THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Enthalpic Breakdown of Water Structure on Protein Active-Site Surfaces

Kamran Haider,[†] Lauren Wickstrom,[‡] Steven Ramsey,[†,∥] Michael K. Gilson,[§] and Tom Kurtzman*[,†,∥,⊥]

[†]Department of Chemistry, Lehman College, The City University of New York, 250 Bedford Park Boulevard West, Bronx, New York 10468, United States

[‡]Borough of Manhattan Community College, Department of Science, The City University of New York, 199 Chambers Street, New York, New York 10007, United States

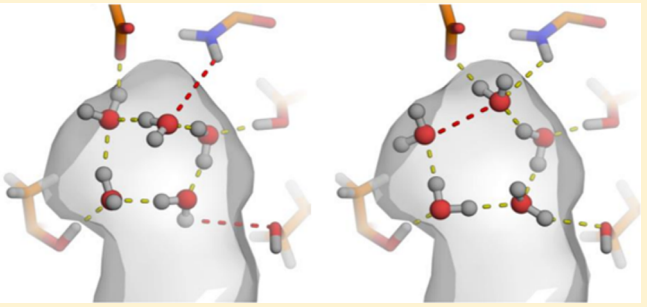
[§]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093-0736, United States

[∥]Ph.D. Program in Biochemistry, The Graduate Center of The City University of New York, 365 Fifth Avenue, New York, New York 10016, United States

[⊥]Ph.D. Program in Chemistry, The Graduate Center of The City University of New York, 365 Fifth Avenue, New York, New York 10016, United States

Supporting Information

ABSTRACT: The principles underlying water reorganization around simple nonpolar solutes are well understood and provide the framework for the classical hydrophobic effect, whereby water molecules structure themselves around solutes so that they maintain favorable energetic contacts with both the solute and the other water molecules. However, for certain solute surface topographies, water molecules, due to their geometry and size, are unable to simultaneously maintain favorable energetic contacts with both the surface and neighboring water molecules. In this study, we analyze the solvation of ligand-binding sites for six structurally diverse proteins using hydration site analysis and measures of local water structure, in order to identify surfaces at which water molecules are unable to structure themselves in a way that maintains favorable enthalpy relative to bulk water. These surfaces are characterized by a high degree of enclosure, weak solute−water interactions, and surface constraints that induce unfavorable pair interactions between neighboring water molecules. Additionally, we find that the solvation of charged side chains in an active site generally results in favorable enthalpy but can also lead to pair interactions between neighboring water molecules that are significantly unfavorable relative to bulk water. We find that frustrated local structure can occur not only in apolar and weakly polar pockets, where overall enthalpy tends to be unfavorable, but also in charged pockets, where overall water enthalpy tends to be favorable. The characterization of local water structure in these terms may prove useful for evaluating the displacement of water from diverse protein active-site environments.



1. INTRODUCTION

The concept of protein−ligand complementarity effectively states that, in order for a ligand to bind with reasonable affinity to a given target protein, it must fit in the cavity to which it binds without steric conflicts (shape complementarity) and make appropriate hydrophilic and hydrophobic contacts with the protein surface (electrostatic complementarity). While the idea of complementarity is a simplified picture of the molecular recognition process, it is exceptionally valuable in that it provides the fundamental conceptual framework on which structure-based drug discovery and design is based. On the other hand, water, while widely recognized as playing an instrumental role in the molecular recognition process, and having a significant, if not dominant, contribution to the free energy of recognition, is not described in the complementarity framework, except indirectly in that the chemical groups of both the ligand and the protein are frequently categorized by their interaction with water (i.e., either hydrophobic or hydrophilic). Unfortunately, for water, there is no simple conceptual framework within which drug designers can estimate whether displacement of water molecules from a region will be free energetically beneficial, detrimental, or neither.

Part of the reason for the lack of such a framework is that water is a complex fluid and presents a number of conundrums, which do not exist when considering only direct protein−ligand

Received: February 1, 2016
Revised: April 12, 2016

ACS Publications © XXXX American Chemical Society

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

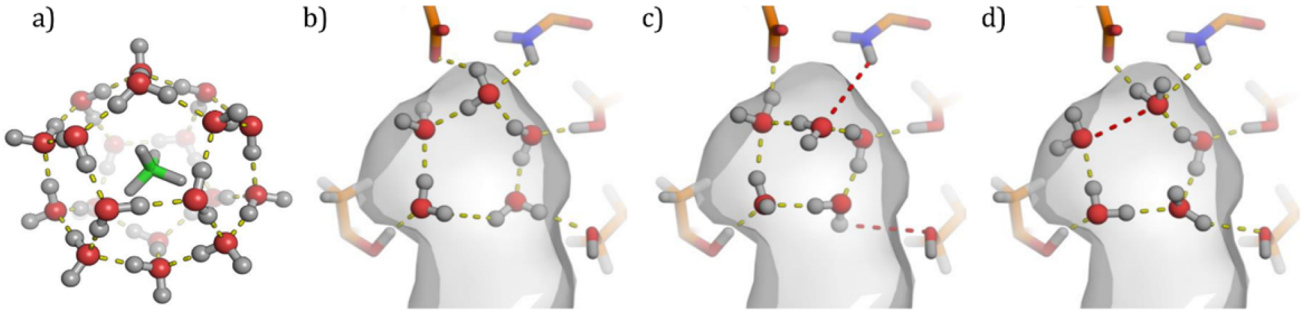


Figure 1. Water structure on simple and complex surfaces. (a) Water clathrate cage around a methane molecule.[58] Optimal (b) and suboptimal (c and d) water structure on a protein active site surface, illustrated using MD simulation snapshots of a part of the streptavidin active site. The red dashes in parts c and d indicate lost interactions that are maintained in the optimal scenario.

interactions. On a given surface, water molecules are in constant fluctuation, translating and reorienting in space, leading to density fluctuations and the regular making and breaking of hydrogen bonds. Although it is generally a reasonable approximation to consider a ligand to be fixed in a pose, thereby allowing the assessment of whether a given protein−ligand hydrophilic and hydrophobic contact is being made or not being made, a water molecule's interactions with a surface are fluctuating too much to support a view that water is *posed* on a surface. Such considerations have made it difficult to construct a simple, yet generally accurate, conceptual framework for considering the role of solvation in molecular recognition with protein surfaces.

Despite these difficulties, the characterization of water structure and thermodynamics on protein surfaces has attracted strong interest due to its impact on protein−ligand binding.[1−9] As a result, a significant number of computational tools have been developed in the past years that calculate the thermodynamics of water on solute surfaces.[10−21] The application of these tools on protein surfaces has led to a better understanding of the role of water molecules in molecular recognition. In other studies, water structure on protein surfaces has been characterized in terms of the geometry and local environment of solvating water molecules. For example, in a series of explicit-water simulations of the melittin tetramer, Rossky and co-workers investigated the effect of surface topography on the structure of solvating water, which indicated the clathrate-like water structure, typical of small convex surfaces, exchanged with a less ordered inverted configuration.[22] Young et al. identified hydrophobic enclosures and correlated hydrogen bonds on protein surfaces as key factors underlying unusual enthalpic and entropic costs of solvation.[11] Sharp used the angular distribution of water−water hydrogen bonds to show that an increased fraction of water molecules engaged in low-angle hydrogen bonds on protein surfaces that contained both polar and apolar residues.[23,24] Several other investigations of the properties of water molecules on protein surfaces are summarized in recent reviews.[1,24−29]

The iceberg model[30] of hydrophobic hydration, which was first laid out by Frank and Evans and later elaborated upon in biological settings by Kauzmann, Tanford, and others,[1,32,28] provides a conceptual framework for understanding water reorganization around solutes. In this model, water structure is dominated by packing around the solute so that it maintains its hydrogen bond interactions with the surrounding environment. In the hydration of small hydrophobic moieties, water molecules pack around the hydrophobic moiety, forming an iceberg or clathrate-like cage, while maintaining hydrogen bond contacts with neighboring water molecules at a level similar to that of water molecules in the bulk (Figure 1a). This is reflected by the small changes in hydration enthalpies for small hydrophobic molecules, such as argon or methane, but a corresponding entropic penalty due to the ordering necessary to form this cage.[33] This strengthening of the water network and the corresponding entropic penalty have remained key features of our understanding of the hydrophobic effect at biological temperatures.

In contrast to simple hydrophobic surfaces,[34] protein active sites vary electrostatically and are composed of amino acids, which can have apolar, neutral, and charged side chains, and can both donate and accept hydrogen bonds. In addition, binding site surfaces are frequently concave and the solvating water molecules are often surrounded by the protein in multiple directions. Despite these differences, we expect the physical principles that guide the water structure in such complex environments to be similar, in that water molecules tend to pack into active sites so that they maintain favorable hydrogen bond contacts with both their water neighbors and the protein surface. An example of optimal structuring of water molecules in a protein active site is illustrated in Figure 1b, which shows a configuration taken from an explicit solvent molecule dynamics (MD) trajectory of streptavidin. In this example, the water molecules form favorable hydrogen bond contacts both with their neighboring water molecules and with the protein surface. However, it is not difficult to imagine surface topographies and charge distributions that cannot be efficiently solvated by water. A relatively trivial example of this is a single water molecule surrounded by three or more hydrogen bond acceptors. In this case, not all of the hydrogen bonds can be satisfied, since a water molecule can only simultaneously donate two hydrogen bonds, one for each hydrogen. Scenarios in which there is a breakdown of such an optimal structure are illustrated in Figure 1c and d, where water molecules do not make all of the favorable interactions with the surface (c) or with their neighbors (d). The relative populations of optimal and suboptimal water configurations are relevant in the context of molecular recognition. For instance, if the predominant water configurations corresponded to Figure 1c and d, as opposed to the configuration in Figure 1b, it would be reasonable to expect that the free energetic cost of displacing water would be lower, and hence lead to higher ligand-binding affinity. Consequently, the determination of whether surface water maintains favorable energetic interactions with its neighbors, with the protein surface, and with its environment as a whole is of crucial value in describing the driving forces of ligand binding and potentially in the optimization of lead compounds in drug design.

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

Table 1. Proteins Investigated in This Study[a]

protein	abbrev	PDB ID	res (Å)	number of sites					
				A.En	A.Fr	P.En	P.Fr	C.En	C.Fr
Abl kinase	Abl	3CS9[35]	2.21	13	1	13	2	2	8
acetylcholinesterase	AChE	4EY7[36]	2.35	14	2	18	4	4	2
caspase 3	Casp	2H5I[37]	1.69	5	0	16	11	0	8
factor Xa	Fxa	1FJS[38]	1.92	8	3	10	6	3	12
HIV-1 protease	HIVP	3NUO[39]	1.35	4	1	8	3	9	8
streptavidin	Strept	1STP[40]	2.60	5	0	10	2	0	3
total				49	7	75	28	18	41

[a]In each case, chain A in the PDB entry was analyzed. The numbers of hydration sites of each type (see main text) are also listed.

The present paper investigates the reorganization of water molecules in a diverse set of six protein active sites. We focus on the energetics of water reorganization, as we aim to identify protein surface features that prevent water from maintaining favorable interactions with neighboring water molecules and with their environments as a whole. Our characterization of water is based on simple measures of local structure that describe water−water interactions in the first solvation shell and are obtained from explicit solvent molecular dynamics simulations of solutes. We compare the structure of water in protein active sites with that of bulk water, and as well as that of water at the surfaces of simple apolar, polar, and charged solutes. This comparison of water molecules on different solute surfaces results in identification of protein surface features that cause strong perturbation in the local structure of water molecules, relative to the bulk. The measures of local structure reported in this study also provide a useful approach for quantifying and visualizing these perturbations and evaluating the displacement of water molecules from protein active sites.

2. METHODS

Explicit water molecular dynamics (MD) simulations were run followed by hydration site analysis (HSA)[2] of trajectories. We first investigated the properties of water molecules in neat water, and on the surfaces of methane, the exterior of cucurbit[7]uril (CB7), and the side chains of phenylalanine (Phe), tyrosine (Tyr), asparagine (Asn), glutamine (Gln), aspartate (Asp), and arginine (Arg). These were followed by analogous analysis of water in the binding sites of six proteins listed in Table 1. Details of the simulation protocols and calculation of water properties are described below.

2.1. Systems and Simulation Setup. *2.1.1. System Preparation.* All of the simulations were run using the TIP3P water model. For the bulk water simulations, a small box containing 521 water molecules was used. The CB7 molecule and amino acids for reference simulations were typed with the OPLS2005[41] force field and amber99SB,[42] respectively. For each protein system, a ligand-bound X-ray structure (Table 1) was used as the starting point after removing the ligand molecule. All protein systems were prepared using Maestro Protein Preparation Wizard.[43] This step involved optimization of side chain orientations, assignment of protonation states, and capping of the N- and C-termini of the peptide chains with N-acetyl and N-methylamide groups. Asp27 in HIV protease and Glu202 in acetylcholinesterase were assigned to be neutral. For the rest of the systems, all aspartate and glutamate side chains were assigned as negatively charged and lysine and arginine side chains were assigned as positively charged, while histidines were assigned as neutral. These assignments were based on the default preparation protocol of Maestro Protein Preparation Wizard and were further confirmed by checking the papers corresponding to each of the above PDB structures for evidence of any alternative protonation states.

2.1.2. Simulation Protocol. Each of the following simulations was carried out using Desmond (GPU version).[44] All systems containing solute atoms were solvated in a periodic box of TIP3P water molecules whose edges were located at least 10 Å from any of the solute atoms. Appropriate numbers of Na^+ or Cl^- ions were placed to neutralize each system. After preparation, each system was subjected to energy minimization which involved 1500 steps of steepest descent, with all solute atoms harmonically restrained using a force constant of 100 kcal/mol Å^2. This was followed by another round of up to 2000 steps of conjugate gradient minimization, using the same restraints but only for solute heavy atoms. A multistage NPT equilibration was performed at a constant pressure of 1 atm, where restraints were gradually decreased from 100 to 10 kcal/mol Å^2, in decrements of 10 kcal/mol Å^2. The final production was run under constant temperature, pressure, and number of particles. The total lengths of equilibration and production runs were 2 and 10 ns, respectively. In all production simulations, solute atoms were harmonically restrained using force constants of 10.0 kcal/mol/Å^2. For all systems, simulations were carried out with the SHAKE algorithm[45] to constrain bond lengths involving hydrogen atoms. Temperature was regulated with the Langevin thermostat with a relaxation time of 1.0 ps^{-1}, and pressure was regulated by isotropic position scaling, with a relaxation time of 2.0 ps^{-1}. During production simulations, snapshots were saved every 1 ps, resulting in 10,000 snapshots in total for each system.

2.2. Hydration Site Analysis. *2.2.1. Identification of Hydration Sites.* High density spherical regions (hydration sites) of 1 Å radius were identified using a clustering procedure reported previously.[2] In brief, for each protein, the spatial coordinates of the oxygen and two hydrogens of every water molecule were obtained for 1000 snapshots, using every 10th of the original 10,000 snapshots, and were stored in a single superconfiguration. From this superconfiguration, for each water molecule, its neighboring water molecules within 1.0 Å oxygen−oxygen distance were counted. The coordinates of the water molecule with the highest number of neighbors were selected as the position of the center of a potential hydration site. This water molecule and all of its neighbors were removed from subsequent iterations of the same procedure which was repeated until all hydration sites were identified with a number of water neighbors at least twice that of bulk water. The resulting hydration sites were each populated by retrieving all water molecules from the full 10,000 snapshots, which had

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

oxygen atoms within 1.0 Å from the corresponding hydration site center. As in the initial clustering step, only hydration sites with twice bulk or higher density (corresponding to a population of 2800 or higher) were retained. The populated hydration sites were then enumerated according to their occupancies, with the highest populated site being given the index 0. In the subsequent text, a hydration site belonging to a protein is referred to by using the abbreviation of the protein (Table 1) followed by its occupancy rank; e.g., Abl 0 is the most occupied site in Abl kinase.

In order to depict a highly occupied water configuration associated with a given hydration site, we grouped all MD snapshots with an instance of a water molecule in the site; computed the distances between all pairs of these instances, in a six-dimensional rotation−translation space; identified the pair with the shortest distance; and stored the configuration of one member of this pair. The hydrogen bonds in the illustrations are computed on the basis of this configuration, using the geometric criteria above. As a consequence, the number of hydrogen bonds displayed does not always match the numbers listed in the data tables. In the study of ligand displacement of hydration sites, a hydration site is considered to be displaced if any heavy atom of the ligand is within 2.0 Å of the center of the site.

2.3. Descriptors of Water Structure and Energetics.

The subsections below define a set of properties used to characterize hydration sites in terms of both structure and energetics. Note that all of these quantities are normalized by the occupancy of the hydration site. For example, when computing the average number of first-shell neighbors for the water in a hydration site which is occupied by a water molecule in only 0.7 of the N_{frames} MD frames, we use a denominator of $0.7N_{\text{frames}}$.

2.3.1. Water Neighbors, Enclosure, and Hydrogen Bonding. Two water molecules are considered to be first-shell neighbors in a given MD frame if their oxygen atoms are separated by less than 3.5 Å. This distance criterion is based on the location of the minima of the oxygen−oxygen radial distribution function for bulk water.[46] The mean number of these first shell neighbors for water in each hydration site is termed N^{nbr}, and this first-shell neighbor count is used to define the fractional enclosure of a hydration site, by referencing the number of its first-shell neighbors to the mean number of first-shell neighbors of a TIP3P water molecule in the bulk, $N^{\text{nbr}}_{\text{bulk}}$:

$$f_{\text{enc}} \equiv 1 - \frac{N^{\text{nbr}}}{N^{\text{nbr}}_{\text{bulk}}}$$

This quantity indicates the degree to which the water in a hydration site is blocked from contact with other water molecules.

Two molecules are considered to be hydrogen bonded if their donor−acceptor heavy atom distance is within 3.5 Å and the hydrogen−donor−acceptor angle is less than or equal to 30°.[46] The mean numbers of hydrogen bonds made by the water in a hydration site with other water molecules and with the solute are written as $N^{\text{HB}}_{\text{ww}}$ and $N^{\text{HB}}_{\text{sw}}$, respectively, while the mean number of hydrogen bonds made by a water in the bulk is $N^{\text{HB}}_{\text{ww,bulk}}$. One may also define the number of water−water hydrogen bonds whose loss is attributable to enclosure (i.e., assuming no restructuring of water relative to bulk) as

$$N^{\text{HB}}_{\text{ww,lost}} \equiv f_{\text{enc}} N^{\text{HB}}_{\text{ww,bulk}}$$

The quantity $N^{\text{HB}}_{\text{sw}} - N^{\text{HB}}_{\text{ww,lost}}$ represents a measure of how well hydrogen bonds from the hydration site to the solute compensate, or even overcompensate, for those lost with other waters in the bulk. Finally, we define an additional measure of local structure, the average number of hydrogen bonds the water in a hydration site makes with its first-shell water neighbors, per first-shell water neighbor:

$$f^{\text{HB}}_{\text{ww}} = \frac{N^{\text{HB}}_{\text{ww}}}{N^{\text{nbr}}}$$

2.3.2. Energy Analysis. The total energy of the water in a hydration site, E_{tot}, is calculated as the sum of its mean water−water E_{ww} and solute−water E_{sw} interaction energies, where E_{ww} is one-half the mean interaction energy of the water in the site with all other waters and E_{sw} is one-half the mean interaction energy of the water in the site with the solute. The factors of one-half follow the convention that half of the interaction energy in a pairwise interaction is assigned to each molecule of the pair. We also compute $E^{\text{nbr}}_{\text{ww}}$, the mean interaction energy of the water in each hydration site with its first-shell neighbors, where the first shell is defined with the same distance cutoff as for N^{nbr}, and the same factor of one-half mentioned above is applied. One may characterize the structuring of the water associated with a hydration site in terms of the quantity

$$\bar{E}^{\text{nbr}}_{\text{ww}} \equiv \frac{E^{\text{nbr}}_{\text{ww}}}{N^{\text{nbr}}}$$

This is the mean interaction energy of water in the site with its first shell, normalized by the mean number of water molecules in the first shell; we consider water structure to be enhanced or frustrated around a hydration site depending on whether $\bar{E}^{\text{nbr}}_{\text{ww}}$ is more or less favorable, respectively, than the corresponding quantity for a water in bulk.

We also characterized the first-shell energetics of selected hydration sites i in more detail by computing the probability distributions of their site-to-first shell interactions. For each MD frame k with a water in hydration site i, we computed the interaction energies, $E_{\text{ww},ijk}$, of the site water with each of its j = 1, ..., $N_{i,k}$ first-shell waters in the frame. For example, if frame k included a water at hydration site i and four first-shell neighbors, then $N_{i,k}$ = 4 and four interaction energies were recorded for the site in this frame. The resulting values of $E_{\text{ww},ijk}$ across all MD frames were binned into energy intervals of δE = 0.06 kcal/mol, where bins were indexed with subscript n, so that the energy range associated with bin n is $[n\delta E, (n+1)\delta E]$, and the number of energy values in bin n is termed N_n. Since the bins are narrow, we associated energy $E_n = n\delta E$ with bin n. The number density, over energy, of the first-shell interaction energies for site i then was computed as

$$\rho_i(E_n) = \frac{N_n}{\delta E \sum_k N_{i,k}} N^{\text{nbr}}$$

Note that $\sum_n N_n = \sum_k N_{i,k}$. The quantity $\rho_i(E_n)$ is a number density, so that summing it over energy bins n gives the average number of first shell neighbors for hydration site i. Additionally, summing $E_n\rho_i(E_n)$ over bins n gives the average total water−water energy in the first shell of site i. It is also worth noting that the precise value of the hydration site radius has a minimal effect on the energetic quantities described above. We confirmed this by computing E_{tot} and $\bar{E}^{\text{nbr}}_{\text{ww}}$ for all hydration sites in factor Xa, using distance cutoffs of 0.8 and 1.2 Å, and

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

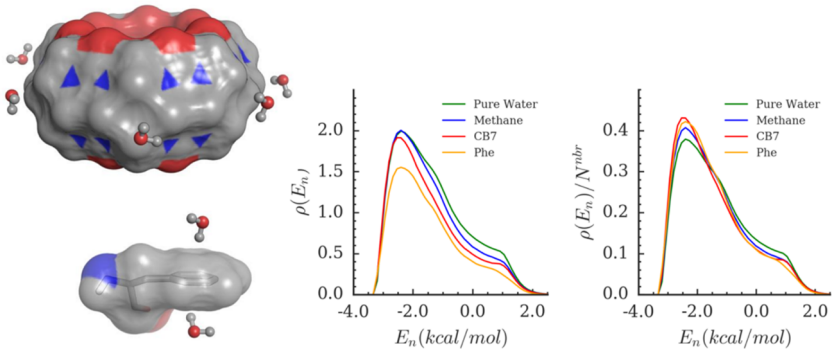


Figure 2. Hydration of simple apolar surfaces. The most probable water configurations in investigated hydration sites on the exterior surface of CB7 (top left) and the aromatic surface of the Phe side chain (bottom left). The molecular surface of CB7 and Phe is shown: gray, carbon; red, oxygen; blue, nitrogen. Number density (middle) and probability distribution (right) of the interaction energy for water molecules on methane, CB7, and Phe surfaces, compared with the same distribution for pure water.

Table 2. Average Energetic (kcal/mol) and Structural Quantities of Water Molecules in Neat Water, at the Surfaces of Simple Solutes and Protein Active Sites[a]

	number	E_{sw}	E_{ww}	E_{tot}	$\bar{E}^{nbr}_{ww}$	N^{nbr}	f_{enc}	N^{HB}_{sw}	N^{HB}_{ww}	f^{HB}_{ww}	$N^{HB}_{ww,lost}$
pure water		0.00	−9.53	−9.53	−1.36	5.26	0.00	0.00	3.33	0.63	0.00
Simple Apolar											
methane		−0.05	−9.42	−9.47	−1.45	4.92	0.12	0.00	3.28	0.67	0.21
CB7		−0.77	−8.79	−9.57	−1.51	4.44	0.18	0.00	3.13	0.70	0.52
Phe		−1.65	−7.46	−9.11	−1.52	3.68	0.31	0.00	2.59	0.69	1.00
Protein Apolar Averages											
A.En.F	32	−2.08	−8.04	−10.11	−1.65	4.24	0.21	0.01	3.20	0.76	0.64
A.En.U	17	−2.31	−6.82	−9.14	−1.67	3.67	0.31	0.01	2.71	0.75	1.00
A.Fr.F	2	−3.85	−6.19	−10.04	−1.33	4.79	0.13	0.05	3.20	0.67	0.30
A.Fr.U	5	−2.59	−6.33	−8.92	−1.28	4.04	0.26	0.02	2.54	0.63	0.77
Simple Polar											
Asn		−2.79	−7.03	−9.81	−1.33	4.02	0.25	0.88	2.51	0.62	0.78
Protein Polar Averages											
P.En.F	45	−3.81	−6.30	−10.11	−1.60	3.23	0.39	0.99	2.31	0.73	1.28
P.En.U	30	−3.08	−5.90	−8.98	−1.69	2.85	0.46	0.72	2.05	0.74	1.52
P.Fr.F	17	−5.59	−4.89	−10.48	−1.16	3.71	0.30	1.18	2.25	0.60	0.98
P.Fr.U	11	−4.51	−4.45	−8.96	−1.18	3.04	0.43	1.11	1.79	0.59	1.40
Simple Charged											
Asp		−6.42	−4.39	−10.82	−1.08	4.46	0.19	1.03	2.47	0.55	0.50
Arg site 1		−5.42	−5.00	−10.42	−1.22	3.85	0.27	0.97	2.39	0.62	0.89
Arg site 2		−6.64	−4.17	−10.82	−0.96	4.10	0.24	1.59	2.15	0.52	0.73
Protein Charged Averages											
C.En.F	18	−5.64	−5.10	−10.75	−1.70	2.77	0.47	1.26	2.08	0.76	1.57
C.Fr.F	41	−7.60	−3.13	−10.73	−0.95	3.34	0.37	1.36	1.88	0.55	1.21
displaced sites	104	−4.19	−5.73	−9.93	−1.51	3.40	0.36	0.68	2.38	0.71	1.19
undisplaced sites	114	−4.49	−5.56	−10.05	−1.39	3.41	0.36	0.93	2.26	0.66	1.20

[a]The amino acid main chain amines and carboxylic acids were capped with N-acetyl and N-methylamide groups, respectively, and water sites were examined only at the side chains, as detailed in the text. Similarly, we examined water only at the apolar outer surface of CB7. The data for Phe are averages of the very similar results for water sites above and below the ring. The data for Asn and Gln are also very similar, and only Asn results are reported here (see the Supporting Information for Gln). See text for definitions of Arg 1 and Arg 2. Averages are reported for protein hydration site groups. See the Supporting Information for standard deviations of the averages listed here.

comparing the results with those obtained using a cutoff of 1.0 Å. The largest differences in these quantities across all sites is found to be 0.12 and 0.04 kcal/mol, for E_{tot} and $\bar{E}^{nbr}_{ww}$, respectively, and the means and standard deviations of these differences are considerably smaller. For example, E_{tot} changes by an average of 0.052 ± 0.03 kcal/mol and $\bar{E}^{nbr}_{ww}$ by 0.009 ± 0.01 kcal/mol when calculated with hydration site radii of 1.0 Å versus 1.2 Å.

3. RESULTS AND DISCUSSION

We consider water at three types of solute surfaces: apolar (A), neutral polar (P), and charged (C). For each type, we examine a simple solute of appropriate type, such as methane for apolar and then contrast this with water sites at the same type of surface in a protein binding pocket. Neat water serves as an additional reference baseline. The water sites at each type of surface are additionally classified by the quantity $\bar{E}^{nbr}_{ww}$ (see Methods), so that a site is considered to have "enhanced" water

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

Table 3. Energetic (kcal/mol) and Structural Properties of Water Molecules in Specific Hydration Sites Belonging to Categories Listed in Table 2

	type	E_{sw}	E_{ww}	E_{tot}	$\overline{E}_{ww}^{nbr}$	N^{nbr}	f_{enc}	N_{sw}^{HB}	N_{ww}^{HB}	f_{ww}^{HB}	$N_{ww,lost}^{HB}$
Specific Apolar Protein Sites											
Abl Site 13	A.En.U	−3.22	−5.32	−8.55	−1.97	3.10	0.41	0.00	2.52	81.50	1.37
Abl site 16	A.En.U	−2.48	−5.87	−8.35	−2.21	2.22	0.58	0.00	2.04	91.87	1.93
FXa site 21	A.Fr.U	−2.63	−6.10	−8.73	−1.31	3.34	0.37	0.00	2.22	0.66	1.22
FXa site 23	A.Fr.U	−2.71	−6.09	−8.80	−1.33	3.73	0.30	0.00	2.39	0.64	0.98
Specific Neutral Polar Protein Sites											
Abl site 6	P.En.U	−5.75	−3.15	−8.91	−2.19	1.14	0.78	1.79	1.02	0.89	2.61
FXa site 32	P.Fr.U	−1.95	−6.99	−8.94	−1.08	4.52	0.18	0.83	2.42	0.53	0.59
Specific Charged Protein Sites											
AChE site 0	C.En.F	−10.69	−1.89	−12.58	−1.72	1.12	0.79	2.12	0.89	0.79	2.62
Casp site 0	C.Fr.F	−13.78	0.55	−13.22	−0.77	2.00	0.62	2.92	1.01	0.51	2.06
Casp site 4	C.Fr.F	−11.57	1.16	−10.41	−0.20	3.29	0.37	1.95	1.11	0.34	1.24

structure if its value of $\overline{E}_{ww}^{nbr}$ is less than that of a water in bulk, and it is considered to have "frustrated" structure if it is greater. Thus, apolar, polar, and charged sites with enhanced water structure are termed A.En, P.En, and C.En, respectively, while sites where water structure is frustrated are termed A.Fr, P.Fr, and C.Fr. Each of these categories is then further divided according to whether the total water energy, E_{tot}, is more or less favorable than the value of E_{tot} for a water molecule in the bulk; category A.En thus separates into A.En.F and A.En.U, and so on. We used a t test to determine whether the total water energy, E_{tot}, of a hydration site is truly distinguishable from the bulk. This resulted in only 11 of 218 sites with E_{tot} indistinguishable from bulk, based on the P-value cutoff of 0.01. Since classifying these few sites as bulk-like, as opposed to favorable or unfavorable, did not affect our conclusions, we left them in their original categories, F or U, depending on whether their value of E_{tot} was less or greater than that of bulk water. Finally, we also examined the properties of water sites, in proteins, that are known to be displaced by small ligands. Detailed results for all of the hydration sites summarized here are provided in the Supporting Information.

3.1. Water at Apolar Surfaces. *3.1.1. Simple Apolar Molecules.* Methane and the outer surface of the synthetic host molecule CB7 represent simple, convex, apolar surfaces. For methane, discrete hydration sites could not be identified, so water molecules within its first solvation shell were studied. The first solvation shell was considered to include all waters within 5.4 Å of the methane carbon, because this distance corresponds to the first minimum of the carbon−oxygen radial distribution function.[47] For CB7, seven equivalent hydration sites were identified on the exterior surface, each near the center of one of the seven glycoluril units (Figure 2, top left), and the present results are averages over these seven sites. As detailed in Table 2, water molecules at both methane and CB7 have fewer first shell neighbors (N^{nbr}) than do water molecules in the bulk, but their interactions with these neighbors are enhanced, as indicated by more favorable average energetic interactions per first-shell neighbor ($\overline{E}_{ww}^{nbr}$) and a greater fraction of neighbors hydrogen bonding to the site water (f_{ww}^{HB}), relative to the bulk. Thus, these waters are both of type A.En. Their mean interactions with all other water molecules (E_{ww}) are somewhat less favorable than that of bulk water, but water at these sites also forms weak but favorable interactions with the solutes (E_{sw}). The net effect is that the total energy (E_{tot}) of these waters remains close to that of bulk water; however, because their values of E_{tot} are slightly above and below that of bulk water, methane and CB7 waters must be classified as type A.En.U and A.En.F, respectively.

The results are similar, but accentuated, for water sites centered above and below the aromatic ring of phenylalanine (Figure 2, bottom left; Table 2). Thus, as for waters near methane and CB7, the mean interaction and fractional hydrogen bonding with first-shell neighbors, $\overline{E}_{ww}^{nbr}$ and $\overline{E}_{ww}^{HB}$, are enhanced relative to the bulk. In addition, because the local surface of the phenyl group is flat, rather than convex like methane or the outer surface of CB7, water at the phenyl surface is more enclosed, so that f_{enc}, is considerably higher, and it has fewer first-shell neighbors, so that N^{nbr} is considerably lower. Not surprisingly, then, the overall interaction of water at the phenyl surface with other waters, E_{ww}, is less favorable than water at methane or CB7. Thus, although water at the phenyl group forms relatively favorable interactions with the solute, as indicated by a low value of E_{sw}, the net effect is that the overall energy, E_{tot}, is unfavorable relative to the bulk, making this clearly a site of type A.En.U.

The water−water interactions for water molecules at these simple apolar surfaces may be further analyzed by examining the distributions, $\rho_i(E_n)$, of their first-shell water−water interaction energies (Figure 2, middle). The area under the curve is somewhat smaller than that of bulk water (see figure) for methane and CB7, and considerably smaller for Phe, consistent with the numerical results summarized above. In addition, the reduction in area stems almost entirely from the parts of the graphs to the right of the peak, indicating selective depletion of interactions that are only weakly favorable and outright unfavorable; that said, some depletion to the left of the peak can be discerned in the case of the phenyl site. In contrast, the parts of the graph to the left of the peak track the bulk graph rather closely, indicating retention of the most favorable water−water interactions. Figure 2 (right) shows the normalized probability distributions of first shell interaction energies for the same systems and illustrates how water structure is enhanced. The curves for the hydrophobic groups have higher peaks on the left, which shows that water on all three of the hydrophobic surfaces has a higher probability of forming strong pair interactions than water in the bulk and correspondingly there is a lower probability of forming weaker water pair interactions, as indicated by the lower height on the right side of the curves.

The particularly marked reduction in neighbor count, N^{nbr}, and in net water−water interactions, E_{ww}, for water at the center of the phenyl group may be rationalized, at least in part,

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

by the fact that the most probable orientation of water molecules at this site point an OH group directly at the center of the ring (Figure 2). This means that only about three neighboring waters can form stabilizing hydrogen bonds with it. Accordingly, the value of $N^{HB}_{ww,lost}$ is precisely 1.00, in contrast with lower values of this quantity observed for water at the surfaces of methane and CB7 (Table 2). This situation contrasts with the convex surfaces of methane and CB7, where the most probable water orientations allow more than three neighbors to interact favorably. It is also consistent with prior observations that water molecules with unsatisfied hydrogen bonding potential at hydrophobic surfaces frequently point an OH group toward the surface, forming what are termed dangling hydrogen bonds.[22,48,49]

Overall, the solvation of convex surfaces is consistent with a traditional view of hydrophobic hydration,[30−32] in which, at biological temperatures, there is little enthalpy cost for solvating a hydrophobic surface, and the free energy cost is attributed to the entropy penalty arising from the enhancement of water structure around the hydrophobic moiety. The solvation of the phenyl group suggests that even on flat surfaces there may be an enthalpic penalty to solvation as well.[50,51] All of these sites are classified as A.En, indicating enhanced structure, rather than A.Fr, which would indicate frustrated structure. It is also worth noting that, in a previous study, we calculated the first order entropic cost of hydrating the outer surface of CB7 to be 0.60 kcal/water molecule.[12] Although tools for estimating the full contribution of water−water correlation to hydration entropy are not yet available, accounting for the increased correlation at surfaces should further supplement the first-order term entropic penalty.

3.1.2. Hydration of Apolar Surfaces in Protein Binding Pockets. A total of 56 hydration sites in the protein pockets studied here are classified as A (apolar type), based on the somewhat arbitrary but physically reasonable criterion that they average fewer than 0.1 hydrogen bonds with the protein, i.e., that $N^{HB}_{sw} < 0.1$. The bulk of these A sites (49 of the 56) are like those on the apolar simple solutes in the sense that they are further classified as having enhanced water structure, because their values of $\bar{E}^{nbr}_{ww}$ are less than that of bulk water. However, the protein A.En sites tend to have even more favorable values of $\bar{E}^{nbr}_{ww}$ than the simple solute ones (Table 2; see A.En.F and A.En.U rows), indicating even more enhanced water structure, on average. They also tend to have more favorable water−solute interactions, E_{sw}, which is not surprising, given the potential for more intimate contacts with the solute. Perhaps surprisingly, however, the average degree of enclosure, as measured by f_{enc}, is not especially high, consistent with average losses of at most 1.00 hydrogen bond between these sites and other waters ($N^{HB}_{ww,lost}$ in Table 2). The total energies of these sites, E_{tot}, remain within about 0.6 kcal/mol of the bulk. Overall, then, the A.En sites in the proteins may be viewed as similar to the A.En sites on the simple solutes but with generally more pronounced deviations from bulk properties.

However, there are also seven A sites in the binding pockets which, unlike water at the simple solutes, have values of $\bar{E}^{nbr}_{ww}$ which are unfavorable relative to the bulk, ranging up to −1.23 kcal/mol, and which therefore are categorized as A.Fr, rather than A.En sites. These, on average, have considerably more favorable interactions with the protein (despite the lack of water−protein hydrogen bonding), and less favorable water−water interactions, relative to the A.En sites (Table 2). The energetically favorable ones (A.Fr.F) average total energies only a little more favorable than that of the bulk, but the energetically unfavorable ones (A.Fr.U) deviate further, on average. The A.Fr sites represent a breakdown of the traditional view of hydrophobic solvation, because the loss of water neighbors at these sites is no longer accompanied by enhanced interactions with the remaining neighbors but rather by at least a modest weakening of these interactions.

It is of interest to inspect and further characterize some of the highest energy A sites in the protein pockets. Accordingly, Figure 3 (top left) shows highly occupied water configurations

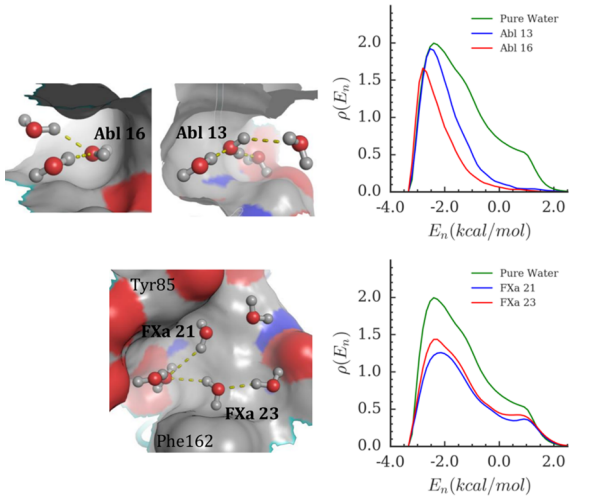


Figure 3. High energy hydration sites in apolar surface pockets. Top row: Abl 13 and Abl 16. Bottom row: FXa 21 and FXa 23. Left: Representative water configurations in labeled and neighboring hydration sites in the context of the active-site molecular surface (gray, carbon; red, oxygen; blue, nitrogen). Right: Number density distribution of the interaction energy of hydration site water molecules with their first shell neighbors, compared with the same distribution for pure water.

for hydration sites 13 and 16 in Abl kinase, two A.En.U sites for which $\bar{E}^{nbr}_{ww}$ is particularly favorable yet E_{tot} is particularly unfavorable (Table 3). Highly occupied configurations of several nearby sites are also shown. These sites are positioned in hydrophobic (gray surface) concavities, within hydrogen-bonding distance of other hydration sites (unlabeled), which afford enhanced first-shell interactions and thus account for the structural enhancement of sites 13 and 16. The selective loss of moderately favorable and outright unfavorable neighbor interactions, relative to the bulk, is clearly evident in the energy distributions of these two sites (Figure 3, top right). The particularly marked depletion for site 16 is consistent with the reduced number of neighbors evident in the top left panel of the figure. The energy distribution functions for these two sites may be viewed as strongly enhanced versions of the corresponding distributions for sites at methane and CB7 (Figure 2, top right).

The S4 pocket of FXa contains two sites of the nontraditional type A.Fr.U. These sites, 21 and 23, are neighbors of each other (Figure 3, bottom left) but are simultaneously occupied in only 18% of the MD frames; , in those frames in which they are both occupied, their mean interaction energy of −0.55 kcal/mol is not particularly favorable. This relatively unfavorable interaction presumably stems in part from competing interactions of these waters with the aromatic surfaces on which they reside (see below) and with other neighboring

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

water sites (Figure 3, bottom left). The frustrated first-shell interactions of these waters correspond to their mildly unfavorable values of $\overline{E}_{ww}^{nbr}$ (Table 3), relative to the bulk, and lead to their A.Fr categorization. Interestingly, sites 21 and 23 are located on the aromatic surfaces of Tyr85 and Phe162, respectively, and they each direct an OH group toward the surface, as seen for Phe in solution (above). The broad loss of first-shell interactions at both sites is evident in the distributions of first-shell water energies associated with FXa 21 and 23 (Figure 3, bottom right). Here, the depletion is relatively uniform, in contrast to the selective depletion of less favorable interactions seen for the two A.En.U sites in Abl kinase (Figure 3, top right). The present pattern is consistent with the lack of structure enhancement for these two FXa sites and resembles what was observed for Phe in solution (Figure 2, bottom right) but with greater and more uniform depletion.

3.2. Water at Neutral, Polar Surfaces. *3.2.1. Simple Polar Molecules.* The asparagine (Asn) side chain serves as a reference model for the hydration of simple, neutral polar solutes. This amide group induces four different hydration sites, for which representative configurations are shown in Figure 4

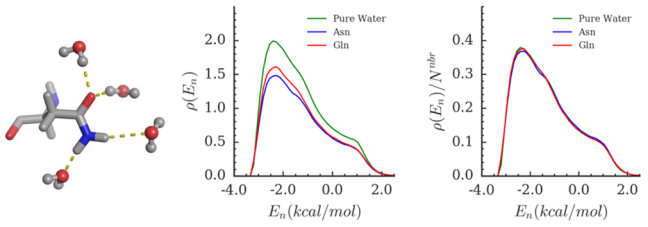


Figure 4. Hydration of the amide side chain of Asn. Left: The most probable water configurations in hydration sites forming hydrogen bonds with the side chain (only shown for Asn). Middle and right: normalized probability distribution and number density distribution of the interaction energy of hydration site water molecules with their first shell neighbors, compared with the same distribution for pure water.

(left). Because the properties of these are very similar to each other, Table 2 and the graphs in Figure 4 present averages over the four sites. (Calculations for the glutamine side chain yielded very similar results.) For these sites, the value of $\overline{E}_{ww}^{nbr}$ remains similar to that of bulk water, rather than becoming more favorable like water at the simple apolar solutes. In addition, the net water−water interaction becomes less favorable than for the simple apolar solutes; however, this loss of stabilizing interactions is more than compensated by the more favorable solute−water interaction, E_{sw}, consistent with the establishment of solute−water hydrogen bonding. The net effect is a total energy, E_{tot}, more favorable than that for the apolar sites and bulk. Also, whereas water in the simple apolar sites was found to have fewer hydrogen bonds, particularly in the case of Phe, for waters at the Asn sites, the net number of hydrogen bonds ($N_{sw}^{HB} + N_{ww}^{HB} = 3.39$) is essentially the same as that of bulk water (3.33). Thus, hydrogen bonding to the amide group replaces the lost hydrogen bonding with other water molecules.

The distribution of first-shell interaction energies of the Asn hydration sites is depleted relative to bulk water (Figure 4, right), as expected due to volume exclusion; less expected is that the shape of the distribution strongly resembles that of bulk water. This marked similarity becomes particularly clear upon normalizing the energy distributions (Figure 4, middle). Indeed, the amide group produces almost no change in water structure, by this measure. This is in contrast to the other types of simple solutes examined here, for which the normalized energy distributions deviate more from the bulk. We speculate that this preservation of structure is related, in part, to the fact that the amide group, like a water molecule, offers both hydrogen bond donors and acceptors.

3.2.2. Hydration of Polar Surfaces in Protein Binding Pockets. Protein hydration sites where water molecules form greater than 0.1 hydrogen bonds with the neutral donor or acceptor groups on the protein surface are categorized as polar (P). As detailed in Table 2, there are about twice as many P sites (103) as apolar ones (56) in the binding pockets studied here. The water at the P sites averages about one hydrogen bond with the protein, consistent with their categorization as polar. As for the A sites, a majority of P sites show enhanced water structure. On the other hand, the fraction of P sites with frustrated water (P.Fr, 27%) is double the corresponding value for the A sites (A.Fr, 13%), consistent with a view that a hydrophobic surface has a particular tendency to generate enhanced water structure. The average degree of frustration at the P.Fr sites is greater than that at the A.Fr sites, in the sense that they have less favorable values of $\overline{E}_{ww}^{nbr}$, but structure at the P.En sites is enhanced about the same as at the A.En sites. Perhaps surprisingly, given that waters at P sites can form stronger interactions with the protein, the fractions of P and A sites that are net unfavorable (based on E_{tot}) are similar, at about 40%. However, the unfavorable P and A sites differ in that the P sites are much more enclosed on average (f_{enc} about 0.40 vs 0.24). Despite being in a polar environment, energetically unfavorable hydration sites lose more water−water hydrogen bonds due to enclosure than they form with the surface; thus, $N_{sw}^{HB} + N_{ww}^{HB}$ is less than the mean number of hydrogen bonds, 3.33, of a water in the bulk. Inspection of the P.Fr.U sites, in particular, reveals the consistent presence of a first-shell water neighbor which forms competing interactions with another patch of the protein's surface.

Examples of two energetically unfavorable P sites are now further analyzed. The first is site 6 in Abl kinase, a P.En.U site where a representative water configuration shows hydrogen bonding with the protein and a single water neighbor (Figure 5,

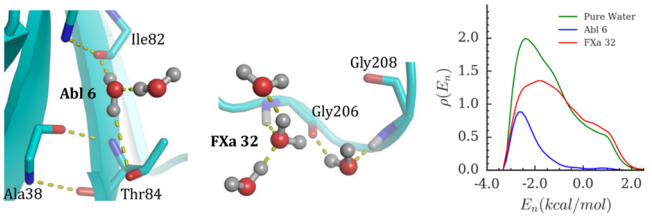


Figure 5. High energy hydration sites in polar surface pockets. The most probable water configurations of labeled and neighboring hydration sites and their interactions with the protein surface are shown for sites Abl 6 (left) and FXa 32 (middle). Number density distribution of the energy of hydration site water molecules with their first shell neighbors, compared with the same distribution for pure water (right).

left). The value of $\overline{E}_{ww}^{nbr}$ is strongly negative, at −2.19 kcal/mol (Table 3), indicating strongly enhanced water structure, but there is only one water−water hydrogen bond, on average, and the net water−water interaction energy, E_{ww}, is correspondingly weak, at only −3.15 kcal/mol, while the water−protein interaction is strong, with E_{sw} of −5.75 kcal/mol. In summary, despite forming strong solute−water interactions, the enhanced

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

water structure at this site is comparable to that typical of an enclosed apolar site, such as Abl sites 13 and 16 (Table 3); indeed, the distribution of first-shell energies (Figure 5, right) is depleted of moderately favorable and unfavorable interactions much as seen previously for Abl 13 and 16 (Figure 3, top right).

Hydration site FXa site 32 is categorized as type P.Fr.U, where, as mentioned above, a water neighbor is constrained by hydrogen bonds to the protein, leading to frustration at site 32 (Figure 5, middle). The first-shell energy distribution (Figure 5, right) is broadened, in striking contrast to that of Abl site 6 (Figure 5, right), with notable enhancement of the probability of forming unfavorable interactions. As a consequence, $\overline{E}_{ww}^{nbr}$ is much less favorable (−1.08 kcal/mol) than that for Abl site 6 (−2.19 kcal/mol; Table 3). However, the overall depletion is not marked, due to only a modest level of enclosure, f_{enc} = 0.18. Thus, although FXa site 32 is surrounded by nearly a full complement of first hydration shell waters, its interactions with these waters are quite unfavorable relative to the bulk.

3.3. Water at Charged Surfaces. *3.3.1. Simple Charged Molecules.* Ionized aspartate (Asp) and arginine (Arg) side chains were chosen as simple reference solutes for the hydration of charged surfaces. Six similar sites, in which the water donates one hydrogen bond to a carboxylate oxygen, were found around Asp (Figure 6, left), and we present

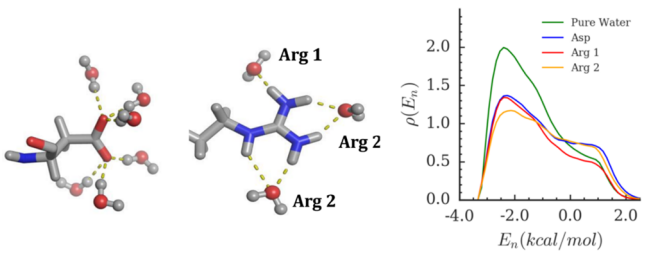


Figure 6. Hydration of charged side chains of Asp and Arg. Left and middle: The most probable water configurations in hydration sites forming hydrogen bonds with the side chain of Asp and Arg, respectively. Right: Number density distribution of the interaction energy of hydration site water molecules with their first shell neighbors, compared with the same distribution for pure water.

averages over these sites. Two types of Arg hydration sites were found (Figure 6, middle): Arg 1 comprises one hydration site where water accepts one hydrogen bond from the guanidinium group; Arg 2 comprises two hydration sites, where the water typically accepts two hydrogen bonds from the solute, and the Arg 2 results are averages over both Arg 2 sites.

As detailed in Table 2, all three sites—Asp, Arg 1, and Arg 2—have frustrated water structure, based on values of E_{ww}^{nbr} considerably less favorable than the bulk; they also have total energies, E_{tot}, much more favorable than that of bulk water. Accordingly, they are all classified as sites of type C.Fr.F. They also have particularly favorable solute−water interactions, ranging stronger than −6 kcal/mol. The overall pattern of energetics is similar to that of the neutral polar solute, Asn, but the deviations from the bulk are considerably stronger for these charged sites. The first-shell energy distributions accordingly deviate quite strongly from those of bulk water (Figure 6, right), with marked reductions in the peak of favorable interactions and actual elevation of the repulsive shoulder at about 1 kcal/mol, in the case of the Asp and Arg 2 sites. These changes, including the new repulsions, presumably represent the cost in unfavorable water−water interactions of orienting waters to form strongly stabilizing interactions with the charged solutes.

3.3.2. Hydration of Charged Surfaces in Protein Binding Pockets. Hydration sites where the water molecules averaged more than 0.1 hydrogen bond with a charged side chain were categorized as charged (C). As detailed in Table 2, there are about as many charged sites, 60, as apolar ones, 56. Consistent with the results for simple charged solutes (above), the fraction of charged sites with frustrated water structure is particularly high, at 70%, and these C.Fr sites are particularly frustrated, with $\overline{E}_{ww}^{nbr}$ averaging −0.96 kcal/mol. The corresponding fractions are 27 and 13% for the P.Fr and A.Fr protein sites (above), and the average values of $\overline{E}_{ww}^{nbr}$ for the A.Fr and P.Fr sites reach only −1.16 kcal/mol. On the other hand, the average enhancement of $\overline{E}_{ww}^{nbr}$ for the C.En sites, −1.7 kcal/mol, is similar to those of the A.En and P.En sites. Unlike the case for both A and P sites, there are no C sites that are overall unfavorable energetically, and the values of E_{tot} for C sites run lower (more negative) than those for either the A or P sites. Of all categories studied here, the C.Fr.F sites form the most favorable average interactions with the solute, and the least favorable interactions with other water molecules. All of these features go to support the broad picture that strong water−solute interactions at charged surfaces tend to break water structure.

As an example, we examine the severe disruption of local water structure at Casp sites 0 and 4, which together contribute to the hydration of two arginine side chains, a histidine side chain, and a glutamine side chain (Figure 7, left), in the S1 specificity pocket of Caspase 3. (Two fractional (0.34, 0.36) hydrogen bonds between Casp 0 and Arg 207 are not depicted because the highly occupied water position of Casp 0 shown

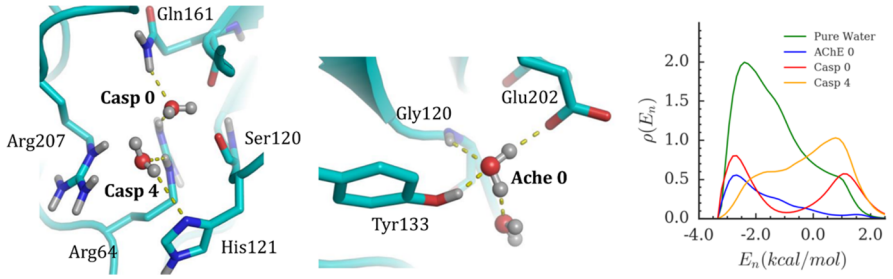


Figure 7. Hydration sites in charged surface pockets. The most probable water configurations of labeled and neighboring hydration sites and their interactions with the protein surface are shown for sites Casp 0 and 4 (left) and AChE 0 (middle). Number density distribution of the interaction energy of hydration site water molecules with their first shell neighbors, compared with the same distribution for pure water (right).

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

does not satisfy hydrogen bonding criteria for these; see Methods.) The first-shell energy distributions of these two sites (Figure 7, right) show not only marked, selective depletion of favorable interactions but also marked enhancement of unfavorable interactions, as indicated by the peaks near 1.5 kcal/mol, and a value of −0.77 kcal/mol for $\overline{E}_{ww}^{nbr}$ (Table 3). This mirrors the similar but less marked addition of a repulsive shoulder in the corresponding graphs for the Arg 2 and Asp hydration sites (Figure 6). Similarly, frustrated water structure is observed at a number of other C.Fr sites, such as FXa sites 6 and 12; see the Supporting Information. It is also worth mentioning that the total water−water energy, E_{ww}, is positive for sites Casp 0 and 4. This is due to unfavorable water−water energy in the long range (>3.5 Å) compensating for favorable water−water interactions in the first shell. Eight other hydration sites display a similar balance between short- and long-ranged interactions (see the Supporting Information for Casp and Abl sites). It appears that the strong interactions with the protein surface cause misalignment between water pairs that are distant neighbors (>3.5 Å). In this way, strong alignment of water molecules by charged solutes can lead to strong new water−water repulsions.

Nonetheless, there are also C.En sites where local water structure is enhanced at a charged surface (Table 2). For example, Figure 7 (middle) shows a representative configuration of water at AChE site 0. The site water forms multiple hydrogen bonds with the protein, along with a hydrogen bond with its only first shell neighbor. The resulting value of $\overline{E}_{ww}^{nbr}$ is quite favorable, at −1.72 kcal/mol (Table 3). Correspondingly, the first shell energy shows broad depletion of neighbors but retention of the most favorable interactions on the left-hand side of the graph (Figure 7, right).

3.4. Characteristics of Displaced Hydration Sites in Proteins. Co-crystal structures with bound small molecule ligands are available for all of the proteins studied here, and some of the bound ligands overlap with the hydration sites examined. Although a statistically rigorous analysis is not possible, due to the anecdotal nature of the data set, it is nonetheless of interest to consider whether any of the site properties studied here correlate with the probability that a ligand will displace the site, and to inquire about the nature of the ligand substituents that displace various types of hydration sites. We also note that the positions of hydration sites are determined from the ligand-bound conformation of the receptor, which may be different from the apo structure. For simplicity, we restrict our calculations to the ligand-bound conformation and do not investigate the effect of protein reorganization on hydration site properties. The utility of such analyses stems from two observations. First, multiple ligands (e.g., from a congeneric series) may bind to essentially the same receptor conformation, so analyzing the hydration of this conformation can be useful to help understand how water may influence the relative thermodynamics of ligand binding. Second, the thermodynamics of protein−ligand binding can be analyzed in terms of a first step of organizing the solvated apo conformation of the protein into the solvated ligand-bound conformation, followed by a second step in which the ligand binds the organized protein. The present study focuses on the second step.

The resulting numbers of sites and displaced sites, as well as the ratio of the two, are reported by site category in Table 4. Perhaps the most striking, if not surprising, trend is that the percentage of sites displaced is greatest for apolar sites (63%), intermediate for polar sites (49%), and lowest for charged sites (34%). A closely related trend is that 59% of energetically unfavorable sites are displaced, compared with only 43% of favorable sites (Table 4). However, enhanced (type En) sites are more frequently displaced than the frustrated (type Fr) sites. This effect is related to the observation that frustrated sites are more frequently found in polar and charged environments, which we noted earlier to be less frequently displaced than the apolar sites. Some weak trends may also be discerned for the average properties of displaced versus nondisplaced sites (Table 2); in particular, the average solute−water energy is somewhat lower for displaced sites, and, perhaps most strikingly, the average number of water−protein hydrogen bonds is smaller for displaced sites, 0.69, than for nondisplaced sites, 0.93.

Table 4. Summary Statistics, by Hydration Site Categories, for All Sites and Sites Displaced by Bound Ligands

type	all sites	displaced sites	fraction displaced
A.En.F	32	18	0.56
A.En.U	17	12	0.71
A.Fr.F	2	1	0.50
A.Fr.U	5	4	0.80
P.En.F	45	20	0.44
P.En.U	30	18	0.60
P.Fr.F	17	7	0.41
P.Fr.U	11	3	0.27
C.En.F	18	7	0.39
C.Fr.F	41	14	0.34
all En	142	75	0.53
all Fr	76	29	0.38
all A	56	35	0.63
all P	103	48	0.47
all C	59	21	0.36
all F	155	67	0.43
all U	63	37	0.59

As might be expected, all of the displaced apolar hydration sites are occupied by apolar ligand substituents, and the same is true for the displaced polar hydration sites that average 0.5 or fewer hydrogen bonds with the surface. As an example, Figure 8 (left) shows a nonpolar ligand substituent overlain on the FXa 21 and 23 apolar sites previously shown in Figure 3 (lower left). The ligand's hydrophobic group packs against the enclosed hydrophobic surface of the S4 pocket of factor Xa, and displaces multiple high energy water molecules. Perhaps more surprising is that there are 21 cases of polar and charged hydration sites where water averages at least 0.5 hydrogen bonds to the protein, yet is displaced by a ligand functional group that does not form the same hydrogen bond(s). The Abl 6 site considered above (Figure 5, left, and Figure 8, middle) is an example of this scenario: the water in this site averages 1.8 hydrogen bonds with the protein (Table 3), making hydrogen bonds with each of the Thr84 side chain and Ile82 backbone 77% of the time, and also making a hydrogen bond with the Ala38 backbone carbonyl 24% of the time. (The hydrogen bond with the backbone carbonyl is not shown in Figure 5 because it is not formed by the most probable water configuration.) The water-displacing ligand places a methyl group proximal to the Ile82 backbone carbonyl, and thus fails

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

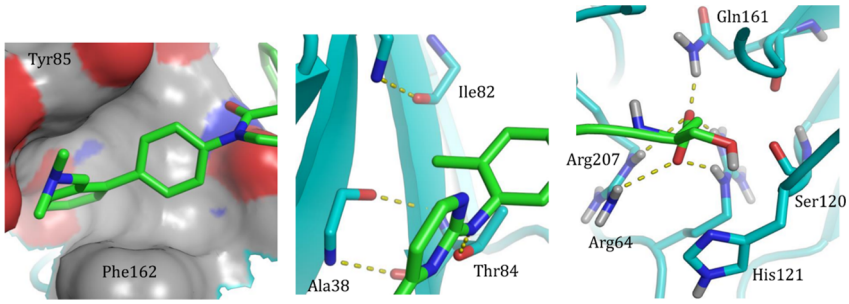


Figure 8. Displacement of hydration sites by ligand functional groups. Apolar ligand moieties displace FXa 21 and 23 (left) and Abl 6 (middle), while a carboxylate displaces sites Casp 0 and 4 (right).

to make a hydrogen bond that is formed by the displaced water. However, the ligand does use a different functional group to form a hydrogen bond with the Thr84 side chain hydroxyl. This case might be rationalized, at least in part, by the fact that the protein groups to which the site water forms its hydrogen bonds also make hydrogen bonds with other protein groups, and thus may substitute greater internal stabilization for interactions with water when the ligand binds. However, there are other instances in which a P-type water is displaced by a ligand which fails to substitute new ligand−protein hydrogen bonds for the pre-existing water−protein hydrogen bonds. For example, water at the P.En.U site FXa 14 with occupancy 0.67 has a mean interaction with the protein (E_{sw}) of −3.1 kcal/mol and averages 1.4 hydrogen bonds with the protein, but there are at least 11 ligands in the PDB database, which displace it with a methyl group.

At the other extreme, we observe that ligands displacing P.Fr and C.Fr hydration sites always form hydrogen bonds to the protein, replacing or surpassing those which had been made by the displaced water molecules. For example, a ligand carboxylate group displacing the Casp 0 and Casp 4 sites presented above (Figure 7) simultaneously accepts five hydrogen bonds from the nearby Arg and Gln side chains (Figure 8, right), matching the average of five hydrogen bonds previously formed by the two water sites (Table 3). This particular example is quite striking, as in the displaced water molecules not only have unfavorable values of $\overline{E}_{ww}^{nbr}$, relative to bulk water, but also form suboptimal hydrogen bonds with the surface compared to the ligand. Thus, despite their location at the surface of several charged side chains, their interactions are suboptimal, and they are displaceable. As noted above, Casp 4 forms two fractional hydrogen bonds with Arg 207, and Casp 0 does not hydrogen bond with Arg 207. On the other hand, the X-ray crystallographic pose of the ligand indicates that the carboxylate functional fully forms these two hydrogen bonds. The additional hydrogen bonds formed by water with His121 and Ser120 are replaced either by ligand interactions involving a different ligand functional group or by structural water molecules observed in the X-ray structure. Taken together, these observations indicate that the carboxylate functional group satisfies the hydrogen bonding groups on the protein surface better than displaced water.

4. CONCLUDING REMARKS

We have described measures of local water structure and energetic stability to characterize how water molecules reorganize on simple solute surfaces and on protein active-site surfaces. Our reference systems show that local water structure is generally enhanced when water makes weak interactions with the surface, remains unchanged when it makes interactions with the surface comparable to that made with other water molecules in the bulk, and is slightly frustrated when it makes stronger interactions with the surface. In the protein active sites studied here, most of the hydration site water molecules behave similarly, in that they maintain a favorable total energy relative to the bulk, while having either enhanced energetic interactions with water neighbors, strong interactions with the protein that compensate for unfavorable energetic interactions with the water neighbors, or both. However, exceptional cases are observed, among both apolar and polar hydration sites, where water's ability to structure itself in such a manner breaks down, resulting in energetically unfavorable water molecules, relative to bulk water. These water molecules (categorized as A.Fr.U and P.Fr.U), in addition to losing neighbors on the surface, are unable to maintain favorable interactions with some of the neighbors which are present, resulting in a frustrated local structure. The frustration results from surface constraints that restrict neighboring water molecules from adopting energetically favorable orientations. Some examples of such constraints, found in the present study, include neighboring water molecules either sandwiched between two aromatic walls or simultaneously hydrogen bonded to the surface. We also note the presence of other high energy water molecules that have enhanced local structure (categorized as A.En.U and P.En.U) but, due to a high degree of enclosure and weak interactions with the surface, remain energetically unfavorable relative to bulk water. Overall, water structure in high energy hydration sites is characterized either by significant loss of neighbors and/or favorable interactions with the remaining neighbors, combined with surface interactions that insufficiently compensate these energetic losses. We expect the displacement of such water molecules to afford an enthalpic advantage, in addition to a favorable entropic contribution. Such cases of enthalpically unfavorable water molecules have been highlighted in previous computational and experimental studies of protein−ligand systems,[2−5,52,53] host−guest complexes,[54−57] and model cavity−ligand systems.[6,50,51]

Comparison of protein−water with protein−ligand contacts for displaced hydration sites shows that, as expected, apolar and weakly polar hydration sites tend to be displaced by hydrophobic functional groups, and strongly polar and charged hydration sites tend to be displaced by hydrophilic functional groups. However, a subset of polar hydration sites is displaced by nonpolar ligand moieties. By definition, a ligand that is fully complementary to the surface is expected to sterically fit and make complementary electrostatic interactions. Hence, it is reasonable to expect that, in a polar hydration site, ligands

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

would make the same hydrogen bonds with the protein that are formed by the water molecules. It is perplexing to see that ligands do not form the hydrogen bonds that are made by water molecules. This prompts further investigation of whether hydrophilic ligands would displace water molecules from these hydration sites with greater affinity or whether the lost hydrogen bonds are compensated by other contributions from ligand binding.

At charged hydration sites, water molecules always have a favorable total energy relative to the bulk, at least in the current data set, due to favorable interactions with both water neighbors and the protein (C.En hydration sites), or to a strong compensation between lost water−water and gained protein−water interactions (C.Fr hydration sites). Interestingly, water molecules at C.Fr hydration sites often have unfavorable interactions with their first-shell water neighbors, far in excess of those typical of water molecules interacting with charged side chains in solution; such unfavorable interactions are almost nonexistent for water molecules in the bulk. These unfavorable pair interactions typically arise from the fact that the neighboring water forms competing hydrogen bonds with nearby charged groups. Although similar constraints occur for water molecules interacting with neutral polar groups, the effect is much stronger for charged groups.

We conjecture that the presence of water with highly frustrated local structure has implications for how the contributions of polar ligand−protein contacts ought to be assessed. In general, any hydrogen bond formed between the protein and the ligand replaces hydrogen bonds that each molecule made with water; thus, if a hydrogen bond is diagrammed as a dash, the number of hydrogen bonds is balanced in the following simplified equation:

$$\mathrm{P{-}W + L{-}W \rightarrow P{-}L + W{-}W}$$

Here, P, L, and W indicate protein, ligand, and water molecules, respectively. As a result, the net affinity contribution of hydrogen bonds formed at protein−ligand interfaces is often considered to be minimal. However, the liberation of water molecules that make unfavorable interactions with their neighbors, from a protein surface into bulk water, may result in a net favorable contribution to the formation of a protein−ligand hydrogen bond. This situation is described by the following simplified equation, which differs only in the replacement of P−W by P−W*, where the asterisk indicates water with relatively unfavorable interactions with other water molecules.

$$\mathrm{P{-}W^{*} + L{-}W \rightarrow P{-}L + W{-}W}$$

Here, the formation of a protein−ligand hydrogen bond may be energetically favorable, in contrast to the prior equation with its simpler balance of hydrogen bonds. Identification of such hydration sites based on local structure measures such as $\overline{E}_{\mathrm{nbr}}$ may, therefore, be a valuable tool to interpret and even predict the influence of hydrogen bonding on protein−ligand binding affinity.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.6b01094.

Table showing standard deviations associated with the average values of structural and energetic quantities reported for protein hydration site groups and figure showing normalized probability distribution plots corresponding to the number density distribution plots shown in Figures 3, 5, 6, and 7 (PDF)

Microsoft Excel spreadsheet containing detailed data for hydration sites belonging to different groups reported in the main text (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: thomas.kurtzman@lehman.cuny.edu. Phone: 718-960-8832.

Notes

The authors declare the following competing financial interest(s): M.K.G. has an equity interest in and is a cofounder and scientific advisor of VeraChem LLC.

■ ACKNOWLEDGMENTS

This publication was made possible in part by grants GM095417 and GM100946 from the National Institute of Health.

■ REFERENCES

(1) Snyder, P. W.; Lockett, M. R.; Moustakas, D. T.; Whitesides, G. M. Is It the Shape of the Cavity, or the Shape of the Water in the Cavity? *Eur. Phys. J.: Spec. Top.* **2014**, *223* (5), 853−891.

(2) Nguyen, C. N.; Cruz, A.; Gilson, M. K.; Kurtzman, T. Thermodynamics of Water in an Enzyme Active Site: Grid-Based Hydration Analysis of Coagulation Factor Xa. *J. Chem. Theory Comput.* **2014**, *10* (7), 2769−2780.

(3) Englert, L.; Biela, A.; Zayed, M.; Heine, A.; Hangauer, D.; Klebe, G. Displacement of Disordered Water Molecules from Hydrophobic Pocket Creates Enthalpic Signature: Binding of Phosphonamidate to the S_1'-Pocket of Thermolysin. *Biochim. Biophys. Acta, Gen. Subj.* **2010**, *1800* (11), 1192−1202.

(4) Biela, A.; Nasief, N. N.; Betz, M.; Heine, A.; Hangauer, D.; Klebe, G. Dissecting the Hydrophobic Effect on the Molecular Level: The Role of Water, Enthalpy, and Entropy in Ligand Binding to Thermolysin. *Angew. Chem., Int. Ed.* **2013**, *52* (6), 1822−1828.

(5) Barillari, C.; Taylor, J.; Viner, R.; Essex, J. W. Classification of Water Molecules in Protein Binding Sites. *J. Am. Chem. Soc.* **2007**, *129* (9), 2577−2587.

(6) Baron, R.; Setny, P.; McCammon, J. A. Hydrophobic Association and Vol.-Confined Water Molecules. In *Protein-Ligand Interactions*; Gohlke, H., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2012; pp 145−170.

(7) García-Sosa, A. T.; Firth-Clark, S.; Mancera, R. L. Including Tightly-Bound Water Molecules in de Novo Drug Design. Exemplification through the in Silico Generation of poly(ADP-Ribose)polymerase Ligands. *J. Chem. Inf. Model.* **2005**, *45* (3), 624−633.

(8) Cozzini, P.; Fornabaio, M.; Marabotti, A.; Abraham, D. J.; Kellogg, G. E.; Mozzarelli, A. Free Energy of Ligand Binding to Protein: Evaluation of the Contribution of Water Molecules by Computational Methods. *Curr. Med. Chem.* **2004**, *11* (23), 3093−3118.

(9) Levy, Y.; Onuchic, J. N. Water Mediation in Protein Folding and Molecular Recognition. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 389−415.

(10) Li, Z.; Lazaridis, T. Computing the Thermodynamic Contributions of Interfacial Water. In *Computational Drug Discovery and Design*; Baron, R., Ed.; Methods in Molecular Biology; Springer: New York, 2012; pp 393−404.

(11) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. Motifs for Molecular Recognition Exploiting Hydrophobic Enclosure in Protein−ligand Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (3), 808−813.

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

(12) Nguyen, C. N.; Young, T. K.; Gilson, M. K. Grid Inhomogeneous Solvation Theory: Hydration Structure and Thermodynamics of the Miniature Receptor cucurbit[7]uril. *J. Chem. Phys.* **2012**, *137* (4), 044101.
(13) Huggins, D. J. Application of Inhomogeneous Fluid Solvation Theory to Model the Distribution and Thermodynamics of Water Molecules around Biomolecules. *Phys. Chem. Chem. Phys.* **2012**, *14* (43), 15106−15117.
(14) Huggins, D. J.; Payne, M. C. Assessing the Accuracy of Inhomogeneous Fluid Solvation Theory in Predicting Hydration Free Energies of Simple Solutes. *J. Phys. Chem. B* **2013**, *117* (27), 8232−8244.
(15) Kovalenko, A.; Hirata, F. Three-Dimensional Density Profiles of Water in Contact with a Solute of Arbitrary Shape: A RISM Approach. *Chem. Phys. Lett.* **1998**, *290* (1−3), 237−244.
(16) Sindhikara, D. J.; Hirata, F. Analysis of Biomolecular Solvation Sites by 3D-RISM Theory. *J. Phys. Chem. B* **2013**, *117* (22), 6718−6723.
(17) Truchon, J.-F.; Pettitt, B. M.; Labute, P. A Cavity Corrected 3D-RISM Functional for Accurate Solvation Free Energies. *J. Chem. Theory Comput.* **2014**, *10* (3), 934−941.
(18) Bodnarchuk, M. S.; Viner, R.; Michel, J.; Essex, J. W. Strategies to Calculate Water Binding Free Energies in Protein−Ligand Complexes. *J. Chem. Inf. Model.* **2014**, *54* (6), 1623−1633.
(19) Ross, G. A.; Morris, G. M.; Biggin, P. C. Rapid and Accurate Prediction and Scoring of Water Molecules in Protein Binding Sites. *PLoS One* **2012**, *7* (3), e32036.
(20) Riniker, S.; Barandun, L. J.; Diederich, F.; Krämer, O.; Steffen, A.; van Gunsteren, W. F. Free Enthalpies of Replacing Water Molecules in Protein Binding Pockets. *J. Comput.-Aided Mol. Des.* **2012**, *26* (12), 1293−1309.
(21) *SZMAP 1.2.0.7*; OpenEye Scientific Software: Santa Fe, NM, 2013.
(22) Cheng, Y.-K.; Rossky, P. J. Surface Topography Dependence of Biomolecular Hydrophobic Hydration. *Nature* **1998**, *392* (6677), 696−699.
(23) Sharp, K. A. The Remarkable Hydration of the Antifreeze Protein Maxi: A Computational Study. *J. Chem. Phys.* **2014**, *141* (22), 22D510.
(24) Sharp, K. A.; Vanderkooi, J. M. Water in the Half Shell: Structure of Water, Focusing on Angular Structure and Solvation. *Acc. Chem. Res.* **2010**, *43* (2), 231−239.
(25) Li, Z.; Lazaridis, T. Water at Biomolecular Binding Interfaces. *Phys. Chem. Chem. Phys.* **2007**, *9* (5), 573−581.
(26) Ball, P. Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2008**, *108* (1), 74−108.
(27) Wong, S. E.; Lightstone, F. C. Accounting for Water Molecules in Drug Design. *Expert Opin. Drug Discovery* **2011**, *6* (1), 65−74.
(28) de Beer, S. B. A.; Vermeulen, N. P. E.; Oostenbrink, C. The Role of Water Molecules in Computational Drug Design. *Curr. Top. Med. Chem.* **2010**, *10* (1), 55−66.
(29) Mancera, R. L. Molecular Modeling of Hydration in Drug Design. *Curr. Opin. Drug Discovery Dev.* **2007**, *10* (3), 275−280.
(30) Frank, H. S.; Evans, M. W. Free Volume and Entropy in Condensed Systems III. Entropy in Binary Liquid Mixtures; Partial Molal Entropy in Dilute Solutions; Structure and Thermodynamics in Aqueous Electrolytes. *J. Chem. Phys.* **1945**, *13* (11), 507−532.
(31) Tanford, C. The Hydrophobic Effect and the Organization of Living Matter. *Science* **1978**, *200* (4345), 1012−1018.
(32) Kauzmann, W. Some Factors in the Interpretation of Protein Denaturation. *Adv. Protein Chem.* **1959**, *14*, 1−63.
(33) Chandler, D. Physical Chemistry: Oil on Troubled Waters. *Nature* **2007**, *445* (7130), 831−832.
(34) Gallicchio, E.; Kubo, M. M.; Levy, R. M. Enthalpy−Entropy and Cavity Decomposition of Alkane Hydration Free Energies: Numerical Results and Implications for Theories of Hydrophobic Solvation. *J. Phys. Chem. B* **2000**, *104* (26), 6271−6285.
(35) Weisberg, E.; Manley, P. W.; Breitenstein, W.; Brüggen, J.; Cowan-Jacob, S. W.; Ray, A.; Huntly, B.; Fabbro, D.; Fendrich, G.; Hall-Meyers, E.; et al. Characterization of AMN107, a Selective Inhibitor of Native and Mutant Bcr-Abl. *Cancer Cell* **2005**, *7* (2), 129−141.
(36) Cheung, J.; Rudolph, M. J.; Burshteyn, F.; Cassidy, M. S.; Gary, E. N.; Love, J.; Franklin, M. C.; Height, J. J. Structures of Human Acetylcholinesterase in Complex with Pharmacologically Important Ligands. *J. Med. Chem.* **2012**, *55* (22), 10282−10286.
(37) Fang, B.; Boross, P. I.; Tozser, J.; Weber, I. T. Structural and Kinetic Analysis of Caspase-3 Reveals Role for S5 Binding Site in Substrate Recognition. *J. Mol. Biol.* **2006**, *360* (3), 654−666.
(38) Adler, M.; Davey, D. D.; Phillips, G. B.; Kim, S. H.; Jancarik, J.; Rumennik, G.; Light, D. R.; Whitlow, M. Preparation, Characterization, and the Crystal Structure of the Inhibitor ZK-807834 (CI-1031) Complexed with Factor Xa. *Biochemistry* **2000**, *39* (41), 12534−12542.
(39) Shen, C.-H.; Wang, Y.-F.; Kovalevsky, A. Y.; Harrison, R. W.; Weber, I. T. Amprenavir Complexes with HIV-1 Protease and Its Drug-Resistant Mutants Altering Hydrophobic Clusters. *FEBS J.* **2010**, *277* (18), 3699−3714.
(40) Weber, P. C.; Ohlendorf, D. H.; Wendoloski, J. J.; Salemme, F. R. Structural Origins of High-Affinity Biotin Binding to Streptavidin. *Science* **1989**, *243* (4887), 85−88.
(41) Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comput. Chem.* **2005**, *26* (16), 1752−1780.
(42) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65* (3), 712−725.
(43) Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27* (3), 221−234.
(44) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; et al. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. *Proceedings of the ACM/IEEE SC 2006 Conference* **2006**, 43−43.
(45) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327−341.
(46) Luzar, A.; Chandler, D. Structure and Hydrogen Bond Dynamics of Water−dimethyl Sulfoxide Mixtures by Computer Simulations. *J. Chem. Phys.* **1993**, *98* (10), 8160−8173.
(47) Lazaridis, T. Solvent Reorganization Energy and Entropy in Hydrophobic Hydration. *J. Phys. Chem. B* **2000**, *104* (20), 4964−4979.
(48) Lee, C.-Y.; McCammon, J. A.; Rossky, P. J. The Structure of Liquid Water at an Extended Hydrophobic Surface. *J. Chem. Phys.* **1984**, *80* (9), 4448−4455.
(49) Du, Q.; Freysz, E.; Shen, Y. R. Surface Vibrational Spectroscopic Studies of Hydrogen Bonding and Hydrophobicity. *Science* **1994**, *264* (5160), 826−828.
(50) Setny, P.; Baron, R.; McCammon, J. A. How Can Hydrophobic Association Be Enthalpy Driven? *J. Chem. Theory Comput.* **2010**, *6* (9), 2866−2871.
(51) Baron, R.; Setny, P.; Andrew McCammon, J. Water in Cavity−Ligand Recognition. *J. Am. Chem. Soc.* **2010**, *132* (34), 12091−12097.
(52) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, *130* (9), 2817−2831.
(53) Snyder, P. W.; Mecinović, J.; Moustakas, D. T.; Thomas, S. W.; Harder, M.; Mack, E. T.; Lockett, M. R.; Héroux, A.; Sherman, W.; Whitesides, G. M. Mechanism of the Hydrophobic Effect in the Biomolecular Recognition of Arylsulfonamides by Carbonic Anhydrase. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (44), 17889−17894.
(54) Smithrud, D. B.; Wyman, T. B.; Diederich, F. Enthalpically Driven Cyclophane-Arene Inclusion Complexation: Solvent-Depend-

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX

ent Calorimetric Studies. *J. Am. Chem. Soc.* **1991**, *113* (14), 5420−5426.
(55) Peterson, B. R.; Wallimann, P.; Carcanague, D. R.; Diederich, F. Steroid Complexation by Cyclophane Receptors in Aqueous Solution: Substrate Selectivity, Enthalpic Driving Force for Cavity Inclusion, and Enthalpy-Entropy Compensation. *Tetrahedron* **1995**, *51* (2), 401−421.
(56) Schneider, H. J.; Kramer, R.; Simova, S.; Schneider, U. Host-Guest Chemistry. 14. Solvent and Salt Effects on Binding Constants of Organic Substrates in Macrocyclic Host Compounds. A General Equation Measuring Hydrophobic Binding Contributions. *J. Am. Chem. Soc.* **1988**, *110* (19), 6442−6448.
(57) Biedermann, F.; Nau, W. M.; Schneider, H.-J. The Hydrophobic Effect Revisited—Studies with Supramolecular Complexes Imply High-Energy Water as a Noncovalent Driving Force. *Angew. Chem., Int. Ed.* **2014**, *53* (42), 11158−11171.
(58) Lenz, A.; Ojamäe, L. Structures of the I-, II- and H-Methane Clathrates and the Ice-Methane Clathrate Phase Transition from Quantum-Chemical Modeling with Force-Field Thermal Corrections. *J. Phys. Chem. A* **2011**, *115*, 6169−6176.

DOI: 10.1021/acs.jpcb.6b01094
J. Phys. Chem. B XXXX, XXX, XXX−XXX