methods that strive to calculate the free energy of binding usually use an energy function also known as a "scoring function" that computes a score directly or indirectly, related to the binding free energy. Scoring functions have traditionally been either simplistic empirical or statistical potentials that relate observables to the free energy of binding by using statistical methods, or they are extremely detailed in nature and use physics-based descriptions of the molecular energetics and extensive sampling of receptor-ligand conformations via molecular simulation. Recently we have reviewed all categories of scoring functions and discussed their pros and cons with respect to RBDD.[75]

The use of quantum mechanics in structure-based drug design has until recently been either qualitative as described in the previous section or peripheral. For example, in large-scale virtual screening of databases using docking programs, semiempirical QM methods have been used during the database preparation phase to calculate atomic charges. A database (ZINC) of commercially available drug-like molecules prepared with QM charges and desolvation penalties has been made publicly available.[76] In a recent study Irwin et al., using ZINC, have successfully enriched known ligands that bind to metalloenzymes over non-binders in retrospective docking screens.[77] Although high-quality charges and desolvation penalties are not the only reason for this success, they no doubt play an important role.

Further evidence of the importance of the quality of charges comes from another study by Cho et al., in which ligand charges calculated using QM/MM methods led to significant improvement in the ability of docking programs to obtain the correct binding mode of the inhibitor.[78] The docking method that employed QM charges performed decisively better than force-field-based charges in ranking native binding modes as the best pose. The difference was more pronounced for poses that were predicted within 0.5 to 1.0 Å rmsd of the native pose. Raha and Merz have also designed a classical scoring function – the molecular recognition model – that used CM2 charges calculated using semiempirical QM for modeling electrostatic and solvation effects during binding.[63] It is noteworthy that charges in this case were computed for the entire protein/ligand complex using linear-scaling methods thus accounting for polarization and charge transfer. The molecular recognition model was able to calculate pK_is that agreed with experimental pK_i (correlation coefficient R^2 of 0.78) for thirty-three inhibitors modeled in the active site of HIV-1 protease.

QM/MM and binding affinity calculation

QM/MM methods are widely used to study mechanistic aspects of enzyme catalysis or in peripheral aspects of RBDD such as small-molecule charge calculation in molecular docking as described above. However, few studies have attempted to use QM/MM, either directly or indirectly, for calculating the free energy of binding between a protein and a ligand. In an earlier study, Mlinsek et al. used QM/MM to generate the MEP on the van der Waals surface of thrombin

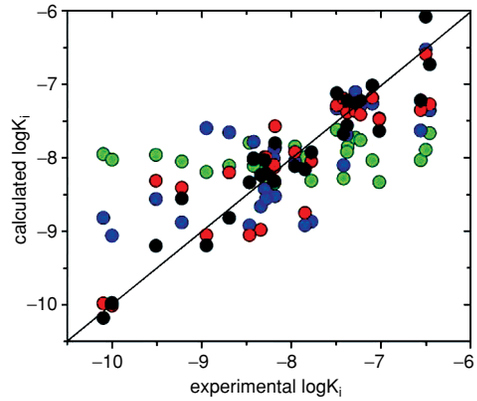

Figure 8.8. Four-tier approach used by Khandelwal et al. Shown is the correlation between the experimental and calculated inhibition constants of a series of hydroxamates against MMP-9 as given by FlexX docking in step 1 (green), QM/MM minimization step 2 (blue), MD simulations with constrained zinc bonds step 3 (red), and by QM/MM energy calculations for the time averaged stuctures from MD simulations step 4 (black).

and then used the MEP as input into an artificial neural network/genetic algorithm engine for data reduction and combination for predicting the pK_i of thrombin inhibitors.[79] Although artificial intelligence methods have shown good success in such studies, often they lack generality. Moreover, the MEPs used in this study were descriptors that ignored other aspects of binding.

In a very recent study, Khandelwal et al. used a four-tier approach that involves docking, QM/MM optimization, MD simulation, and QM/MM interaction energy calculation to predict binding affinity.[80] The authors use a modified version of extended linear response (ELR) theory where the van der Waals and electrostatic terms are replaced by QM/MM interaction energy:

$$\Delta G_{\text{binding}} = \alpha \times \Delta \langle E_{\text{QM/MM}} \rangle + \gamma \times \Delta \langle SASA \rangle + \kappa,$$

where $\langle E_{\text{QM/MM}} \rangle$ is the time average of single-point QM/MM interaction energies obtained from MD simulations. The authors calculated the binding affinity of twenty-eight hydroxamate-based inhibitors of matrix metalloprotease (MMP-9) using this approach with impressive accuracy. The agreement between the calculated and experimental pK_i is excellent ($R^2 = 0.9$ and cross-validated R^2 ranging from 0.77 to 0.88). What is also noteworthy is that the authors clearly demonstrate an improvement in predictive accuracy with every step of their four-tier approach. As shown in Figure 8.8, the agreement with experimental pK_i improves from poor, after the first step of docking ($R^2 = 0.044$; green circles in Figure 8.8) to very good ($R^2 = 0.90$; black circles in Figure 8.2) after the final step of QM/MM single-point interaction energy calculation. This points toward the importance of a quantum mechanical treatment and the sampling of active conformations in accurate binding prediction. Specifically, the QM/MM treatment of the active site is very important (step 3) because it was shown that a proton is transferred from the hydroxamate hydroxyl to the active-site glutamate.