

important to keep in mind when analyzing docking results, and when presenting them to medicinal chemists.

One can speculate as to why scoring functions are better in terms of predicting the ligand pose. One possibility is that this is partly due to cancellation of errors because the ligand structure is a constant and the scoring function merely compares different poses. In addition, although the classical electrostatics are probably a good guide to the orientation of a ligand in the active site, determining their energetic contribution is a much more complex problem.

### LIGAND BINDING AFFINITY

The Holy Grail, at least for protein-structure-based modeling, is the accurate calculation of relative binding free energies. This is at the heart of virtually all mechanistic approaches to drug discovery, at least as it pertains to potency. Unfortunately, the free energy of binding of a ligand to a protein is a very complex calculation,<sup>48</sup> as mentioned already. Nevertheless there are a number of approaches that have been employed with some success.

The most theoretically rigorous, at least from a sampling point of view, are the free-energy perturbation methods that were described in Chapter 5. These methods have two significant drawbacks that limit their impact in typical drug discovery efforts. They are extremely expensive in terms of computer resources. This becomes less of an issue with each passing year as computers continue to become faster and faster, but it still stretches the computer resources of most companies. The other issue is that these calculations work best for small perturbations and so are limited to very conservative changes in the ligands of interest. In recent work Jorgensen<sup>49</sup> has proposed a paradigm for using small structural changes and free-energy perturbation (FEP) to guide the development of structure/activity relationships (SAR).

There have been many approaches developed that, unlike FEP, rely only on calculations for the “end points” of ligand, protein, and complex. These methods might all be described as linear interaction energy (LIE) models<sup>50–55</sup> because they all rely on some fitted model based on computed interaction energies,<sup>56</sup> but people differ in their precise terminology. Some of the most well-known approaches are the linear response method<sup>50,51</sup> or the molecular mechanics-Poisson-Boltzman [(MM-PB) or molecular mechanics/generalized Born model/solvent accessibility (MM-GB/SA)] methods.<sup>57,58</sup> Approaches that use interaction energies for the reactants and products of ligand binding are much more commonly pursued in drug discovery because of their more tenable computational cost. They have also in many cases proven themselves to be quite effective.

We have used simple linear interaction energy (LIE) calculations as a tool for comparing the relative affinities of prospective ligands. For example, we made great use of this approach in our BACE program at J&J.<sup>56,59,60</sup> BACE

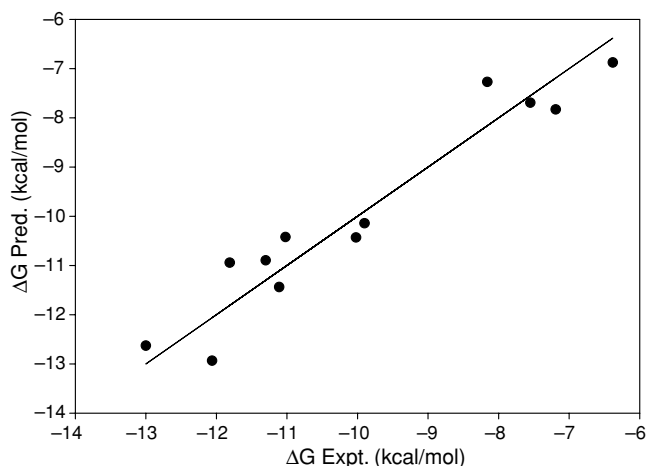


Figure 12.2. LIE model for BACE [Equation (12.2)].

is an aspartyl protease that plays a critical role in processing the amyloid precursor protein (APP) that has been implicated as a causative factor in Alzheimer’s disease. We, and other groups, have shown that it is possible to construct reasonable models for BACE using some variation of the LIE approach. Our initial models were derived using a series of peptidic inhibitors reported by Ghosh and his collaborators.<sup>41</sup> These structures and their corresponding affinities are given in Table 12.2. Our procedure was to optimize the ligand in the protein-binding site using OPLS and GB/SA water. Either the protein can be held frozen or a subset of residues near the active site can be allowed to relax. The optimized ligand is then extracted from the protein and allowed to minimize in GB/SA<sup>61</sup> water. The van der Waals and electrostatic interaction energies are then computed and used to fit the LIE model.

We used OPLS and GB/SA calculations for the protein, ligands, and protein-ligand complexes to fit the model for BACE binding given in Equation (12.2):

$$\Delta G_{\text{Bind}} = 0.2228^* \Delta U_{\text{vdw}} + 0.0577^* \Delta U_{\text{ele}} + 12.7464. \quad (12.2)$$

This model provides a root-mean-squared deviation (rmsd) of 0.58 kcal/mol and an  $r^2$  of 0.92 (Figure 12.2), and it has shown itself to be reasonably predictive with respect to compounds outside the training set.<sup>56,59,60</sup>

A similar model was derived for our internal series of BACE inhibitors<sup>63</sup> in spite of the fact that the structures are very different. The initial low micromolar lead for our BACE program (11) was identified in a high-throughput screen. Unlike most of the previously reported ligands, it was a non-peptidic inhibitor:

