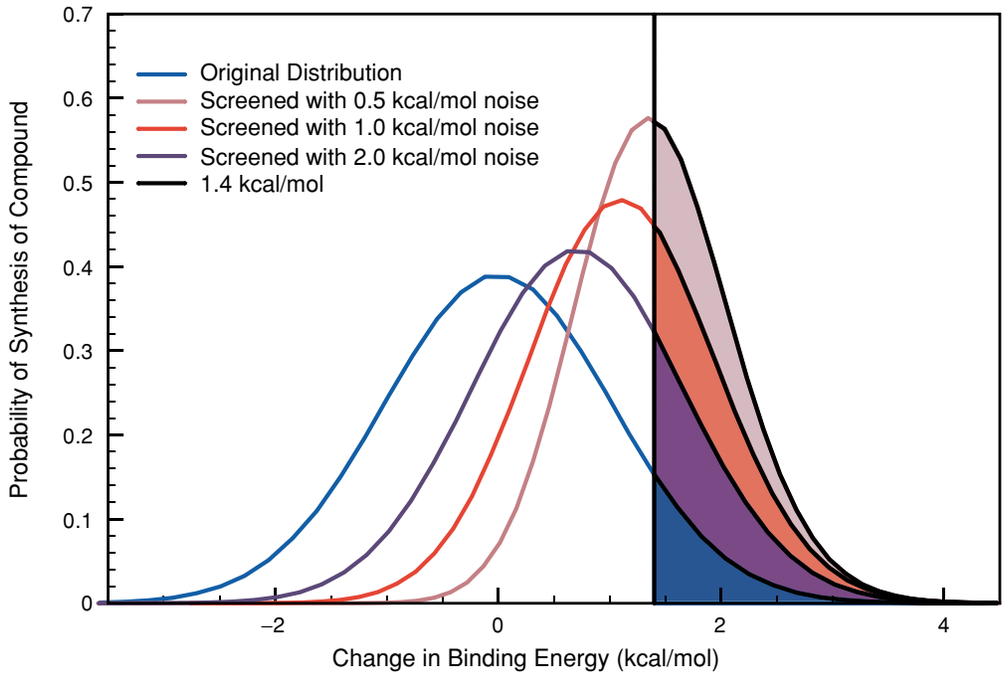

Figure 5.1. Modeled distribution of affinity changes of the proposed modifications (blue) compared to the distribution of affinity changes after computational screening with Gaussian error $\epsilon = 0.5$ (pink), $\epsilon = 1.0$ (red), and $\epsilon = 2.0$ (purple). The shaded area represents the total probability of a proposed modification with affinity gain greater than 1.4 kcal/mol. Hence, in many situations, even with moderate error, a reliable method of filtering compounds could significantly improve the efficiency of synthesis in lead optimization.

So even relatively small numbers of moderately accurate computer predictions may be able to give significant advantage in the pharmaceutical work flow. When we translate the chance of obtaining binding improvement into the number of rounds of synthesis required to obtain that improvement, then screening 100 molecules with 2 kcal/mol noise or 10 screened molecules with 0.5 kcal/mol noise in this model reduces the number of molecules to be synthesized by almost an order of magnitude. Clearly, these calculations assume the simulations are not biased against active compounds, and errors that are highly dependent on the binding system would result in less reliable advantages. The type of computation matters as well – computing relative binding affinities would require only one calculation to compare affinity changes, whereas absolute binding affinities would require two, increasing the effective error. But physically based prediction methods should in principle be more reliable than parameterized methods, as the basic physics and the atomistic details are transferable between drug targets.

This analysis is in agreement with informal discussions with pharmaceutical chemists, who mentioned reliability as being more important than pure speed or the highest accuracy. Many thought they could fit methods that took as much as a month into a work flow, as long as they truly converged reliably with 1 kcal/mol variance error. Even a slight decrease in reliability, for example, being off by several kcal/mol more than 20% of the time, greatly decreased the amount of time that scientists would be willing to wait, perhaps down to a day or two.

FREE-ENERGY METHODOLOGIES

A very large number of methods for computing binding free energies with atomistic molecular models have been developed. Most of them are still under active study, and each has different trade-offs between accuracy and computational efficiency. Because of the scale, complexity, and speed of methodological developments, choosing and applying methods can be confusing even to experienced practitioners. Here, we focus on an overview of some of the key methods available for computing binding affinities, emphasizing references to primary literature. A number of useful recent reviews have focused specifically on free energy methods.[8–14] Of particular note is a recent, fairly comprehensive book on free-energy methods, specifically Chapters 1–7.[15] Several molecular simulation and modeling textbooks have useful introductions to free-energy calculations as well.[16–18]

In this discussion of methods, we will assume standard classical molecular mechanics models, with harmonic bond and angle terms, periodic dihedral terms, and nonbonded terms consisting of point charges and Lennard–Jones repulsion/dispersion terms. In the vast majority of ligand-binding free-energy methods, calculations have been performed with these types of models. Adding