Water at biomolecular binding interfaces

Zheng Li and Themis Lazaridis*

Received 29th August 2006, Accepted 8th November 2006
First published as an Advance Article on the web 24th November 2006
DOI: 10.1039/b612449f

Water molecules are often found at the binding interface of biomolecular complexes mediating the interaction between polar groups *via* hydrogen bonds, or simply filling space providing van der Waals interactions. Recent studies have demonstrated the importance of taking such water molecules into account in docking and binding affinity prediction. Here, we review the recent experimental and theoretical work aimed at quantifying the influence of interfacial water on the thermodynamic properties of binding. We highlight especially our recent results obtained by inhomogeneous fluid solvation theory in several systems and the prediction of the thermodynamic consequences of displacement of the bound water molecule by ligand modification. Finally, we discuss possible directions for further progress in this field.

Introduction

Crystal structures of biomolecules typically reveal many water molecules on the biomolecular surface, in crevices, and sometimes buried within the biomolecular interior or at binding interfaces. The water molecules at binding interfaces can modify the shape and flexibility of the protein binding site, improving the steric complementarity between protein and ligand, or mediate the binding between the biomolecules with a hydrogen bond network.[1–13] Analysis of 109 protein–DNA complexes showed that 6% of the crystallographic water mediates the binding.[14] For instance, in the trp-repressor/operator complex, in addition to the direct interactions between the phosphates of DNA and the protein, six water-mediated polar contacts to the bases were observed,[1] which are the determinants of specificity in this system.[6] An analysis of 75 protein–protein complexes showed, on average, one water molecule per 100 Å^2 of interface area.[15] A more recent analysis found that water-mediated polar interactions are as abundant as direct protein–protein hydrogen bonds.[16]

Several previous reviews are relevant to this topic. Levitt and Park surveyed the available information on the location and dynamics of water molecules in and around protein structures.[17] Ladbury focused on the role of water in protein–ligand binding and its potential applications in drug design.[18] Janin also covered water-mediated interactions in protein–protein and protein–DNA complexes and their effect on affinity and specificity.[19] Cozzini *et al*. described different computational methods used to predict protein–ligand binding free energy with some emphasis on the contributions of water molecules.[20] Jayaram and Jain reviewed the role of water in protein–DNA recognition.[21]

Despite much work, many issues regarding interfacial water molecules remain unclear: Are water-mediated interactions equally strong as direct interactions? Do they provide more or less specificity? How does the presence of interfacial water molecules affect the enthalpy, entropy, and heat capacity of binding? What is the best way to incorporate water-mediated interactions in docking and drug design? Here, we will review recent work that addresses some of these questions. It is not meant to be a comprehensive review, but an exposition of the diversity of approaches taken to tackle this problem.

Department of Chemistry, City College of New York/CUNY, Convent Ave. & 138th Street, New York, New York 10031.
E-mail: tlazaridis@ccny.cuny.edu; Fax: +1-212-650-6107;
Tel: +1-212-650-8364

Experimental studies of the thermodynamics of interfacial water

The standard approach for studying the effect of interfacial water experimentally is to introduce modifications to the ligand or the protein that will introduce or displace it, and measure the effect on binding thermodynamics. In some cases, ligands designed to displace the water molecules exhibit higher binding affinity. The best known example can be found among HIV-1 protease inhibitors. Several water molecules were observed crystallographically at the binding interface between HIV-1 protease and KNI-272 or its analogs.[22–26] Cyclic urea inhibitors designed to displace and mimic the interactions of one of the bound water molecules were found to bind more strongly to the protein.[27–29] However, the cyclic urea inhibitors are entirely different from KNI-272, and therefore it is not certain that the gain in binding affinity really comes from water displacement. More clear-cut is the example of scytalone dehydratase. The crystal structure of this protein complexed with a salicylamide inhibitor revealed two water molecules at the binding interface. Analogs of this inhibitor with an additional nitrile group were found to displace one of the crystallographic water molecules and to have higher inhibitory potency.[30]

Crystal structures of OppA–dipeptide complexes revealed several ordered water molecules mediating the interactions between ligand and protein.[7,10] These water molecules are displaced when the ligand changes from dipeptide to tri- and tetrapeptides. Isothermal titration calorimetry measurements of the binding of the peptides with different lengths to OppA indicate that the dipeptide is bound with about 60-fold lower affinity than related tri- and tetrapeptides.[10] The higher affinity for tri- or tetrapeptides *versus* dipeptides is due to a more

favorable binding entropy, which is only partly offset by a more unfavorable enthalpy. The favorable entropy was suggested to result from the displacement of the ordered water.[10]

The streptavidin–HABA (2-[(4′-hydroxyphenyl)-azo]-benzoate) crystal structure shows an immobilized water molecule bridging two hydroxyl groups. Analogs of HABA with additional aliphatic groups displaced this water molecule and exhibited higher binding affinity, due to a more favorable binding entropy. The enhanced affinity was attributed to water displacement and ligand flexibility.[31]

Higher binding affinity by displacement of interfacial water molecules can also be achieved by protein mutations. Crystallographic studies revealed that two water molecules bound to the tyrosine-82 hydroxyl group in unliganded wild type FKBP-12 are displaced upon formation of a complex with FK506.[32] In the Y82F mutant of the protein, no ordered water molecules are observed. Thermodynamic measurements showed that the enthalpy of binding is 4.2 kcal mol^{-1} more negative for the mutant, but partial compensation by entropy led to a slightly more negative binding free energy (0.6 kcal mol^{-1}) for the mutant.[33] A simulation study of the E60A mutant found a reorientation of the water and led to an intuitive prediction that the mutant will exhibit reduced binding affinity.[34]

In the examples given above, water displacement was found to be beneficial for binding. In other cases, however, the displacement of interfacial water weakens protein–ligand binding. The structure of the complex of Concanavalin A (Con A) with a trimannoside shows a conserved water molecule bridging the protein and the ligand with several hydrogen bonds.[35] To study the role of this water molecule in binding, an analog of the trimannoside was designed to replace the hydroxyl at C-2 of the central mannose with a hydroxyethyl group, which was expected to displace and mimic the interactions of this water molecule with the protein.[36] Calorimetric measurements showed that this displacement leads to a more favorable binding entropy but also a more unfavorable and larger enthalpy term, and thus weaker binding.[36] A crystallographic study of cyclosporin A bound to cyclophilin A revealed a cluster of water molecules mediating the protein–ligand binding.[3] One of the bound water molecules was found displaced and others reorganized in going from cyclosporin A to (5-hydroxynorvaline)-2-cyclosporin.[37] The binding affinity decreased by 1.3 kcal mol^{-1}.

Sharrow *et al.* studied the thermodynamic contribution of the hydrogen bond network formed between the interfacial waters and the mouse urinary protein and its ligands, employing calorimetry and mutational studies.[38] Mutation of a protein Y to F disrupts a hydrogen bond between one bound water molecule and the protein and leads to a loss of enthalpy, which is only partially compensated by a favorable entropy change.

Several experimental thermodynamic studies investigated the role of interfacial water in antibody–antigen binding using site-directed mutagenesis. In mutational studies of the antilysozyme–antibody D1.3, the interfacial waters in two of the mutant complexes (Y50S and W92D) appear to reorganize to partially alleviate the loss of antibody–antigen interactions.[39] Still, the binding is weakened. Yokota *et al.* constructed mutants of an antilysozyme antibody changing Y to F and S to A.[40] Crystallographic studies showed that the interfacial water molecules mediating the binding were not affected by the mutations. Instead, the mutations create an unfavorable gap between the waters and the antibody, which induces the loss of binding affinity due to loss of binding enthalpy. An Arg-to-Lys mutant of a different antilysozyme antibody introduces an additional water molecule at the binding interface and leads to a 1000-fold loss of binding affinity.[41]

In other studies, the introduction of water molecules at the interface by mutation was found to be energetically neutral. Watson *et al.* found similar binding affinities for different inhibitors of glycogen phosphorylase regardless of the presence of water-mediated interactions.[42] Dall'Acqua *et al.* found that an Asp18 → Ala mutation introduced 3 water molecules at the interface but had negligible impact on the binding affinity.[43]

Mutagenesis has also been used to explore the contribution of bound water molecules to protein stability. One finding was that incorporation of water into an internal protein cavity is energy-neutral.[44] Another study found that a water-mediated interaction is equivalent in the folded and unfolded states.[45] Yet another group concluded that protein–water hydrogen bonds contribute less to stability than protein–protein hydrogen bonds.[46,47]

Denisov *et al.* performed an entropy analysis on experimental bound waters in BPTI. The amplitude and anisotropy of three water molecules in the protein cavities were assessed based on the nuclear magnetic relaxation dispersion data. The results indicated that the binding of some of these water molecules is entropically driven, even when they formed three or four hydrogen bonds with the protein.[48]

Several experimental thermodynamic studies focused on water mediating protein–DNA interactions. Osmotic stress has shown the importance of water molecules for specific DNA recognition by EcoRI.[49] The same approach applied to the complexes of the tryptophan and lac-repressor with their operators showed a number of water molecules at the protein–DNA interface, in good agreement with the number determined crystallographically.[50,51]

Isothermal titration calorimetry has been used to measure the thermodynamics of a TATA-box binding protein with TATA-box and found that the bridging water molecules at highly polar biomolecular surfaces, as well as those in the binding interface, contribute substantially to the negative heat capacity change upon protein–DNA association.[52] This complements earlier results on the contribution of interfacial water molecules to the heat capacity of binding.[53,54]

Prediction of interfacial water in biological systems

The identification and localization of water molecules in the crystal structures of biomolecules can be problematic due to crystallographic uncertainties. Some of the interfacial water in biological systems can be artifacts, particularly when they do not hydrogen-bond to the biomolecules.[55] Thus, in many instances, one needs to determine the location of bound water molecules theoretically. Several methods have been developed for this purpose. AQUARIUS[56] is a knowledge-based

approach specially aimed at identifying water sites in proteins from the electron density maps generated by protein crystallography. GRID was developed to identify binding sites for water and other functional groups.[57] It involves calculating the interaction energy between the probe and the protein using an empirical energy function based on Lennard-Jones, electrostatic and hydrogen-bond terms. GRID has been reported to perform well in predicting protein–ligand interfacial water.[58–60] MCSS determines the location of functional groups *via* molecular dynamics simulations of multiple copies of such groups;[61] SuperStar is an empirical method based on the thermodynamics of small molecules and their crystal structures that treats protein binding site residues as small molecules or small structure fragments.[62] CS-Map predicts the most favorable binding positions of probes on a protein surface by calculating an interaction potential that accounts for van der Waals, electrostatic, and solvation contributions. The last two were calculated by an implicit solvation model.[63] A new version of the Fold-X force field allows the prediction of the position of bound water molecules, and can pick up 76% of water molecules interacting with at least two polar atoms of biomolecules in crystal structures with an average uncertainty of 0.8 Å.[64]

Garcia-Sosa *et al.* performed a multivariate regression analysis to establish a statistical correlation between the structural properties of water molecules in the binding site of a free protein crystal structure and the probability of observing the water molecules in the same location in the crystal structure of the complex.[65] The B-factor, the solvent contact surface area, the total hydrogen bond energy and the number of protein water contacts were found to correlate with the conservation of an ordered water molecule upon complex formation.

Interfacial water in scoring, docking, and ligand design

Interfacial water molecules have begun to be taken into consideration in *de novo* drug design, protein–ligand docking programs and the development of scoring functions for prediction of binding affinity.[60,65–75]

A computational approach based on the HINT empirical energy function was developed to study the binding and energetic roles of water molecules located in the binding interface of 23 HIV-1 protease–ligand complexes.[60] The correlation coefficients between the calculated HINT scores and the experimental binding affinity of these complexes were significantly improved by taking the contributions of water into account. HINT was also used to estimate the energetics of water molecules in a wider variety of binding sites.[73] The water molecules bridging protein and ligand were found to have the largest binding energies (-1.3 kcal mol^{-1}). The HINT score together with the number and quality of hydrogen bonds of water molecules at protein active sites were suggested as criteria for predicting whether they would be favorably displaced by the ligand.

Inclusion of the interfacial water molecules in deriving both binding site structure- and ligand-based pharmacophores shows a profound effect on ligand design.[70] Wang *et al.* used the number of bridging water molecules in a QSAR-type predictor of the binding affinity of tyrosine phosphatase inhibitors.[68] Mancera found that inclusion of explicit water greatly facilitated *de novo* ligand design for bacterial neuraminidase.[75]

A few docking programs have developed ways to include discrete water molecules. Rarey *et al.* developed an approach for taking water into account in the docking program Flex-X. The water molecule is represented by a particle that can form, at most, four hydrogen bonds. Possible positions for a water molecule in the protein binding site are precomputed.[66] The geometry of the hydrogen bonds and the steric constraints resulting from the water is used to optimize the orientation of the ligand during the docking process. Osterberg treated water as a variable in an ensemble of protein structures that are combined to produce a single target representation for use with Autodock.[67] Schnecke and Kuhn predicted possible water molecules from the crystal structure of the ligand-free protein and considered their effects on molecular docking, allowing them to interact with the ligands or incorporating a desolvation penalty into the scoring function when they are displaced by ligand atoms.[76] Verdonk *et al.* extended the program GOLD to include the bound water molecules in protein–ligand docking. In this approach, the water molecules in all-atom representation are allowed to rotate and switch between the bound and displaced status during the docking protocol. The loss of rigid-body entropy of water is added to the scoring function as a penalty when water is present, thus rewarding water displacement.[72] The program GLIDE docks explicit water molecules into the binding site and employs empirical scoring terms to account for their impact.[77] Moitessier *et al.* made crystallographic water molecules displaceable by truncating the water–ligand short range repulsion.[78]

Significant improvements in docking performance by considering the contributions from interfacial water have been reported.[66,67,71,72,76,79,80] In general, when crystallographic or even predicted water molecules are retained in protein structures, docking performance is improved.[71,72,76,79,80] These water molecules can usually accept two and donate two hydrogen bonds, at most, in the complex. However, other studies found that including water molecules had little effect on the quality of docking.[81]

A "solvated rotamer" approach was introduced by Jiang *et al.* to account for water-mediated hydrogen bonds in designing protein–protein interfaces.[82] They expanded a rotamer library by building fixed water molecules onto polar-side chain atoms, as well as backbone atoms. These rotamers were used to calculate protein–water interactions that include a Lennard-Jones repulsive potential and hydrogen bonding potential. It appears that the simple water-mediated hydrogen-bonding model can significantly improve prediction of amino acid identities. This model can also be applied to calculate protein–protein interaction energies. However, the predicted water molecules are not as helpful as crystallographic ones in predicting interaction energies. One problem with the approach is the increased combinatorial complexity of the new rotamer library.

Wolynes and coworkers derived direct and water mediated contact potentials for proteins and found that the combination of the two performs best in discriminating native

protein–protein interfaces.[83] In more recent work, water-mediated contacts along with a new one-dimensional burial term improved the quality of protein structure prediction, especially for larger proteins.[84] Water-mediated potentials also improved the quality of β-sheets in α/β protein structure prediction.[85]

Theoretical studies of the thermodynamics of interfacial water

The simplest estimate of the entropy cost of tying up a water molecule into a protein cavity was given by Dunitz, based on the entropies of inorganic salts.[86] He gave an upper bound of the entropy cost as 7 kcal mol^{-1} (about 2 kcal mol^{-1} at 298 K). Similarly, an estimated upper bound of the enthalpy gain of water ordering was given as -3.8 kcal mol^{-1}.[18] Combining these two estimates one arrives at -1.8 kcal mol^{-1} for the free energy of inclusion of a single water molecule in a protein cavity or binding interface. Nevertheless, these are simple, generic estimates which are of limited utility in practical drug-design situations.

Covell and Wallqvist used a knowledge-based energy function, including crystal waters, to compute the effect of mutations at protein–protein interfaces on binding affinity. 25% of the total binding affinity was calculated to arise from the crystallographic water molecules at the interface.[87]

Petrone and Garcia calculated the contributions of bound water molecules in Major Histocompatibility Complex (MHC)—peptide binding. This is a very interesting system because it is characterized by low specificity (promiscuity). They used an approximate method for calculating the chemical potential of water based on a Gaussian model for the water interaction energy distribution. They found that water that simply fills space can bind at the interface with a net gain in entropy relative to the bulk.[88]

An entirely different methodology—normal mode analysis—was used by Fischer *et al*. to measure the effect of an isolated and well-ordered water molecule on the vibrational entropy of Bovine Pancreatic Trypsin Inhibitor (BPTI).[89] The increase in entropy upon addition of the water molecule was decomposed into the librational contributions of the water (9.4 kcal mol^{-1}) and an increase in protein flexibility (4 kcal mol^{-1}). This work shows that the low frequency vibrational modes of the water are important for the protein–ligand binding processes. Later work found that the water–protein electrostatic interactions are key to this phenomenon.[90]

On a totally different theory level, a quantum chemical *ab initio* method was used to study the geometry and energetics of six water-mediated base pairs in RNA, including their geometries and interaction energies.[91] In most cases, the geometry of the water-mediated base pairs was similar to that observed in crystal structures. The energies of the water-mediated and direct base pairs were compared. It was found that the absolute energy per hydrogen bond was higher for water-mediated interactions in all cases except for the UA pair. Of course, these studies only give interaction energies and do not include entropic effects, unlike the other methods mentioned above. Another quantum mechanical study aimed to quantify the differential binding energies of water *versus* carbohydrates to concanavalin A, where the carbohydrates were modeled by methanol and the protein by a limited number of amino acid side chains.[92] The binding energies of water and methanol to a cluster including three key amino acid side chains of concanavalin A and the key water were found similar to each other.

Free energy perturbation and thermodynamic integration studies of interfacial water

Free Energy Perturbation (FEP) and Thermodynamic Integration (TI) are rigorous methods for computing the relative free energy between two thermodynamic states. Over the past two decades, they have been applied to a wide variety of biological problems, including molecular recognition and protein stability. Studies on the contributions of ordered water molecules to free energy employing FEP or TI usually involve a hypothetical process by which a bound water molecule is transferred from the binding site or cavity of the protein to the bulk.[93–97] Applying these techniques, Wade *et al*. found that the free energy of transferring water molecules from bulk solvent into protein cavities is positive for empty cavities and negative for cavities containing water molecules.[93] Roux *et al*. used free energy perturbation to study the stability of water molecules in bacteriorhodopsin.[94] Tashiro and Stuchebrukhov calculated the thermodynamic properties of water molecules in a hydrophobic cavity around the catalytic center of cytochrome c oxidase employing TI.[98] Zhang and Hermans studied the energetics of buried water molecules in the subtilisin Carlsberg-eglin C complex by calculating the free energies of transferring water molecules into the protein cavities employing molecular dynamics simulations and a harmonic restraint potential for each water.[95] Their results show that such free energies for the hydrated cavities (with a negative value up to -8.3 kcal mol^{-1}) are dramatically different from those for the empty cavities (with a positive value up to $+14.6$ kcal mol^{-1}). Using FEP, Olano and Rick calculated the free energies of transferring water molecules from pure liquid into two different protein cavities, *i.e*., BPTI (hydrophilic) and the I76A mutant of barnase (hydrophobic), which were obtained as -4.7 and $+4.7$ kcal mol^{-1} at 298 K, respectively.[97] They also calculated the entropy of transfer and found it to be unfavorable for the polar cavity and favorable for the non-polar one.

Hamelberg and McCammon studied interfacial water molecules in two biological systems using the double-decoupling approach.[96] In this approach, the binding thermodynamics of two molecules (A and B) is calculated from the free energy of transferring B from solution to the gas phase and from the complex AB in solution to the gas phase.[99] The free energy change of the first process is calculated by simply removing B from the solution applying free energy perturbation, while the other one is evaluated by thermodynamic integration applying constraints on the coordinates of B. They applied this method to the crucial water molecule bridging HIV-1 protease and the ligand KNI-272 and a similar water molecule in the complex of trypsin with benzyldiamine. The free energy of transfer of water into the protein–ligand complex was calculated to be -3.1 kcal mol^{-1} in the first system and -1.9 in the second.[96]

Archontis *et al*. used free energy simulations to study the binding affinity between Glycogen Phosphorylase (GP) and its

inhibitors: glucopyranose spirohydantoin (hydan) and its analogues (m-hydan and n-hydan).[100] Both this study and crystallography show that in the binding interface of the GP–hydan complex one water molecule prefers to stay at a well defined position, while in the GP–m-hydan (or n-hydan) complexes this water prefers another position. The free energy profile for the translocation of the water between the two positions showed that the calculated binding free energies are sensitive to the preference of this interfacial water to occupy the two positions. A free energy decomposition analysis employing the thermodynamic integration method along with a restraining potential on the water showed the free energy of moving a water molecule from one position to another in the GP–hydan complex to be 3.0 kcal mol^{-1}.[100]

A method based on free energy simulations has been proposed to compute the number of water molecules in the interfaces of protein–protein complexes.[101] This approach determines the optimal number of water molecules from the condition that the chemical potential of water in the cavity be equal to that in the bulk. It has been applied to determine the water content in 211 interfacial cavities in 26 antigen–antibody complexes.

Thermodynamic contributions of interfacial water revealed by IFST

In this section, we briefly describe the Inhomogeneous Fluid Solvation Theory (IFST)[102,103] and its recent applications to the study of interfacial water molecules.[104–106] In IFST, the solvation free energy is decomposed into four terms: the solute–solvent energy (E_{sw}), the solute–solvent entropy (S_{sw}), the solvent reorganization energy (ΔE_{ww}) and solvent reorganization entropy (ΔS_{ww}). Each of these can be expressed as integrals of the position-dependent solute–solvent and solvent–solvent correlation functions ($g_{sw}(\mathbf{r},\omega)$ and $g_{ww}(\mathbf{r}, \mathbf{r}', \omega, \omega')$).

$$S_{sw} = -k\rho/\Omega \int g_{sw}(\mathbf{r},\omega) \ln g_{sw}(\mathbf{r},\omega)\, d\mathbf{r}\, d\omega. \quad (1)$$

$$E_{sw} = \rho/\Omega \int g_{sw}(\mathbf{r},\omega) u_{sw}(\mathbf{r},\omega)\, d\mathbf{r}\, d\omega. \quad (2)$$

$$\Delta S_{ww} = -\frac{1}{2} k \frac{\rho^2}{\Omega^2} \int g_{sw}(\mathbf{r},\omega)[g_{sw}(\mathbf{r}',\omega')\{g^{inh}_{ww}(\mathbf{r},\mathbf{r}',\omega,\omega') \ln g^{inh}_{ww}(\mathbf{r},\mathbf{r}',\omega,\omega') - g^{inh}_{ww}(\mathbf{r},\mathbf{r}',\omega,\omega') + 1\} - \{g^{o}_{ww}(R,\omega^{rel}) \ln g^{o}_{ww}(R,\omega^{rel}) - g^{o}_{ww}(R,\omega^{rel}) + 1\}] d\mathbf{r} d\mathbf{r}' d\omega d\omega' \quad (3)$$

$$\Delta E_{ww} = -\frac{1}{2} \frac{\rho^2}{\Omega^2} \int g_{sw}(\mathbf{r},\omega)[g_{sw}(\mathbf{r}',\omega') g^{inh}_{ww}(\mathbf{r},\mathbf{r}',\omega,\omega') - g^{o}_{ww}(R,\omega^{rel})] u_{ww}(R,\omega^{rel})\, d\mathbf{r} d\mathbf{r}' d\omega d\omega' \quad (4)$$

where k is Boltzmann's constant, ρ is the density of bulk solvent, $\mathbf{r}$ and $\mathbf{r}'$ denote the position of two water molecules; ω, and ω' denote the orientation of these two water molecules with respect to the solute, each of which is expressed as three Euler angles; Ω is the integral over ω ($\Omega = 8\pi^2$), R is the distance between two water molecules, ω^{rel} are the five angles that describe the relative orientation of two water molecules, and $\Omega^{rel} = \int d\omega^{rel} = 32\pi^3$; $g^{o}_{ww}(R, \omega^{rel})$ and $g^{inh}_{ww}(\mathbf{r},\mathbf{r}',\omega,\omega')$ are the solvent–solvent correlation functions in the pure solvent and in the complex, respectively; $u_{sw}(\mathbf{r}, \omega)$ and $u_{ww}(R,\omega^{rel})$ are water–solute and water–water potentials, respectively.

The solute–solvent correlation function $g_{sw}(\mathbf{r}, \omega)$ is zero over the regions occupied by the solute, and thus all contributions to the solvation energy and entropy come from regions occupied by the solvent. This approach allows one to focus on the vicinity of the solute, especially the biomolecular interface region, where the largest contributions to solvation thermodynamics originate, whereas the standard homogeneous correlation functions are averages over the entire body of the fluid. The advantage of IFST is that it allows a rigorous decomposition of the solvation free energy into contributions from different solvent regions[107] and provides unprecedented detail: the contribution of each interfacial water molecule to the enthalpy and entropy can be obtained no matter how tightly it binds to the biomolecules and whether or not it is fully buried. In contrast, the removal of water molecules from surface sites upon FE simulations is not meaningful, because another water molecule will immediately take their place. The disadvantages of IFST are that it involves numerous approximations, and requires the evaluation of complicated integrals, which makes it less "user friendly".[108]

IFST has been applied to the isolated water molecules in HIV-1 protease–KNI-272[104] and Concanavalin A–trimannoside 1 complexes,[105] and to interfacial water clusters in complexes of Cyclophilin A with cyclosporin A and (5-hydroxynorvaline)-2-cyclosporin.[106] The approach requires performing MD simulations of the water with the protein and ligand fixed, calculation of solute–solvent and solvent–solvent correlation functions, and evaluation of the IFST integrals to obtain the thermodynamic properties. Table 1 summarizes the results obtained in all three systems. The solvation entropy (ΔS_{solv}) of the waters ranges from −9.8 kcal mol^{-1} to 0, reflecting their different degree of ordering. Thus, the entropic cost of binding for some highly ordered water molecules at protein–ligand binding interfaces can be larger than the upper bound (7 kcal mol^{-1}) proposed by Dunitz.[104] However, the entropic penalty is always outweighed by the favorable enthalpy change, which mainly originates from the direct interactions between the water and the protein–ligand, leading to a favorable contribution to the solvation free energy. In some cases, the contribution is large, *e.g.*, the contribution of the interfacial water in HIV-1 protease–KNI-272 is −15.2 kcal mol^{-1}.[104] In other cases, the contribution is very small, *e.g.*, the contribution of water 133 in the complex of Cyclophilin A–cyclosporin A was calculated as −1.9 kcal mol^{-1}.[106] Among the interfacial water molecules we studied, water 133 is the only one forming no hydrogen bond with the protein or ligand.

Based on the values of the thermodynamics parameters of water at different temperatures, we found that the contribution of several bound water molecules to the heat capacity is negative. The negative value shows that the increase in pure water energy and entropy with temperature is faster than the increase of the protein–ligand interaction energy and entropy.

The values of the solvation free energy for many of the interfacial water molecules we calculated are much larger than the value of standard free energy of transferring the water

Table 1 Thermodynamic parameters of the bound water molecules in the HIV-1 protease–inhibitor complex, Con A–trimannoside complex, cyclophilin A–cyclosporin complexes. Adapted from ref. 106

Complex	Bound water*	E_{ww}	E_{sw}	ΔE_{solv}	S_{ww}	S_{sw}	$-T\Delta S_{solv}$	ΔG_{solv}	ΔC_p
HIV-1 protease	WTR1	0	−28.2	−18.1	0	−25.0	+2.9	−15.2	−11 ± 1
Con A–trimannoside **1**	WTR1	0	−30.2	−19.2	0	−22.1	+2.0	−17.2	−10 ± 3
Cyclophilin A–CsA **1**	WTR5	−3.8	−25.6	−17.4	−5.5	−18.3	+2.6	−14.8	—
	WTR6	−15.4	−8.3	−5.9	−3.4	−13.4	+0.5	−5.4	—
	WTR7	−5.1	−14.1	−6.6	−3.8	−12.2	+0.2	−6.4	—
	WTR133	−15.6	−3.7	−1.4	−3.5	−10.1	−0.5	−1.9	—
Cyclophilin A–CsA **2**	WTR5	+0.4	−23.0	−12.7	+0.2	−20.9	+1.6	−11.1	—
	WTR6	−11.0	−17.7	−13.1	−1.2	−20.5	+2.0	−11.1	—
	WTR7	+3.4	−24.3	−12.5	−0.2	−21.3	+1.9	−10.6	—

Note: Units for the energy are kcal mol^{-1}; Units for the entropy and heat capacity are cal mol^{-1} K^{-1}. ΔE_{solv} and ΔS_{solv} are calculated by subtracting the energy and entropy of pure water from the sum of the solute–solvent and solvent–solvent terms. The bound waters listed here are the key water molecules we previously studied in different systems.[104–106]

molecule from the bulk to the binding interface obtained by Hamelberg and McCammon. For example, for the water in HIV-1 protease they obtained −3.1 kcal mol^{-1}, while we obtained −15.2 kcal mol^{-1}. It should be clarified, however, that the results of these two methods are not comparable. FEP simulates the actual removal of a water molecule, while IFST does not; it performs a decomposition of the free energy of the system and assigns a certain value to each water molecule. In the case of FEP, the removal of the water molecule may be accompanied by conformational changes of the protein and ligand. The contribution of such changes is not included in the IFST result. For example, in the HIV-1–KNI272 complex, 1600 steps of energy minimization showed that the protein–ligand interaction energy is 4.2 kcal mol^{-1} more negative when the key water molecule is not present. Another factor that may contribute to the large values for the free energy contribution of interfacial waters by IFST is that, in the calculations, the protein and ligand are fixed at their optimized configuration. Thermal motion of the protein and ligand is expected to reduce their interaction energy with the bound water. In principle, the IFST approach should be applied to a thermal ensemble of protein–ligand configurations.

To connect the thermodynamic contributions of interfacial water to the protein–ligand binding thermodynamics, we studied two pairs of similar complexes, *i.e.*, Con A–trimannoside complexes[105] and the complexes of Cyclophilin A with cyclosporin A and analogues,[106] where the thermodynamic consequences of displacing key water molecules have been measured experimentally. For each pair, the ligand structures and the binding modes are very similar to each other (see Fig. 1 for Con A). As a result, we can approximately attribute the binding free energy differences to only a few factors, namely, the contributions of the interfacial water, protein–ligand interaction energies, ligand desolvation and entropy of conformational restriction.

The calculated contributions for Con A are given in Table 2. To calculate the difference of the interaction energy of the two trimannosides with concanavalin A, we assumed that only the central mannose contributes to the different interaction energy. The difference in interaction energy of the C-2 substituents of the two ligands with the protein is −18.4 kcal mol^{-1}, which is comparable to a favorable contribution to the solvation free energy given by the bound water molecule in Con A–trimannoside **1** complex (−17.2 kcal mol^{-1}). To calculate the different desolvation enthalpy and entropy of the ligands, we assumed that the solvation effects of polar and apolar groups of trimannoside **1** and **2** are additive and proportional to their solvent accessible surfaces. We also calculated the entropy of conformational restriction of the hydroxyethyl side chain of trimannoside **2** by comparing the distributions of the dihedral angles at the ligand side chain in the bound and free state. Similar calculations of these factors were performed for the complexes of Cyclophilin A with cyclosporin A and analogues. The final calculated results for the different binding affinity were both in good agreement with the experimental data.

Outlook

From the published experimental and theoretical studies on interfacial water, the following conclusions can be reached: (a) interfacial water involved in hydrogen bonds with the protein and ligand makes a negative contribution to the entropy, enthalpy, and heat capacity of binding. However, water that does not form hydrogen bonds can have higher entropy than in the bulk; (b) Water-mediated interactions can be as strong as direct interactions. Displacement of bound water by ligand modification can increase or decrease binding affinity

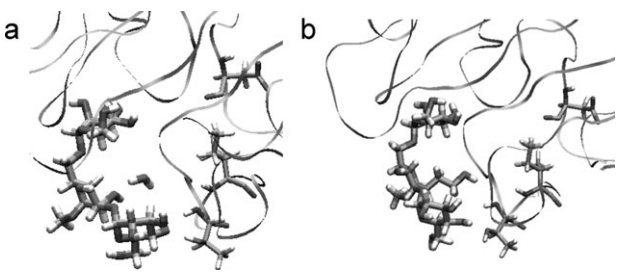

Fig. 1 Binding of trimannoside **1** and **2** to Con A (structures obtained after 8 ns MD simulation). a: the ordered water molecule in the binding interface of Con A–trimannoside **1** (in thicker sticks) complex and the 3 residues (Asp14, Asn16, and Arg228 in thinner sticks) at the protein binding site; b shows the similar structure of Con A–trimannoside **2** complex.

Table 2 Contributions to the enthalpy, entropy and heat capacity of the binding of two trimannosides to concanavalin A. Adapted from ref. 105

	Con A–trimannoside **1**		Con A–trimannoside **2**		Difference between **1** and **2**	
Temp.	300 K	330 K	300 K	330 K	300 K	330 K
$E_{\text{L2OH-P}}$	-9.4 ± 0.1	-9.3 ± 0.03	-27.8 ± 0.02	-27.6 ± 0.05	-18.4 ± 0.1	-18.3 ± 0.05
ΔE_{solv} (ordered water)	-19.2 ± 0.1	-19.5 ± 0.1	0	0	$+19.2 \pm 0.1$	$+19.5 \pm 0.1$
ΔS_{solv} (ordered water)	-6.8 ± 0.1	-7.8 ± 0.1	0	0	$+6.8 \pm 0.1$	$+7.8 \pm 0.1$
ΔH_{desolv} [a]	+15.4	+15.7	+18.3	+19.3	+2.9	+3.6
ΔS_{desolv} [a]	+18.0	+18.9	+23.1	+26.4	+5.1	+7.5
ΔS_{config}	0	0	−2.9	−2.9	−2.9	−2.9
ΔE (total)[b]	-13.2 ± 0.1	-13.1 ± 0.1	-9.5 ± 0.02	-8.3 ± 0.05	$+3.7 \pm 0.1$	$+4.8 \pm 0.1$
ΔS (total)	$+11.2 \pm 0.1$	$+11.1 \pm 0.1$	+20.2	+23.5	$+9.0 \pm 0.1$	$+12.3 \pm 0.1$
ΔG (total)	-16.6 ± 0.1	-16.8 ± 0.1	-15.6 ± 0.02	-16.1 ± 0.05	$+1.0 \pm 0.1$	$+0.7 \pm 0.1$
ΔC_{p} [c]	3 ± 3		40 ± 1		$+37 \pm 3$	

*Note: Units for enthalpy and free energy are kcal mol^{-1}. Units for entropy and heat capacity are cal mol^{-1} K^{-1}.[a] ΔH_{desolv} stands for the contribution of enthalpies of dehydration from the hydroxyl group in trimannoside **1** and from the hydroxyethyl group in trimannoside **2**; Similarly for ΔS_{desolv}. [b] ΔE (total) = $E_{\text{L2OH-P}}$ + ΔE_{solv} (ordered water) + ΔH_{desolv}, and ΔG (total) = ΔE (total) $-T\Delta S$ (total). [c] ΔC_{p} was calculated with $\Delta\Delta E$ (total)/ΔT.

depending on the detailed balance of the various contributions involved. The effect of water-mediated interactions on binding specificity is much less clear. Some systems with bridging waters, such as Trp-repressor/operator, exhibit specificity and others, such as MHC, do not.

On the experimental side, the most widely used techniques so far have been mutagenesis and calorimetry. While powerful, these methods only provide thermodynamic data that result from a multitude of contributions. One would wish for more detailed information. One promising technique that should be further explored toward that goal is NMR.[48]

On the theoretical side, a number of methods have been applied to obtain the thermodynamics of interfacial water: FEP simulations,[96] the Gaussian model,[88] and IFST.[104–106] These methods all have advantages and disadvantages. Furthermore, their theoretical formulations are very different from each other. It would be useful in the future to explore the relationship between them both theoretically and by applying them all to a wider range of problems and comparing closely the results. IFST provides the greatest level of detail but is more cumbersome to use. The authors will be happy to share their programs with colleagues who might be interested in applying it to different systems. One issue that should be explored is the effect of thermal motion (small fluctuations in the protein–ligand configurations) on the values of the thermodynamic contributions obtained for water.

The above methods are too expensive in many practical situations. Much work still needs to be done to find practical ways to include water-mediated interactions in protein–ligand docking, scoring, and drug design. Most work so far appears to be somewhat *ad hoc*. One direction is to use discrete, explicit water molecules at key positions in the binding interface. This will require improvements in the methodology of predicting water locations and has the disadvantage of neglecting entropic effects. The second direction is towards implicit modeling of water. A start has been made with the knowledge-based water-mediated potentials of Wolynes and coworkers,[83,84] but it would be useful to develop physics-based potentials as well. What is essentially needed is the extension of implicit solvent models, such as ACE,[109] Generalized Born,[110] or EEF1,[111] to make them applicable to water in restricted environments. It is a difficult task, because the thermodynamics of bound water is sensitive to its surroundings. To capture this complexity correctly, one would most likely need many-body potentials and an extensive set of data on the dependence of the free energy of the bound water on the configuration of surrounding polar atoms.

Note added at proof

The double decoupling free energy method has recently been applied to four interfacial water molecules in the HIV-1 protease–KNI-272 complex.[112]

Acknowledgements

This work was supported by the National Science Foundation (MCB-0615552). Infrastructure support was provided in part by RCMI grant RR03060 from NIH.

References

1 Z. Otwinowski, R. W. Schevitz, R.-G. Zhang, C. L. Lawson, A. Joachimiak, R. Q. Marmorstein, B. F. Luisi and P. B. Sigler, *Nature*, 1988, **335**, 321–329.
2 F. A. Quiocho, D. K. Wilson and N. K. Vyas, *Nature*, 1989, **340**, 404–407.
3 V. Mikol, J. Kallen, G. Pflugl and D. Walkinshaw, *J. Mol. Biol.*, 1993, **234**, 1119–1130.
4 J. Zheng, E. A. Trafny, D. R. Knighton, N. Xuong, S. S. Taylor and J. M. Sowadski, *Acta Crystallogr., Sect. D*, 1993, **49**, 362–365.
5 T. N. Bhat and R. J. Poljak, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 1089–1093.
6 A. Joachimiak, T. E. Haran and P. B. Sigler, *EMBO J.*, 1994, **13**, 367–372.
7 J. R. H. Tame, G. N. Murshudov, E. J. Dodson, T. K. Neil, G. G. Dodson, C. F. Higgins and A. J. Wilkinson, *Science*, 1994, **264**, 1578–1581.
8 K. Huang, W. Y. Lu, S. Anderson, M. Laskowski and M. N. G. James, *Protein Sci.*, 1995, **4**, 1985–1987.
9 G. A. Holdgate, A. Tunnicliffe, W. H. J. Ward, S. A. Weston, G. Rosenbrack, P. T. Barth, I. W. F. Taylor, R. A. Pauptit and D. Timms, *Biochemistry*, 1997, **36**, 9663–9673.

10 S. H. Sleigh, J. R. H. Tame, E. J. Dodson and A. J. Wilkinson, *Biochemistry*, 1997, **36**, 9747–9758.
11 S. Shaltiel, S. Cox and S. S. Taylor, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 484–491.
12 K. Niefind, M. Putter, B. Guerra, O. G. Issinger and D. Schomburg, *Nat. Struct. Biol.*, 1999, **6**, 1100–1103.
13 M. Babor, V. Sobolev and M. Edelman, *J. Mol. Biol.*, 2002, **323**, 523–532.
14 C. K. Reddy, A. Das and B. Jayaram, *J. Mol. Biol.*, 2001, **314**, 619–632.
15 L. Lo Conte, C. Chothia and J. Janin, *J. Mol. Biol.*, 1999, **285**, 2177–2198.
16 F. Rodier, R. P. Bahadur, P. Chakrabarti and J. Janin, *Proteins: Struct., Funct., Bioinf.*, 2005, **60**, 36–45.
17 M. Levitt and B. H. Park, *Structure*, 1993, **1**, 223–226.
18 J. E. Ladbury, *Chem. Biol.*, 1996, **3**, 973–980.
19 J. Janin, *Struct. Fold. Des.*, 1999, **7**, R277–R279.
20 P. Cozzini, M. Fornabaio, A. Marabotti, D. J. Abraham, G. E. Kellogg and A. Mozzarelli, *Curr. Med. Chem.*, 2004, **11**, 3093–3118.
21 B. Jayaram and T. Jain, *Annu. Rev. Biophys. Biomol. Struct.*, 2004, **33**, 343–361.
22 E. T. Baldwin, T. N. Bhat, S. Gunik, B. S. Liu, I. A. Topol, Y. Kiso, T. Mimoto, H. Mitsuya and J. W. Erickson, *Structure*, 1995, **3**, 581–590.
23 Y.-X. Wang, D. I. Freedberg, P. T. Wingfield, S. J. Stahl, J. D. Kaufman, Y. Kiso, T. N. Bhat, J. W. Erickson and D. A. Torchia, *J. Am. Chem. Soc.*, 1996, **128**, 12287–12290.
24 J. Kervinen, N. Thanki, A. Zdanov, J. Tina, J. Barrish, P. F. Lin, R. Colonno, K. Riccardi, H. Samanta and A. Wlodawer, *Protein Pept. Lett.*, 1996, **3**, 399–406.
25 L. Hong, X. J. Zhang, S. Foundling, J. A. Hartsuck and J. Tang, *FEBS Lett.*, 1997, **420**, 11–16.
26 J. M. Louis, F. Dyda, N. T. Nashed, A. R. Kimmel and D. R. Davies, *Biochemistry*, 1998, **37**, 2105–2110.
27 P. Y. S. Lam, P. K. Jadhav, C. J. Eyermann, C. N. Hodge, Y. Ru, L. T. Bacheler, J. L. Meek, M. J. Otto, M. M. Rayner, Y. N. Wong, C. H. Chang, P. C. Weber, D. A. Jackson, T. R. Sharpe and S. EricksonViitanen, *Science*, 1994, **263**, 380–284.
28 S. Grzesiek, A. Bax, L. K. Nicholson, T. Yamazaki, P. T. Wingfield, S. J. Stahl, C. J. Eyermann, D. A. Torchia, C. N. Hodge, P. Y. S. Lam, P. K. Jadhav and C. H. Chang, *J. Am. Chem. Soc.*, 1994, **116**, 1581–1582.
29 C. N. Hodge, P. E. Aldrich, L. T. Bacheler, C. H. Chang, C. J. Eyermann, S. Garber, M. Grubb, D. A. Jackson, P. K. Jadhav, B. Korant, P. Y. S. Lam, M. B. Maurin, J. L. Meek, M. J. Otto, M. M. Rayner, C. Reid, T. R. Sharpe, L. Shum, D. L. Winslow and S. EricksonViitanen, *Chem. Biol.*, 1996, **3**, 301–314.
30 J. M. Chen, S. L. Xu, Z. Wawrzak, G. S. Basarab and B. D. Jordan, *Biochemistry*, 1998, **37**, 17735–17744.
31 P. C. Weber, M. W. Pantoliano, D. M. Simons and F. R. Salemme, *J. Am. Chem. Soc.*, 1994, **116**, 2717–2724.
32 S. W. Michnick, M. K. Rosen, T. J. Wandless, M. Karplus and S. L. Schreiber, *Science*, 1991, **252**, 836–839.
33 P. R. Connelly, R. A. Aldape, F. J. Bruzzese, S. P. Chambers, M. J. Fitzgibbon, M. A. Fleming, S. Itoh, D. J. Livingston, M. A. Navia, J. A. Thomson and K. P. Wilson, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 1964–1968.
34 S. Park and J. Saven, *Proteins: Struct., Funct., Bioinf.*, 2005, **60**, 450–463.
35 R. Loris, D. Maes, F. Poortmans, L. Wyns and J. Bouckaert, *J. Biol. Chem.*, 1996, **271**, 30614–30618.
36 C. Clarke, R. J. Woods, J. Gluska, A. Cooper, M. A. Nutley and G. J. Boons, *J. Am. Chem. Soc.*, 2001, **123**, 12238–12247.
37 V. Mikol, C. Papageorgiou and X. Borer, *J. Med. Chem.*, 1995, **38**, 3361–3367.
38 S. D. Sharrow, K. A. Edmonds, M. A. Goodman, M. V. Novotny and M. J. Stone, *Protein Sci.*, 2005, **14**, 249–256.
39 B. C. Braden, E. R. Goldman, R. A. Mariuzza and R. J. Poljak, *Immunol. Rev.*, 1998, **163**, 45–57.
40 A. Yokota, K. Tsumoto, M. Shiroishi, H. Kondo and I. Kumagai, *J. Biol. Chem.*, 2003, **278**, 5410–5418.
41 S. Chacko, E. Silverton, L. Kam-morgan, S. Smith-Gill, G. Cohen and D. Davies, *J. Mol. Biol.*, 1995, **245**, 261–274.
42 K. A. Watson and A. Papageorgiou, *Biochemistry*, 1994, **19**, 5745–5758.
43 W. Dall'Acqua, E. R. Goldman, W. H. Lin, C. Teng, D. Tsuchiya, H. M. Li, X. Ysern, B. C. Braden, Y. L. Li, S. Smith-Gill and R. A. Mariuzza, *Biochemistry*, 1998, **37**, 7981–7991.
44 J. Xu, W. A. Basse, M. L. Quillin, E. P. Baldwin and B. W. Matthews, *Protein Sci.*, 2001, **10**, 1067–1078.
45 U. Langhorst, J. Backmann, R. Loris and J. Steyaert, *Biochemistry*, 2000, **39**, 6586–6593.
46 K. Takano, Y. Yamagata, M. Kubota, J. Funahashi, S. Fujii and K. Yutani, *Biochemistry*, 1999, **38**, 6623–6629.
47 K. Takano, Y. Yamagata, J. Funahashi, Y. Hioki, S. Kuramitsu and K. Yutani, *Biochemistry*, 1999, **38**, 12968–12708.
48 V. P. Denisov, K. Venu, J. Peters, H. D. Horlein and B. Halle, *J. Phys. Chem. B*, 1997, **101**, 9380–9389.
49 C. R. Robinson and S. G. Sligar, *J. Mol. Biol.*, 1993, **234**, 302–306.
50 M. P. Brown, a. O. Grillo, M. Boyer and C. A. Royer, *Protein Sci.*, 1999, **8**, 1276–1285.
51 M. G. Fried, D. F. Stickle, K. V. Smirnakis, C. Adams, D. MacDonald and P. Z. Lu, *J. Biol. Chem.*, 2002, **277**, 50676–50682.
52 S. Bergqvist, M. A. Williams, R. O'Brien and J. E. Ladbury, *J. Mol. Biol.*, 2004, **336**, 829–842.
53 J. E. Ladbury, J. G. Wright, J. M. Sturtevant and P. B. Sigler, *J. Mol. Biol.*, 1994, **238**, 669–681.
54 C. J. Morton and J. E. Ladbury, *Protein Sci.*, 1996, **5**, 2115–2118.
55 A. M. Davis, S. J. Teague and G. J. Kleywegt, *Angew. Chem., Int. Ed.*, 2003, **42**, 2718–2736.
56 R. W. Pitt and J. M. Goodfellow, *Protein Eng.*, 1991, **4**, 531–537.
57 P. J. Goodford, *J. Med. Chem.*, 1985, **28**, 849–857.
58 R. C. Wade and P. J. Goodford, *J. Med. Chem.*, 1993, **36**, 148–156.
59 M. Pastor, G. Cruciani and K. A. Watson, *J. Med. Chem.*, 1997, **40**, 4089–4102.
60 M. Fornabaio, F. Spyrakis, A. Mozzarelli, P. Cozzini, D. J. Abraham and G. E. Kellogg, *J. Med. Chem.*, 2004, **47**, 4507–4516.
61 A. Miranker and M. Karplus, *Proteins*, 1991, **11**, 29–34.
62 M. L. Verdonk, J. C. Cole and R. Taylor, *J. Mol. Biol.*, 1999, **289**, 1093–1108.
63 T. Kortvelyesi, S. Dennis, M. Silberstein, l. I. Brown and S. Vajda, *Proteins*, 2003, **51**, 340–351.
64 J. W. H. Schymkowitz, F. Rousseau, I. C. Martins, J. Ferkinghoff-Borg, F. Stricher and L. Serrano, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 10147–10152.
65 A. T. Garcia-Sosa, R. L. Mancera and P. M. Dean, *J. Mol. Model.*, 2003, **9**, 172–182.
66 M. Rarey, B. Kramer and T. Lengauer, *Proteins: Struct., Funct., Bioinf.*, 1999, **34**, 17–28.
67 F. Osterberg, G. M. Morris, M. F. Sanner, A. J. Olson and D. S. Goodsell, *Proteins: Struct., Funct., Bioinf.*, 2002, **46**, 34–40.
68 J. Wang, S. L. Chan and K. Ramnarayan, *J. Comput. Aided Mol. Des.*, 2003, **18**, 495–513.
69 R. X. Wang, Y. P. Lu and S. M. Wang, *J. Med. Chem.*, 2003, **46**, 2287–2303.
70 D. G. Lloyd, A. T. Garcia-Sosa, I. L. Alberts, N. P. Todorov and R. L. Mancera, *J. Comput. Aided Mol. Des.*, 2004, **18**, 89–100.
71 C. De Graaf, P. Pospisil, W. Pos, G. Folkers and N. P. E. Vermeulen, *J. Med. Chem.*, 2005, **48**, 2308–2318.
72 M. L. Verdonk, G. Chessari, J. C. Cole, M. J. Hartschorn, C. W. Murray, J. W. M. Nissink, R. D. Taylor and R. Taylor, *J. Med. Chem.*, 2005, **48**, 6504–6515.
73 A. Amadasi, F. Spyrakis, P. Cozzini, D. J. Abraham, G. E. Kellogg and A. Mozzarelli, *J. Mol. Biol.*, 2006, **358**, 289–309.
74 P. Cozzini, M. Fornabaio, A. Mozzarelli, F. Spyrakis, G. E. Kellogg and D. J. Abraham, *Int. J. Quantum Chem.*, 2006, **106**, 647–651.
75 R. L. Mancera, *J. Comput. Aided Mol. Des.*, 2002, **16**, 479–499.
76 V. Schnecke and L. A. Kuhn, *Perspect. Drug Discovery Des.*, 2000, **20**, 171–190.
77 R. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739–1749.

78 N. Moitessier, E. Westhof and S. Hanessian, *J. Med. Chem.*, 2006, **49**, 1023–1033.
79 P. Pospisil, T. Kuoni, L. Scapozza and G. Folkers, *J. Recept. Signal Transduction Res.*, 2002, **22**, 141–154.
80 J. M. Yang and C. C. Chen, *Protein Sci.*, 2004, **55**, 288–304.
81 L. Birch, C. W. Murray, M. J. Hartschorn, I. J. Tickle and M. L. Verdonk, *J. Comput. Aided Mol. Des.*, 2002, **16**, 855–869.
82 L. Jiang, B. Kuhlman, T. Kortemme and D. Baker, *Proteins: Struct., Funct., Bioinf.*, 2005, **58**, 893–904.
83 G. A. Papoian, J. Ulander and P. G. Wolynes, *J. Am. Chem. Soc.*, 2003, **125**, 9170–9178.
84 G. A. Papoian, J. Ulander, M. P. Eastwook, Z. Luthey-Schulten and P. G. Wolynes, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 3352–3357.
85 C. Zong, G. A. Papoian, J. Ulander and P. G. Wolynes, *J. Am. Chem. Soc.*, 2006, **128**, 5168–5176.
86 J. D. Dunitz, *Science*, 1994, **264**, 670.
87 D. G. Covell and A. Wallquist, *J. Mol. Biol.*, 1997, **269**, 281–297.
88 P. M. Petrone and A. E. Garcia, *J. Mol. Biol.*, 2004, **338**, 419–435.
89 S. Fischer, J. C. Smith and C. S. Verma, *J. Phys. Chem. B*, 2001, **105**, 8050–8055.
90 C. S. Verma and S. Fischer, *Biophysical Chemistry*, 2005, **115**, 295–302.
91 M. Brandl, M. Meyer and J. Suhnel, *J. Phys. Chem. A*, 2000, **104**, 11177–11187.
92 S. M. Tschampel and R. J. Woods, *J. Phys. Chem. A*, 2003, **107**, 9175–9181.
93 R. C. Wade, M. H. Mazor, J. A. McCammon and F. A. Quiocho, *Biopolymers*, 1991, **31**, 919–931.
94 B. Roux, M. Nina, R. Pomes and J. C. Smith, *Biophys. J.*, 1996, **71**, 670–681.
95 L. Zhang and J. Hermans, *Proteins: Struct., Funct., Bioinf.*, 1996, **24**, 433–438.
96 D. Hamelberg and J. A. McCammon, *J. Am. Chem. Soc.*, 2004, **126**, 7683–7689.
97 R. L. Olano and S. W. Rick, *J. Am. Chem. Soc.*, 2004, **126**, 7991–8000.
98 M. Tashiro and A. A. Stuchebrukhov, *J. Phys. Chem. B*, 2005, **109**, 1015–1022.
99 M. K. Gilson, J. A. Given, B. L. Bush and J. A. McCammon, *Biophys. J.*, 1997, **72**, 1047–1069.
100 G. Archontis, K. A. watson, Q. Xie, G. Andreou, E. D. Chrysina, S. E. Zographos, N. G. Oikonomakos and M. Karplus, *Proteins: Struct., Funct., Bioinf.*, 2005, **61**, 984–998.
101 P. Monecke, T. Borosch, J. Brickmann and S. M. Kast, *Biophys. J.*, 2006, **90**, 841–850.
102 T. Lazaridis, *J. Phys. Chem. B*, 1998, **102**, 3531–3541.
103 T. Lazaridis, *J. Phys. Chem. B*, 1998, **102**, 3542–3550.
104 Z. Li and T. Lazaridis, *J. Am. Chem. Soc.*, 2003, **125**, 6636–6637.
105 Z. Li and T. Lazaridis, *J. Phys. Chem. B*, 2005, **109**, 662–670.
106 Z. Li and T. Lazaridis, *J. Phys. Chem. B*, 2006, **110**, 1464–1475.
107 T. Lazaridis and M. Karplus, *Biophys. Chem.*, 2003, **100**, 367–395.
108 T. Lazaridis and M. Karplus, *Curr. Opin. Struct. Biol.*, 2000, **10**, 139–145.
109 M. Schaefer and M. Karplus, *J. Phys. Chem.*, 1996, **100**, 1578–1599.
110 D. Qiu, P. S. Shenkin, F. P. Hollinger and W. C. Still, *J. Phys. Chem. A*, 1997, **101**, 3005–3014.
111 T. Lazaridis and M. Karplus, *Proteins: Struct., Funct., Bioinf.*, 1999, **35**, 133–152.
112 Y. Lu, C. Y. Yang and S. Wang, *J. Am. Chem. Soc.*, 2006, **128**, 11830–11839.