A great deal of experimental and computational effort has been expended to try to determine the correct protonation state for these residues.[32–38]

My experience suggests that many structures also contain bad contacts and/or structural elements that are highly strained. Structural problems are often most prevalent in the bound ligand.[39,40] Unfortunately this is precisely the region of the protein that is of the greatest interest from a modeling point of view. Many modeling studies have probably been done with the coordinates as they come from the crystal structure, but it is generally a good idea to do at least a small amount of geometry optimization to remediate the most serious structural problems. In most cases limited minimization will result in a protein structure that is little changed in terms of the overall geometry, and even fit to the x-ray data, but the overall strain and the most severe geometric problems can be greatly reduced.

One of the final decisions is the following: water or no water? For docking and many modeling tasks it is common to remove the water molecules from the protein. This makes sense from the point of view of having an unbiased structure for docking, for example. It is also simple – just remove them all. However, there are circumstances where waters do make critical contacts not only with the protein, but in some cases also with the ligand. In such cases removing these critical waters is suspect. Leaving them all in can also pose problems. For example, most docking programs will not displace waters automatically, so any waters in the protein are essentially treated as an unchangeable part of the protein. Selective removal of waters is an intermediate course. This makes sense when a few waters are seen consistently across a number of crystal structures and/or where the water is involved in a key mediating interaction between protein and ligand. Of course this approach is not without hazard because the choice of waters for inclusion can have a big effect on the protein-binding site.

I have personally been bitten by this particular bug in a study of the protonation state of BACE. In our first publication[31] we elected to include a key water in the active site. Although this had no effect on our prediction that the preferred overall protonation state of the two catalytic asps in BACE was the −1 state (i.e., monoprotonated), it did lead to a surprising difference in energy between two specific states involving protonation of Asp32 or Asp228. In retrospect, this significant asymmetry in the active site appears to be due to this water that cooperates with Ser35 in stabilizing Asp32 when it is deprotonated. A later quantum mechanics/molecular mechanics (QM/MM) structure refinement study[30] showed that the energy for protonation of these aspartates were very nearly equivalent. It seems likely that one of the primary reasons for this discrepancy was the treatment of waters that were all included or all replaced by a continuum model in the later QM/MM refinement. This serves as a cautionary tale with respect to selective inclusion of waters in the active site. The protonation state question also highlights one of the other potential pitfalls in crystallography, namely that the crystallization conditions and environment where the enzyme is actually active may be quite different. In the case of BACE, many, if not most, of the crystals have been grown[34,41,42] at a pH near 7 while the enzyme itself is most active under more acidic conditions. This pH difference may be inconsequential, but that cannot be known for certain, particularly for a property as sensitive to medium effects as the protonation state.

In summary, it is important to carefully consider how the protein structure has been prepared. Protein preparation can have a meaningful impact on the quality of any subsequent modeling using that structure. This probably becomes more critical as one asks more from the results, such as computing relative binding affinities, or, as in the case mentioned above, energetically sensitive properties such as protonation states.

DOCKING AND SCORING

Docking and scoring is described in great detail in Chapter 7, but it is worth reiterating some of the practical conclusions of the many validation studies that have been done.[43–47] First, in most drug discovery efforts docking is used for one of two purposes. It may be used to determine the most probable docking pose for a ligand in a protein-binding site where a crystal structure is unavailable. Here the goal is just to find the correct orientation and conformation of the ligand in the protein. The second use for docking is in a virtual screening mode. In this mode the specific docking pose is not as important as ranking a set, usually a very large set, of chemical structures in terms of their propensity to bind to the target of interest. In the first mode the correct structure is the key result. In the second mode, enrichment of a screening set, in terms of potent ligands, is the desired result.

The literature suggests that it is possible in many cases to predict the best binding pose. At least the best binding pose is often near the top of possibilities unless there is a significant change in the protein structure, such as an allosteric modification or significant induced fit on the part of the ligand. Apparently available scoring algorithms are able to differentiate between good and bad binding poses for any one specific ligand. It is also true that decent enrichment factors often result from docking-based virtual screening, albeit perhaps not as significant as we might hope in most cases. It is very clear that scoring functions used in evaluating docking poses are exceedingly poor when asked to rank compounds in terms of their affinities. Most evaluations show essentially no correlation between docking score and affinity. Our experience at J&J has been consistent with this result. Therefore, it is important to consider docking-based virtual screening more as a filter than any kind of ordered list. Docking does filter out compounds that fit the active site poorly, but it does not differentiate well between weak and potent binders. This fact is