Water: The Most Important Molecule in Drug Discovery

Gregory Ross

Abstract: *Diseases can often be attributed to malfunctioning of proteins; the molecules that constitute the vast majority the machinery and scaffolding within our bodies. As a result, we take medicinal drugs to bind to and modulate the functionality of faulty proteins. However, because the body is mostly an aqueous environment, water effects are an important consideration in the development of all new drugs. In particular, when a small molecule binds to a protein, water within the protein's binding site can be expelled or rearranged so as to facilitate or hinder binding. Knowledge of this process can help medicinal chemists design selective and stronger binding drugs to combat diseases.*

Water is a substance that is utterly fundamental to life. It is well known that water constitutes up to 70% of an adult's body weight, so it is hardly surprising that water plays a key role in almost all biological functions. Water is the fluid through which molecular machinery is transported within cells and around the body, and its unique properties facilitate chemical reactions. Importantly, as this article will discuss, water plays a critical role in molecular binding and informs the creation of new drugs to combat diseases.

Proteins are biology's molecular machines that carry out all of the functions necessary for life. Endogenous diseases occur due to the malfunctioning of one or more of our proteins. We use drugs (medicinal or recreational) to inhibit, activate or modulate the working of particular proteins. For instance, ibuprofen and aspirin are believed to inhibit cyclooxygenase,[1] a protein that is associated with inflammation. In the case of bacteria and viruses, drugs are designed to target microbial proteins and hinder their functions so they cannot proliferate or produce harmful substances in the body. Proteins catalyse reactions, relay signals, and carry out all their functions through binding events. These take place in grooves or pockets on the proteins' surfaces or in the proteins' interiors, which are known as binding sites. Drugs can interfere with a protein's functionality by fastening onto its binding site and preventing other binding events. Small molecules that bind to proteins are known as ligands.

Once a protein has been identified as being integral for a disease, it is typical for a pharmaceutical company to virtually screen hundreds of thousands to many millions of compounds against it, with the aim of finding compounds which bind strongly to the protein's binding site, and thereby identifying the most promising candidates for inhibiting the protein. This is an important initial step in drug discovery as subsequent chemical screens run on the candidates are time-consuming and very expensive, so by screening ligands virtually, both time and money can be saved. Despite decades of research, however, accurately predicting the binding affinity of a ligand is still very difficult, even once the 3D structure of the complex is known. It is harder still to predict a ligand's binding strength at the rapid speed required by virtual screens.[2] This shortcoming can be partly attributed to the misrepresentation of water in existing binding models. Rather than being a passive medium through which binding processes take place, water

plays an important part of the recognition process and in the affinity of the binding pair. One of the most important factors that drive a ligand to bind with a protein is the tendency of non-polar substances to cluster together in an aqueous (water-based) solution. It is called the hydrophobic effect and is a process that merits an entire article of its own.[3] Water, however impacts the binding process in subtler ways as well, such as through the rearrangement of water molecules in the binding site when a ligand binds.

A classic example is the bacterial oligopeptide binding protein (OppA). This protein is responsible for transporting small peptides, or chains of amino acids (Figure 1a) from outside a bacterial cell to its interior. Due to the nature of its task, OppA must exhibit binding plasticity, or be able to accommodate a large variety of ligand shapes and chemical types. It does this partly by changing the number and distribution of water molecules in its binding site. When a ligand is too small to bond directly with the protein, as in Figure 1b, four water molecules surround the ligand, bind to the polar group and bridge the interaction between the peptide and the protein. In Figure 1c, where the peptide side chain is large and hydrophobic, the opposite is true: to accommodate the ligand, all but one of the water molecules are expelled from the binding site. Water molecules acting like those in Figure 1b are very common; it has been reported that 85% of protein-ligand complexes have at least one water molecule bridging the interaction between the ligand and the protein.[4]

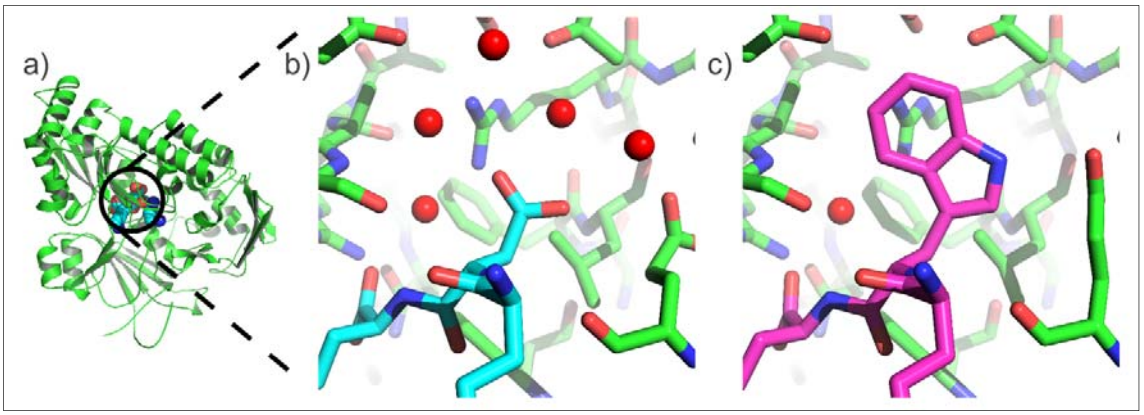

Figure 1. a) The bacterial peptide transporter OppA binds to a myriad of small peptides. b) As the peptide (in blue) is too small, four water molecules (red spheres) are able to bridge the interaction to the protein (green sticks). c) All but one of the water molecules has been expelled to accommodate the larger peptide side chain (purple sticks).

Whether or not a ligand will displace a water molecule is of particular interest to medicinal chemists, as water displacement is seen as a route to improve the affinity of a drug. A classic example of this can be seen with HIV protease, a protein necessary for replication of the HIV virus (Figure 2a). X-ray crystal structures of HIV protease revealed that a water molecule bridged the interaction between the ligand and the protein (Figure 2b). As a result, a new class of potent drugs were created to displace that same water molecule (Figure 2c).[5] In fact, this was the first example of the targeted displacement of a water molecule.

We expect the affinity of a ligand to increase when it displaces a water molecule because the process increases the *entropy* of the system. Entropy, a measure of disorder, is one of two constituents that determine the binding strength of a reaction. If a water molecule is displaced

from a protein's binding site, it is more able to freely rotate and explore its surroundings, thus making the system more disordered. The other factor determining binding strength is *enthalpy*, a measure of the intermolecular energies of the protein-ligand complex. To understand the respective roles of entropy and enthalpy in a binding event, we can imagine two ligands which are identical, except that one has a chemical group which displaces a water molecule upon binding with the protein. If this ligand can fully replace the bonds between the protein and the dislodged water molecule, then the change in enthalpy caused by its binding will be roughly the same as that of the other ligand; but in terms of entropy, the water-displacing ligand is the stronger binder because of the increase in disorder provided by freeing the water molecule. However, the targeted displacement of a water molecule may not always result in stronger binding, especially if the ligand cannot compensate fully for the stabilising role played by the water molecule within the protein.[6]

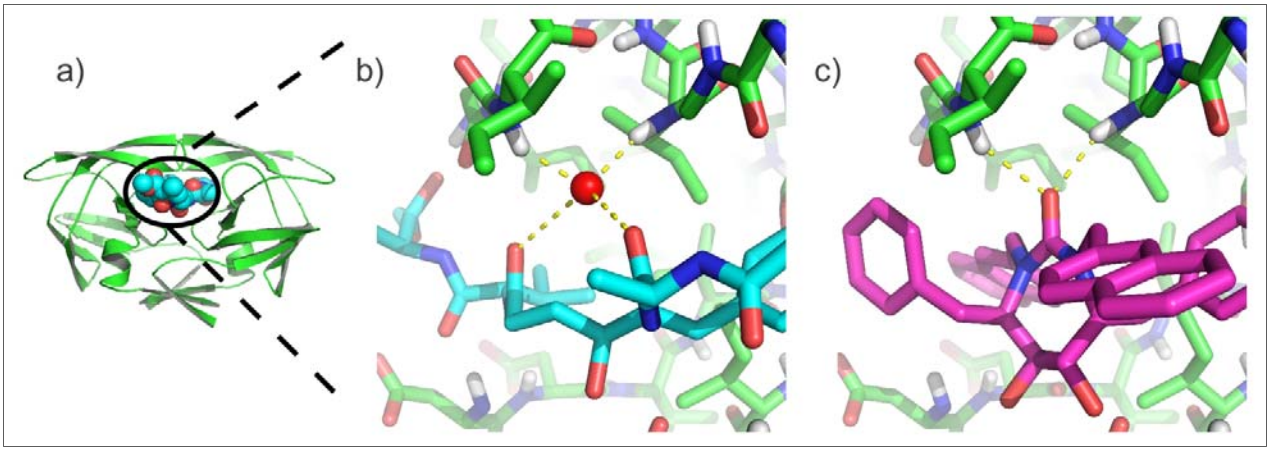

Figure 2. a) The protein HIV protease is required by the virus to help create a complete viron. b) An X-ray crystal structure of HIV protease with an inhibitor reveals that a water molecule is bonded to both the protein (green sticks) and the inhibitor (blue sticks). c) As the first example of its kind, a new inhibitor was designed to displace the same water molecule to improve its binding affinity.

We have seen with OppA that water molecules in protein binding sites can help ligands bind to proteins either by bridging the interaction or by being dislocated to make room in the binding pocket. Further, with HIV protease we have seen that the accommodating nature of water can be exploited to create more selective and strongly-binding drugs. Hence, our ability to predict the role of water in a protein binding site may hold vital importance for a drug discovery program. Before this can be done, however, we must be able to accurately determine the locations of water molecules; after all, we cannot possibly predict what role the molecules play if we have not first identified them.

My work focuses on developing methodologies to elucidate the role of water in protein-ligand binding. The end goal of my research is to inform better ways to calculate the binding strength of a compound accurately and within the time constraints demanded by virtual screening. As a starting point for my approach, which combines molecular simulations and empirical analysis, I use 3D structures of many different proteins that have been determined with X-ray crystallography. Thus far I have established a method for predicting water locations accurately and rapidly (Figure 3). Furthermore, I have used this new, rapid method to study a

large range of protein-ligand X-ray crystal structures, and have established empirical ways to predict whether a water molecule will be displaced by a chemical group on a ligand. I hope that my work will aid the faster, more affordable development of potential new drugs.

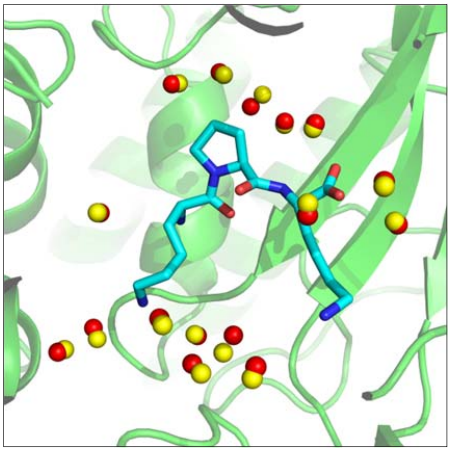

Figure 3. An example of my rapid water prediction method. A ligand (blue sticks) is shown bound to the protein OppA (green cartoon). My predictions (yellow spheres) are overlaid on top of water molecules from the crystal structure (red spheres). All water locations are correctly determined.

References

[1] Busson, M. (1986). Update on ibuprofen: review article. *The Journal of International Medical Research*, *14*, 53-62.

[2] Subramaniam, S., Mehrotra, M., & Gupta, D. (2008). Virtual high throughput screening (vHTS)-a perspective. *Bioinformation*, *3*, 14-17.

Chandler, D. (2005). Interfaces and the driving force of hydrophobic assembly. *Nature*, *437*, 640-647
[3] Lu, Y., Wang, R., Yang, C.-Y., & Wang, S. (2007). Analysis of ligand-bound water molecules in high-resolution crystal structures of protein−ligand complexes. *Journal of Chemical Information and Modeling*, *47*, 668-675.

[4] Lam, P. Y., Jadhav, P. K., Eyermann, C. J., Hodge, C. N., Ru, Y., Bacheler, L. T., Meek, J. L., Otto, M. J., Rayner, M. M., & Wong, Y. N. (1994). Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science*, *263*, 380-384.

[5] Clarke, C., Woods, R. J., Gluska, J., Cooper, A., Nutley, M. A., & Boons, G. J. (2001). Involvement of water in carbohydrate-protein binding. *Journal of the American Chemical Society*, *123*, 12238-12247.