Just add water! The effect of water on the specificity of protein–ligand binding sites and its potential application to drug design

John E Ladbury

Recent data have highlighted the enigmatic role that water plays in biomolecular complexes. Water at the interface of a complex can increase the promiscuity of an interaction, yet it can also provide increased specificity and affinity. The ability to engineer water-binding sites into the interface between a drug and its target might prove useful in drug design.

Address: Department of Biochemistry and Molecular Biology, University College London, 91 Riding House Street, London W1P 8BT, UK.

Chemistry & Biology December 1996, **3**:973–980

 ISSN 1074-5521

Introduction

Water is a highly versatile component at the interface of biomolecular complexes; it can act both as a hydrogen bond donor and acceptor, it imposes few steric constraints on bond formation and although it takes up less space than the polar sidechains of a protein it can take part in multiple hydrogen bonds. Water can thus confer a high level of adaptability to a surface, allowing promiscuous binding, yet it can also provide exquisite specificity and increased affinity to an interaction. Structure-based drug design strategies currently largely ignore the effects of water, because the structural and thermodynamic effects of water's inclusion in binding interfaces are hard to determine and hard to model. There are, however, several well-studied examples in which water is extremely important in defining the interactions between molecules, and so leaving water out of a drug-design strategy clearly reduces the chance that the strategy will be successful.

Recent structural and thermodynamic data indicate that water at a protein–ligand interface may make a favourable contribution to binding. Of course, any water molecule in such an interface makes a favourable contribution to the free energy compared to a void in the same position. But in general it has been assumed that, because the entropic cost of trapping highly mobile water molecules is so large, an interface that leaves no space between the interacting protein and the ligand will necessarily give a higher binding affinity than one in which the interface contains gaps filled by water. And yet, in some cases the enthalpy gain that results from making extra water-mediated hydrogen bonds is greater than the entropic penalty that must be paid for immobilizing the water involved. Thus, to understand the role of water in protein–ligand interactions one must establish whether it is possible to make a more favourable interaction by substituting a ligand, or part thereof, for the water molecule. If not, then an ideal ligand would not be one in which the binding interface is structurally complementary, but one that includes water-binding sites in appropriate places. It would therefore be potentially valuable to drug design to understand exactly how water is involved in biomolecular interactions. One approach to this is to identify and characterize the binding sites where water improves the overall binding of an interaction. This review considers this question from two points of view. First, I illustrate the versatility of water using examples of protein–ligand interactions in which the presence of water appears to contribute to the stability of the complex. Second, I discuss the ways in which this improvement in binding can be used by applying the structural and thermodynamic principles deduced from

these observations to the problem of how to include water in structure-based ligand design strategies.

Energetics of water at the interface

The effects of water on protein–ligand interactions are very complex and therefore accurate measurement of these effects is extremely difficult. The only hope for an accurate measurement is to combine high resolution structural detail with accurate thermodynamic data. To define a binding event fully, the interacting molecules have to be characterized in their individual free states and in the bound state. Clearly, when one considers the water molecules as an integral part of the definition of these states, the problem becomes more complex. A full description of the binding interaction now requires an understanding of the change in hydration states of the protein and ligand when the interface forms, and an assessment of the entropic and enthalpic effects of these changes. In the examples given below, and in ligand design generally, these effects are not easily quantified and uncertainties in the determination of these contributions will affect any conclusions drawn.

The interactions of the water molecules that are observed in biomolecular complexes in high resolution crystallographic studies (and to a lesser extent in solution structural studies [1,2]) are energetically more favourable than bulk solvent interactions (see [3] for review). Buried water molecules have long residence times in protein structures (10^{-2}–10^{-8} s), and the interactions of water in a bimolecular interface are therefore thought to occur over periods greater than a nanosecond. As noted above, in many biomolecular interactions removing water from a binding site has a highly favourable effect. This is due to the entropic gain when surface-associated solvent molecules are released into bulk solvent. It is, however, possible to rationalize the fact that water sometimes provides a favourable contribution to the free energy of binding of the interface using thermodynamical considerations; although the restriction of the degrees of freedom of a water molecule incurs an entropic penalty, this cost can be compensated for by the enthalpic gain resulting from the formation of hydrogen bonds. It is clear that the balance between the enthalpic and entropic contributions is a fine one; for a water molecule to contribute to increasing the binding affinity, it has to be in a binding site which provides the maximum number of hydrogen bonding partners at the right proximity and orientation.

Observing water molecules in protein–ligand binding sites

Before any conclusions can be drawn on the effects of the inclusion of water molecules in a protein–ligand interface the positions and orientations of the bound waters have to be ascertained. To describe water molecules in an interface accurately is a formidable task. Data from studies of solution structures are inherently prone to interference caused by bulk solvent molecules and the experimental time resolution is seldom sufficient to allow the detection of bound water. The solvents required to induce crystal formation for X-ray studies, however, often lead to the determination of structures of complexes in solutions with a low water concentration. In X-ray spectroscopic determinations, water is observed as regions of electron density in a time-averaged map for which the data collection can take several hours. The position of the electron density fluctuates over the period of data collection and the water molecule is then assigned to the point that corresponds to a local minimum of electron density [3]. Generally, crystallographic structural data with a resolution better than 2.0Å are required to permit reliable discrimination of water molecules. Because of the range of possible hydrogen bonding angles, assignment of the exact position of a water molecule is uncertain unless neutron diffraction studies can be performed, although the location of donor and acceptor atoms in the binding site provide circumstantial evidence. The identification of water molecules is often the final step in structural refinement and, as their inclusion can help to improve the apparent fit of the data, it can be tempting to position them where they will most improve the statistics of the structure. Thus, any attempt to describe the effects of water molecules in detail, or to draw inferences from such a description, must bear these caveats in mind [4,5].

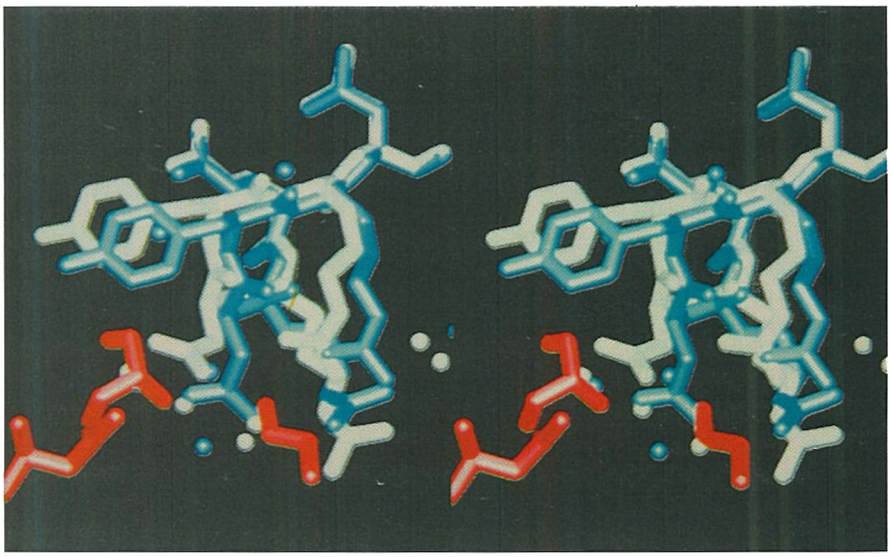
The incorporation of water molecules into a binding interface is generally believed to be energetically unfavourable. Indeed, much of the theory on molecular interactions is based on the fact that removing water from surfaces provides a driving force to cause intermolecular association, and ligands designed to displace water molecules can be highly potent inhibitors [6–9]. Nevertheless, in the systems that I discuss below, the inclusion of water appears to have the opposite effect, acting to stabilize an interaction. These examples primarily show how water increases the range of specificity of a binding site (nature uses water to extend the range of ligands that can be bound by a protein), and one could, therefore, view the improvements in binding that result from the inclusion of water as coincidental. In these cases, the ligands are not tightly bound — nature has used water to extend the range of ligands but the range of binding affinities is small. The important point for the purposes of this review, however, is that where additional water molecules are included improvements in the binding occur. If the causes of such improvements can be identified, they could be useful in ligand design.

Antibody–antigen interactions

X-ray crystallographic studies of the free and complexed forms of the bacterially expressed Fv fragment from a monoclonal antibody (mAbD1.3) that binds to hen egg-white lysozyme (HEWL) clearly show that a large number of water molecules are situated in and around

Figure 1

Stereoview of the superposition of the structures of the free (blue) and complexed (white) Fv fragment of the antibody mAbD1.3 showing the negligible change in structure of the binding site which occurs on binding. The water molecules preclude conformational rearrangements of the proteins. The HEWL antigen in the complex structure is shown in red. Water molecules are shown and coloured according to the structure in which they were found. (Figure kindly provided by R.J. Poljak.)



the antibody–antigen interface [10,11]. These water molecules form an intricate network that helps to make the antibody–antigen interface fully complementary. A comparison of the structures of free and bound Fv shows that some of the ordered water molecules in the free antibody are retained, but additional water molecules are required to complete the binding interface. No significant conformational change in the structures of either the antibody domains or the antigen is observed (Fig. 1), although a slight relative displacement of the V_L and V_H domains was apparent [11]. The residues in the binding interface appeared to show reduced mobility on complex formation, as detected by a decrease in the temperature factors; this was also observed in water-mediated protein–DNA interactions [12,13]. Calorimetric determination of the thermodynamics of binding showed that the interaction at 297 K was accompanied by a large favourable enthalpy ($\Delta H^\circ = -90.0$ kJ mol^{-1}) which more than compensates for the unfavourable entropy ($T\Delta S^\circ = -42.8$ kJ mol^{-1}). It can therefore be concluded that the network of water molecules that mediate this interaction is significantly stabilizing. The effects of water concentration on binding have been assessed under different conditions of water activity in a range of co-solutes. As the concentration of water in the complex interface was increased, so did the size of the favourable enthalpic and unfavourable entropic contributions. The overall result was tighter binding, supporting the notion that the water molecules in the interface stabilize the interaction [14].

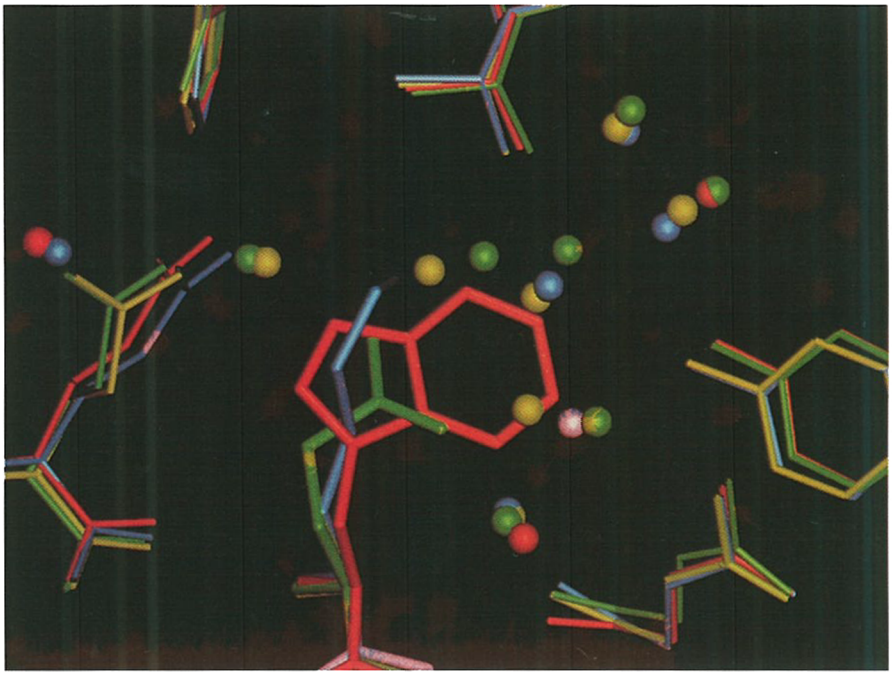
Protein–carbohydrate interactions

The effect of water on binding-site specificity is clearly demonstrated in the interaction of L-arabinose binding protein (ABP) with two different sugars: D-galactose (Gal) and L-arabinose (Ara). The structures of the complexes have been determined to high resolution (1.7 Å and 1.8 Å, respectively) [15]. On binding, each monopyranoside is completely enclosed in a cavity formed between the two globular domains of ABP, and each sugar adopts a similar orientation. The replacement of D-galactose by L-arabinose results in one water molecule being included in the position initially occupied by the -CH_2OH group. This substitution allows one additional hydrogen bond to form in the ABP–Ara complex as a result of a local structural rearrangement in the binding site and a slight increase in affinity (K_d(ABP–Ara) = 0.98×10^{-7} M; K_d(ABP–Gal) = 2.3×10^{-7} M). The thermodynamic effect of this change is negligible, indicating that the entropic cost of including the water molecule in the binding site is similar to the cost for the immobilization of the -CH_2OH group of D-galactose. Water is clearly able to change the specificity of this binding site and moreover can increase, albeit slightly, the affinity of the protein for the ligand.

Protein–peptide interactions

The oligopeptide binding protein, OppA, binds peptides of two to five residues no matter what the amino acid sequence. This promiscuous binding is mediated by two types of interaction: first, the protein binds directly to the peptide mainchain, and second, the peptide residue sidechains are accommodated in binding cavities through interactions largely mediated by water [16]. The thermodynamic effect (at 298 K) of including water molecules in the binding site was explored by taking a series of tripeptides of the form LysXxxLys (where Xxx is a naturally occurring amino acid) [17]. The central residue of this peptide was shown to bind in a buried cavity isolated from bulk solvent. Insertion of different residues into the cavity changed the water content but had a negligible effect on the shape of the cavity itself (Fig. 2). The

Figure 2



The superposition of four structures of the binding site of OppA for the central residue of tripeptides LysXxxLys. Xxx is Ala (yellow), Glu (green), Lys (blue) and Trp (red). The water molecules incorporated into the protein–ligand interface in each complex are colour-coded appropriately. On binding the tripeptide, the protein remains invariant (except for a slight twisting of the Glu residue, seen in the left of the binding site) but the water content of the binding site changes to accommodate the ligand. Note how the conserved waters in the complexes form distinct clusters (the water associated with the binding of the Trp sidechain (red) is present but obscured in the cluster in the top right hand corner).

binding constants for the LysXxxLys peptides are spread over less than two orders of magnitude (S.H. Sleigh, J.R.H. Tame, A.J. Wilkinson and J.E.L. unpublished data), demonstrating the remarkable level of adaptability that results from the inclusion of water molecules in the binding site. The inclusion of water molecules into the binding interface was also shown to improve the binding of the peptide (Table 1). When a tryptophan residue is replaced by an alanine in the central residue binding cavity, three extra water molecules are required to fill the void left by the removal of the bulky hydrophobic group. When these water molecules are included in the cavity, the affinity of the protein for the peptide increases (K_d(LysTrpLys) = 0.11 μM; K_d(LysAlaLys) = 0.06 μM). The thermodynamic effects of this change are a small entropy change ($\Delta\Delta S°_{Trp\rightarrow Ala}$ = 1.7 J mol^{-1} K^{-1}; $T\Delta\Delta S°_{Trp\rightarrow Ala}$ = 0.5 kJ mol^{-1}) and a slightly larger, favourable enthalpy change ($\Delta\Delta H°_{Trp\rightarrow Ala}$ = −1.2 kJ mol^{-1}). The same overall effect is observed when the central residue is changed from one that is positively charged to one that is negatively charged (Lys→Glu, see Table 1). In this case, the affinity is increased by an order of magnitude as a result of a significant decrease in ΔH° that more than compensates for the unfavourable entropic effect caused by the inclusion of one extra water molecule ($\Delta\Delta H°_{Lys\rightarrow Glu}$ = −15.5 kJ mol^{-1}; $T\Delta\Delta S°_{Lys\rightarrow Glu}$ = −8.9 kJ mol^{-1}). The water appears to cushion the effects of switching the charge of this residue of the tripeptide [17].

Table 1

Thermodynamic data for the binding of LysXxxLys tripeptides to OppA.

Peptide	Water*	K_B (10^6) M^{-1}	K_d μM	ΔH° kJ mol^{-1}	ΔG° kJ mol^{-1}	ΔS° J mol^{-1} K^{-1}
KAK	7	17.8±3.0	0.06	20.1±0.2	−41.4	206.3
KWK	4	9.0±2.5	0.11	21.3±0.6	−39.7	204.6
$\Delta_{Trp\rightarrow Ala}$	3	–	–	−1.2	−1.7	1.7
KEK	7	6.5±1.0	0.15	11.8±0.1	−38.9	170.3
KKK	6	0.5±0.1	2.0	27.3±0.6	−32.3	200.4
$\Delta_{Lys\rightarrow Glu}$	1	–	–	−15.5	−6.6	−30.1

*Waters observed in the pocket that accommodates the central residue sidechain. Data are for 25 °C and pH 7.0.

As emphasized above, the conclusions drawn from these types of studies have to be considered carefully. The thermodynamic parameters for binding that are measured can only be attributed to the addition of water molecules to the closed system if the initial state of all the molecules is equivalent in all cases. Although this is a reasonable assumption for the protein, it may not be for the peptides. For example, the initial entropy of the different central residues of the tripeptides will differ; these can be assessed from theoretical principles [18,19]. Furthermore, in the case of the Trp→Ala substitution the free LysTrpLys peptide is likely to have a significantly greater number of water molecules interacting with it than does the LysAlaLys peptide. These water molecules will be liberated on binding, producing a more favourable entropy. It is currently impossible to deconvolute the available data sufficiently to assess the individual value of each of these contributions.

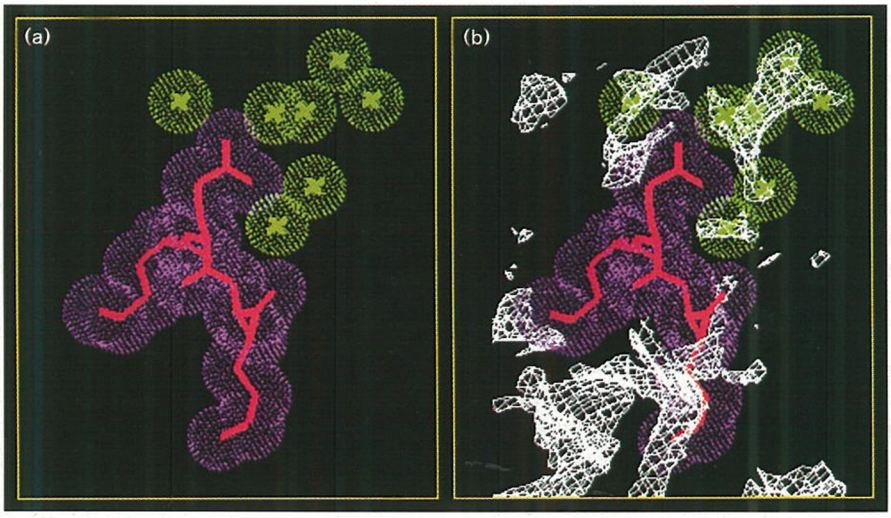
The high resolution structural detail (at < 1.4 Å) of the water molecules that mediate the OppA–peptide interactions show that, although the amino acid sidechain of the peptide is accommodated by changing the water content of the binding site, the water molecules that remain occupy the same positions from one ligand to another [17] (Fig. 2). Thus, in a sense, the water acts as part of the protein structure. Although the water changes the specificity of the protein, it does not do so by moving around to change the shape of the binding site, but by adopting a subset of the possible conserved positions. This implies that the water can make favourable hydrogen bonding arrangements within the binding interface that are not greatly affected by the insertion of a ligand. Analysis of the conserved water-binding sites indicates that they are indeed sites where the interactions are highly favourable in terms of hydrogen bond lengths, positioning and electrostatic and Lennard-Jones potentials (Fig. 3). Thus, the binding surface of OppA provides well defined, highly favourable sites for water.

Water in structure-based ligand design

So far, it has not proven possible to precisely determine the thermodynamic effects of incorporating a water molecule into a protein–ligand interface. Data from other sources, however, suggest that the inclusion of water molecules can increase the affinity of small-molecule binding. Using data from crystalline salt hydrates, the upper limits of the entropy (–9 kJ mol^{-1} at 298 K) and enthalpy (–16 kJ mol^{-1} at 298 K) of transferring a water molecule from bulk solvent and including it in an interface can be inferred [20–22]. Based on these values, it is clear that inclusion of a water molecule into an interface can provide a significant contribution to the free energy of an interaction (some idea of this can be gained from the upper limit values above at 298 K, which give an estimated free energy change of –7.0 kJ mol^{-1}, translating to an increase in affinity of more than one order of magnitude).

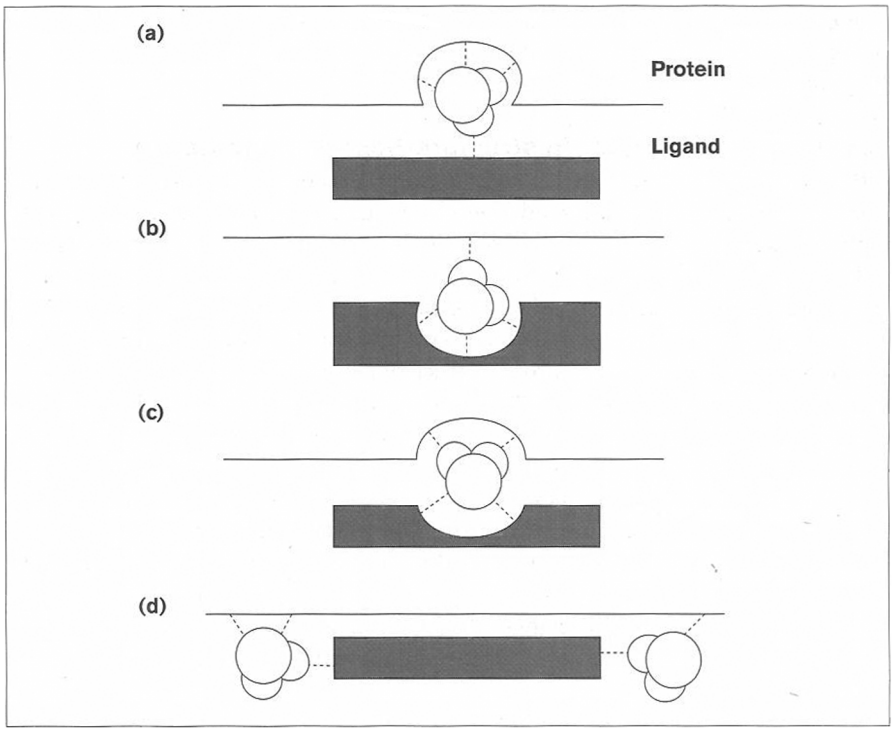
If the presence of a water molecule in the interface can give increased binding, an awareness of the uses of water molecules could assist in ligand design. It is important to assess in each case whether the free energy gain from including a water molecule into an interface will be sufficient to have a significant effect on the binding. It is likely that the conditions have to be carefully chosen for a water molecule to improve the binding characteristics of an interaction [23]. The combined effect of multiple water

Figure 3



Conserved water-binding sites are found at positions that would be predicted by considering electrostatic and Lennard-Jones potentials. **(a)** The structure of trilysine (red) and the van der Waals surfaces as found in the complex with OppA, showing the positions of water molecules (green) in the central residue binding cavity [17]. The OppA binding cavity has been removed. **(b)** The same structure as shown in (a), depicted with the energy surfaces derived from the program GRID [38,39]. The contour surfaces marked as grid-lines (white) correspond to positions of most favourable energy for water, based on Lennard-Jones and electrostatic functions within the binding cavity in the presence of the lysine sidechain. In this case the contours designate positions of interaction energy theoretically calculated as being < 50 kJ mol^{-1} on a scale of 0–67 kJ mol^{-1}. Interestingly, the contours fit very well to the positions of the water molecules actually found in the OppA–trilysine interface. This suggests that these water molecules are positioned in energetically favourable sites, where the enthalpic contribution to the free energy is largest.

Figure 4



Schematic of the modes in which water can be incorporated into a binding site. The water molecules and hydrogen bonds (broken lines) are arbitrarily positioned. The water can bind in several different ways; it may be largely bound to **(a)** the protein or **(b)** the ligand, or may bind approximately equally to both, either **(c)** in the binding site or **(d)** at the periphery of the binding site.

molecules in an interface has the potential to be large, however; for example, some water molecules at the binding site do not directly bridge the protein and the ligand, but may instead contribute to the stability of the complex by holding bridging water molecules in the right place through an aqueous network.

Water has the potential to be incorporated into protein–ligand interactions in several distinctly different ways (Fig. 4): water may be largely bound to (a) the protein or (b) the ligand, or it may be bound approximately equally to both, either (c) in the binding site or (d) at the periphery of the binding site. These arrangements are structurally very different and also require a different approach to ligand design. The first case, (a), is currently the most readily addressed since the majority of the information for identification of potential water binding sites is on the target protein structure.

To take advantage of the (b) and (c) binding modes we would need to know what constitutes a suitable binding site for water. With this information, one could reconstruct the binding site, either fully or partially, in the ligand design. If such an approach were possible, it could be more effective than an approach based on binding mode (a). Instead of attempting to find a water-binding site on the protein target and designing the ligand to interact with the protein-bound water, if the drug is designed to contain a water site (or part of a water site) then all that is required to complete the binding interaction is the protein. The chance of finding an optimal water binding site on the protein which provides the majority of the hydrogen bonds is low compared to finding a site which provides the complementary hydrogen bond(s) to an optimal water binding site designed into the ligand.

It has been suggested that it should also be productive to make use of (d)-type interactions [24], since in this approach water can be used to gain binding energy via interactions outside the original binding site. In this way, water can extend the binding region. Whatever the interaction mode chosen, it is essential to ask the following two questions: if a water molecule is already present on the binding surface in the free state of the interacting molecules, will the formation of the complex give sufficient additional hydrogen bonds (or sufficiently strengthen existing bonds) to give a favourable free energy change? Can the ligand surfaces be designed in such a way as to allow the water to make an optimal arrangement of hydrogen bonds?

In each case, the problem is tractable only when high resolution structural detail of the surface of the target protein is available. It is important to establish when a water molecule is contributing favourably to an interaction and should, therefore, be left in a binding site, or when it should be removed. One approach to this is to survey the distribution of water molecules in a broad sample of high resolution structural information in protein databases.

A large amount of data on water molecules on protein surfaces has been gathered, ostensibly with a view to addressing issues of the function, folding and structural refinement of these molecules [25–31]. Data from unbound proteins can also provide useful information regarding the inherent trends in water-binding sites. For example, water has a greater tendency to act as a hydrogen bond donor than as an acceptor since the most important sidechains for hydration are those of Asp and Glu (which accept hydrogen bonds from, on average, two water molecules per carboxylate group).

Unfortunately, much of the data gleaned from studies of the surface hydration of proteins is not useful for assessments of ligand binding. For instance, the observation that longer sidechains with the same functional groups appear to be more likely to bind water than shorter sidechains does not necessarily apply to ligand binding. This is because incorporating a longer sidechain into a binding site can lead to a greater entropic penalty; the degrees of freedom are reduced more when a longer sidechain is incorporated than when the structure is tightened on formation of a protein–ligand interface. Futhermore, it is somewhat difficult to judge whether a water molecule observed on a protein surface is in fact tightly bound since in many cases these are not highly conserved in different crystal forms of the same protein [1].

The water molecules found in ligand-binding sites appear to provide more useful information for the process of drug design, however [31–34]. The study of protein complexes has revealed some structurally important determinants of water-binding sites. The identification of these sites is based on criteria of surface-shape properties where water molecules are determined from atomic density, hydrogen bonding interactions, the mobility of water molecules in the binding site (from isotropic temperature-factors) and, in some cases, proximal atom hydrophilicity [31,34]. Although the database of protein–ligand structures with water molecules at the interface is small, some useful information has emerged. For instance, a large majority of water molecules involved in complex formation appear to make at least three hydrogen bonds [31] suggesting that this will be an important element of the successful design of a site that will incorporate water. Ultimately, through detailed examination of a large number of protein–ligand complexes it may be possible to determine the general characteristics of an optimal binding site, correlating the positions of water molecules with hydrogen bond donators and acceptors so as to give the maximum thermodynamic contribution to binding. These principles could then be applied to a design program that searches for these characteristics on target molecule surfaces (an algorithm of this type has in fact been developed very recently [34]), or enables them to be synthesised into designed ligands.

Recent developments in computer programs for drug design and structure prediction have also made a theoretical approach to the problem of identifying water sites feasible. Currently available computer programs which probe molecular surfaces with respect to topology, electrostatic and Lennard-Jones potential, [35,36] can be used to determine sites where water can reside (Fig. 3). Although these programs can give some qualitative idea of the 'desire' of water to be in a particular position, they generally suffer from the exclusion of entropic effects, giving an inherently incomplete thermodynamic picture. Theoretical determinations of the positioning of water molecules on protein surfaces are, in many cases, contradictory to the observations from database surveys [3]. For example, a model that incorporates surface tension effects in the solvent predicted that water will be less tightly bound in narrow crevices [37].

Conclusions

Water in protein–ligand interactions can function as an extension of protein structure, allowing varied ligands to be accommodated in a given binding site. Using water to adapt the binding site to a ligand has the potential to overcome some of the problems in structure-based drug strategies, and may provide a route to drugs with affinities that are higher than that of the natural substrate. The thermodynamic result of the inclusion of water is, at present, highly unpredictable and it is not yet clear whether the potential gains in free energy are large enough to make this approach realistic. Only by further characterization, both structural and thermodynamic, of water-binding sites in protein structures will we be able to assess whether adding water can lead to improvements in ligand binding.

Acknowledgements

I am grateful to Paul Sigler, Patrick Connelly, Debora Renzoni and the Protein Structure Group at the University of York for leading me into these uncharted waters, and to Leslie Kuhn and Fred Schwarz for their very important contributions to this review. I also thank Chris Johnson, Craig Morton, Terry Mulhern, Ronan O'Brien and Jeremy Tame for their comments and criticisms on the manuscript.

References

1. Otting, G., Liepinsh, E. & Wüthrich, K. (1991). Protein hydration in aqueous solution. *Science* **254**, 974–980.
2. Brunne, R.M., Liepinsh, E., Otting, G., Wüthrich, K. & van Gunsteren, W.F. (1993). Hydration of proteins: a comparison of experimental residence times of water molecules solvating the bovine pancreatic trypsin inhibitor with theoretical model calculations. *J. Mol. Biol.* **231**, 1040–1048.
3. Levitt, M. & Park, B.H. (1993). Water: now you see it, now you don't. *Structure* **1**, 223–226.
4. Meyer, E. (1992). Internal water molecules and H-bonding in biological macromolecules: a review of structural features with functional implications. *Protein Sci.* **1**, 1543–1562.
5. Karplus, P. A. & Faerman, C. (1994). Ordered water in macromolecular structure. *Curr. Opin. Struct. Biol.* **4**, 770–776.
6. Lam, P.Y.S., *et al.*, & Erickson-Viitanen, S. (1994). Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* **263**, 380–384.
7. Connelly, P.R., *et al.*, & Wilson, K.P. (1994). Enthalpy of hydrogen bond formation in a protein–ligand binding reaction. *Proc. Natl. Acad. Sci. USA.* **91**, 1964–1968.
8. Watson, K.A., *et al.*, & Papageorgiou, A. (1994). Design of inhibitors of glycogen phosphorylase: a study of α- and β-C-Glucosides and

1-thio-β-D-glucose compounds. *Biochemistry* **33**, 5745–5758.
9. Ringe, D. (1995). What makes a binding site a binding site? *Curr. Opin. Struct. Biol.* **5**, 825–829.
10. Bhat, T.N., Bentley, G.A., Fischmann, T.O., Boulot, G. & Poljak, R.J. (1990). Small rearrangements in structures of Fv and Fab fragments of antibody D1.3 on antigen binding. *Nature* **347**, 483–485.
11. Bhat, T.N., *et al.*, & Poljak, R.J. (1994). Bound water molecules and confromational stabilization help mediate an antigen–antibody association. *Proc. Natl. Acad. Sci. USA.* **91**, 1089–1093.
12. Ladbury, J.E., Wright, J.G., Sturtevant, J.M. & Sigler, P.B. (1994). A thermodynamic study of the *trp* repressor–operator interaction. *J. Mol. Biol.* **238**, 669–681.
13. Morton, C.J. & Ladbury, J.E. (1996). Water-mediated protein–DNA interactions: the relationship of thermodynamics to structural detail. *Protein Sci.* **5**, 2115–2118.
14. Goldbaum, F.A., Schwarz, F.P., Eisenstein, E., Cauerhff, A., Mariuzza, R. A. & Poljak, R.J. (1996). The effect of water on the association constant and the enthalpy of reaction between lysozyme and the specific antibodies D1.3 and D44.1. *J. Molec. Recogn.* **9**, 6–12.
15. Quiocho, F.A., Wilson, D.K. & Vyas, N.K. (1989). Substrate specificity and affinity of a protein modulated by bound water molecules. *Nature* **340**, 404–407.
16. Tame, J.R.H., *et al.*, & Wilkinson, A.J. (1994). The structural basis of sequence-independent peptide binding by OppA. *Science* **264**, 1578–1581.
17. Tame, J.R.H., Sleigh, S.H. Wilkinson, A.J. & Ladbury, J.E. (1996). The role of water in sequence-independent ligand binding by an oligopeptide transporter protein. *Nat. Struct. Biol.* 3, 998–1001.
18. Sternberg, M.J.E. & Chickos, J.S. (1994). Protein sidechain conformational entropy derived from fusion – a comparison data-comparison with other empirical scales. *Protein Eng.* **7**, 149–155.
19. Doig, A.J. & Sternberg, M.J.E. (1995). Sidechain entropy in protein folding. *Protein Sci.* **4**, 2247–2251.
20. Dunitz, J.D. (1994). The cost of bound water in crystals and biomolecules. *Science* **264**, 670.
21. Dunitz, J.D. (1995). Win some, lose some: enthalpy–entropy compensation in weak intermolecular interactions. *Chemistry & Biology* **2**, 709–712.
22. Connelly, P.R. (1997). In *Structural-based drug design: thermodynamics, modeling and strategy.* (Ladbury J.E. & Connelly, P.R. eds) R.G. Landes, Austin, TX. USA
23. Connelly, P.R. (1994). Acquisition and use of calorimetric data for prediction of the thermodynamics of ligand-binding and folding reactions of proteins. *Curr. Opin. Biotechnol.* **5**, 381–388.
24. Wang, H. & Ben-Naim, A. (1996). A possible involvement of solvent-induced interactions in drug design. *J. Med. Chem.* **39**, 1531–1539.
25. Thanki, N., Thornton, J., M., Goodfellow, J. M. (1988). Distribution of water around amino acid residues in proteins. *J. Mol. Biol.* **202**, 637–657.
26. Thanki, N., Umrania, Y., Thornton, J. & Goodfellow, J.M. (1991). Analysis of protein mainchain solvation as a function of secondary structure. *J. Mol. Biol.* **221**, 669–691.
27. Kuhn, L.A., Siani, M.A., Pique, M.E., Fischer, C.L., Getzoff, E.D. & Tainer, J.A. (1992). The interdependence of protein surface topography and bound water revealed by surface accessibility and fractal density measures. *J. Mol. Biol.* **228**, 13–22.
28. Kuhn, L.A., Swanson, C.A., Pique, M.E., Tainer, J.A. & Getzoff, E.D. (1995). Atomic and residue hydrophilicity in the context of folded protein structures. *Proteins* **23**, 536–547.
29. Baker, E.N. & Hubbard, R.E. (1984). Hydrogen bonding in globuar proteins. *Prog. Biophys. Molec. Biol.* **44**, 97–179.
30. Williams, M.A., Goodfellow, J.M. & Thornton, J.M. (1994). Buried waters and internal cavities in monomeric proteins. *Protein Sci.* **3**, 1224–1235.
31. Poornima, C.S. & Dean, P.M. (1995). Hydration in drug design. 1. Multiple hydrogen-bonding features of water molecules in mediating protein-ligand interactions. *J. Comp. Aided Mol. Design.* **9**, 500–512.
32. Poornima, C.S. & Dean, P.M. (1995). Hydration in drug design. 2. Influence of local site surface shape on water binding. *J. Comp.-Aided Mol. Design.* **9**, 513–520.
33. Poornima, C.S. & Dean, P.M. (1995). Hydration in drug design. 3. Conserved water molecules at the ligand-binding sites of homologous proteins. *J. Comp. Aided Mol. Design.* **9**, 521–531.
34. Raymer. M.L., Sanschagrin, P.C., Punch, W.F., Venkataraman, S., Goodman, E.D. & Kuhn, L.A. (1996). Predicting conserved water-mediated and polar ligand interactions in proteins using a K-nearest-neighbor genetic algorithm. *J. Mol. Biol.*, in press.
35. Blundell, T.L. (1996). Structure-based drug design. *Nature* **384**, 23–26.
36. Verlinde, C.L.M.J. & Hol, W.G.J. (1994). Structure-based drug design: progress, results and challenges. *Structure* **2**, 577–587.
37. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296.
38. Goodford, P.J. (1985). A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **28**, 849–857.
39. Wade, R.C., Clark, K.J. & Goodford, P.J. (1993). Further developments of hydrogen bond functions for use in determining favourable binding sites of known structure. *J. Med. Chem.* **36**, 140–147.