang et al.[97] found that computed binding free energies could differ by more than 1 kcal/mol depending on the choice of starting protein/ligand configuration, and Fujitani et al. also observed a dependence on starting structure.[208] An even larger dependence on the starting protein conformation was observed in lysozyme, where computed values could