Abstract

Protein crystal structures provide a valuable source of information on the internal interactions of a protein and provide a starting point for simulations. In this thesis, we examine how large-scale analysis of protein structures can explain unexpected geometries and interactions and provide a starting point for further modeling. The large-scale analysis takes two forms: large datasets and large calculations. We first investigate unexpectedly short non-covalent distances in published protein crystal structures. We observe over 75 000 close contacts in a curated dataset of high quality protein structures, and examine the trends in which residues and atoms are involved in these close contacts. We characterize a subset of over 5000 CC with quantum mechanical and molecular mechanical methods to understand their stability. We examine a particular protein, acyl carrier protein, to see how charge parameterization affects its behavior in long-time molecular dynamics simulations. Finally, we test the Fukui shift analysis (FSA) method, which identifies how frontier states of an active site are altered by the presence of an additional QM residue to identify when QM treatment of a residue is essential.