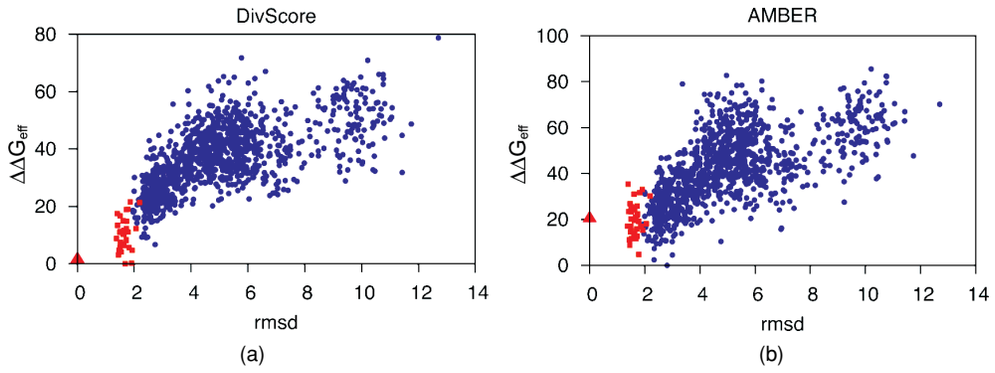

Figure 8.7. Energy versus rmsd plots for the fructose repressor DNA binding domain (PDB id: 1uxd) scored using the DivScore potential (a) and AMBER (b). (▲) NMR minimized mean (taken as the reference for rmsd calculations); (■) individual NMR models; (●) Rosetta decoy. Energies are reported as the difference in energy for a state compared to the lowest energy structure in the decoy set. AMBER scores several decoys better than native models while DivScore correctly identifies several NMR models as native.

although interestingly it is the x-ray structures and not the NMR structures that are generally lowest scoring. The Z-scores for all native structures are large, indicating that the potential function scores the native structure much better than the set of decoys. The energy gaps between the native structure and the best-scoring decoy are large for the four-state-reduced set, although noticeably smaller for the Rosetta decoys. A sample of the resulting plots of DivScore versus rmsd from native is given in Figure 8.7.

It is interesting that semiempirical methods work so well in identifying native structures from nonnative, considering that proteins contain functional groups that were not explicitly parameterized in the semiempirical Hamiltonians used (AM1[48] and PM3[49]). Furthermore, classical approaches have an advantage in that they have been parameterized for a focused set of functional groups found in biological molecules,[13] whereas SE parameters are implemented at the level of individual elements. In addition, macromolecular effects such as nonlocal van der Waals interactions and multiple charged-charged interactions become significant.

The ability to use semiempirical single-point measurements to discriminate native structures from nonnative indicates that these methods are suitable for applications involving proteins and may be capturing important interactions that lead to protein stability. It is worth considering why semiempirical models score protein decoys as well as we have found in the present study. Semiempirical methods are known not to give phi-psi plots that agree with high-quality ab initio results,[50] while force fields are generally parameterized to reproduce these plots at some level of accuracy. This suggests that other factors play a role like long-range electrostatics or cooperativity effects observed in the folding of secondary structural elements.[57] Possibly these effects are overwhelming the conformational effects when using semiempirical methods in scoring native and decoy protein structures. With this approach in place one can start to consider using semiempirical methods to validate homology models or to use it to study the preferred conformation of loops in proteins.

STRUCTURE-BASED DRUG DESIGN

Qualitative uses of QM in RBDD

The ability to characterize a macromolecule such as a protein using QM opens up a whole new range of descriptors that can aid drug discovery. Many of these descriptors are beyond the reach of classical potentials and by their very nature can be used to gain a qualitative understanding of protein/ligand interactions and then be used in the rational design of drug molecules. Linear-scaling QM methods have made therapeutically important protein targets accessible to qualitative analysis, from a rational drug design perspective. These qualitative insights are often used to predict ligand binding or metal binding "hot spots" that can be targets for small-molecule inhibitors. Workers have made use of descriptors such as molecular electrostatic potential (ESP) maps, local hardness and softness, Fukui indices, frontier orbital analysis, density of electronic state analysis, and so on, to probe proteins. Below we concentrate on recent studies that have employed QM derived descriptors.

ESP and relative proton potential

ESP maps have been widely used as a tool for characterizing protein or DNA binding sites in RBDD. However, these maps have traditionally been derived from classical point charge models (for example, PARSE) that were used to compute the electrostatic potential on the surface of proteins by solving the linear or nonlinear PB equation. With the advent of linear-scaling QM algorithms, combined with self-consistent reaction field methods to model solvation, ESP maps can now be computed quantum mechanically. Khandogin and York, using linear-scaling QM technology to generate ESP maps, have probed properties of therapeutically important protein targets such as HIV-1 nucleocapsid (NC) protein.[58,59] These authors have clearly demonstrated the advantage of using the PM3/COSMO computed molecular electrostatic potential (MEP) map over the PARSE/PB map, in discerning between the electronegativity of the