12. Structural Bioinformatics of Protein-Bound Water

Christopher Bottoms, Tommi White, John J. Tanner

University of Missouri-Columbia, United States of America

Keywords: Structurally conserved water, protein crystallography, structural data mining, solvent

We describe a computational method of finding structurally conserved water molecules within families of protein structures. The search is simplified by using a density map to represent the distribution of water molecules in protein families. This method identifies sites previously described in the literature as well as novel ones.

Water plays essential roles in protein structure and function. Of special interest are those water molecules bound to specific sites within protein structures. They can stabilize protein conformations by filling pockets in the protein structure or by bridging secondary structure elements via hydrogen bonding. They often contribute to protein-ligand recognition and catalysis, and can therefore influence the success or failure of structure-based drug design. Certain water-binding sites are particularly important, as evidenced by their conservation among proteins sharing a common three-dimensional fold. We have developed a computational method that allows the rapid identification and preliminary analyses of such sites.
The algorithm of this method involves several steps. First, structurally related protein structures are superimposed. Next, the water molecules in this superposition are used to create a density map. This step was performed using either CNS or fft and mr_map from the CCP4 suite.

Water molecules at the highest peaks of the map are analyzed for their interactions. From this, we develop generalized interaction descriptions of the various conserved solvent sites. Percent conservation is based on (1) similar locations of solvent molecules and (2) similar interactions. Two atoms were considered to interact with each other if they were within 3.2Å of each other and one could serve as a hydrogen bond donor and the other as an acceptor. We have used this method to find structurally conserved solvent sites in multiple protein families. In a recently submitted paper, this method was used to describe structurally conserved solvent sites in fatty acid binding proteins, monodomain cytochrome c proteins, lactate/malate dehydrogenases, and parvalbumins.

This method was also used in the study described by "Conserved solvent and side-chain interactions in the 1.35 Å structure of the Kelch domain of Keap1" by L. J. Beamer, X. Li, C. A. Bottoms and M. Hannink. This article was accepted 14 July 2005 for publication in Acta Crystallographica Section D: Biological Crystallography. This study describes ten different solvent sites that are conserved among all six blades of a beta-propeller.