

Molecular Dynamics in Mixed Solvents Reveals Protein–Ligand Interactions, Improves Docking, and Allows Accurate Binding Free Energy Predictions

Juan Pablo Arcon,[†] Lucas A. Defelipe,[†] Carlos P. Modenutti,[†] Elias D. López,[†] Daniel Alvarez-Garcia,[‡] Xavier Barril,^{§,||} Adrián G. Turjanski,^{*,†} and Marcelo A. Martí^{*,†}

[†]Departamento de Química Biológica e IQUIBICEN-CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón 2, C1428EHA Ciudad de Buenos Aires, Argentina

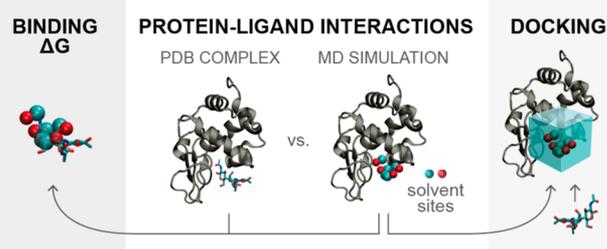
[‡]Discngine, 33 rue du Faubourg Saint-Antoine, 75011 Paris, France

[§]Institut de Biomedicina de la Universitat de Barcelona (IBUB) and Facultat de Farmàcia, Universitat de Barcelona, Av. Joan XXIII s/n, 08028 Barcelona, Spain

^{||}Catalan Institution for Research and Advanced Studies (ICREA), Passeig Lluís Companys 23, 08010 Barcelona, Spain

S Supporting Information

ABSTRACT: One of the most important biological processes at the molecular level is the formation of protein–ligand complexes. Therefore, determining their structure and underlying key interactions is of paramount relevance and has direct applications in drug development. Because of its low cost relative to its experimental sibling, molecular dynamics (MD) simulations in the presence of different solvent probes mimicking specific types of interactions have been increasingly used to analyze protein binding sites and reveal protein–ligand interaction hot spots. However, a systematic comparison of different probes and their real predictive power from a quantitative and thermodynamic point of view is still missing. In the present work, we have performed MD simulations of 18 different proteins in pure water as well as water mixtures of ethanol, acetamide, acetonitrile and methylammonium acetate, leading to a total of 5.4 μ s simulation time. For each system, we determined the corresponding solvent sites, defined as space regions adjacent to the protein surface where the probability of finding a probe atom is higher than that in the bulk solvent. Finally, we compared the identified solvent sites with 121 different protein–ligand complexes and used them to perform molecular docking and ligand binding free energy estimates. Our results show that combining solely water and ethanol sites allows sampling over 70% of all possible protein–ligand interactions, especially those that coincide with ligand-based pharmacophoric points. Most important, we also show how the solvent sites can be used to significantly improve ligand docking in terms of both accuracy and precision, and that accurate predictions of ligand binding free energies, along with relative ranking of ligand affinity, can be performed.



I INTRODUCTION

One of the most important biological processes at the molecular level is the formation of protein–ligand complexes. Proteins need to recognize, bind and discriminate properly among a universe of possible ligands in order to perform their function, and thus determining the structure and underlying key interactions of particular protein–ligand complexes is of paramount relevance.^{1,2} This knowledge is also very important for the rational design of new and more effective drugs,^{3–6} therefore several experimental and *in silico* tools have been developed in the last decades to determine and analyze them.^{7,8} Together with X-ray crystallography (and NMR to a lesser extent) of protein–ligand complexes, small molecular fragments and/or water miscible solvents can be used to identify specific types (e.g., hydrophilic, hydrophobic, charged) of potential protein–ligand interaction sites.^{9–12} For example, solvent

mapping, as used in the multiple solvent crystal structures (MSCS) technique,^{9,10} consists in solving crystal structures of proteins in the presence of several organic cosolvents, enforcing the idea that small molecules binding to specific sites on the protein reveal key interaction sites. Therefore, optimal ligands should be built displaying a chemical structure that is able to perform these same interactions. Although these methodologies successfully contributed to the identification and characterization of several protein pockets,^{13,14} they are demanding and require significant resources, thereby limiting their applicability to a wider range of cases.

Molecular dynamics (MD) simulations in the presence of different solvents with functional groups that probe specific type

Received: November 3, 2016

Published: March 20, 2017

Table 1. Protein–Ligand Complexes Dataset

protein	complex pdb codes ^a	pocket descriptors ^b	DUD-E ³³ target?
acetylcholinesterase	1E66, 1EVE , 1GPN, 1H22, 1VOT, 1W4L, 1W76, 1ZGC, 2CKM, 2V96	druggability, 0.75; volume, 1160; hydrophobicity, 26.3	yes
adenosylmethionine-8-amino-7-oxononanoate aminotransferase	3LV2 , 3TFU	druggability, 0.77; ^c volume, 3768; ^c hydrophobicity, 27.8 ^c	no
β -lactamase	1IEL, 1FSY, 1LL5, 1LL9, 1MY8, 1O07, 1XGJ , 2HDU, 2R9X, 4OLG	druggability, 0.42; volume, 2002; hydrophobicity, 16.5	yes
cGMP phosphodiesterase 5a	1RKP, 1T9S, 1XOZ, 1XP0 , 2CHM, 2H44, 3HDZ, 3SIE, 3TGE, 3TSE	druggability, 0.65; volume, 1195; hydrophobicity, 41.5	yes
coagulation factor XA	1F0R , 1KSN, 1LPG, 2J34, 2J94, 2UWP, 2VH6, 2Y7X	druggability, 0.51; ^c volume, 3237; ^c hydrophobicity, 15.9 ^c	yes
cyclin-dependent protein kinase 2	1CKP , 1PXN, 2BTS, 2DS1, 2FVD, 2R3P, 2XMY, 3LFS, 4ACM	druggability, 0.66; volume, 1773; hydrophobicity, 37.0	yes
cytochrome P450 125	2XC3 , 3IW1	druggability, 0.82; volume, 2192; hydrophobicity, 33.3	no
dihydrofolate reductase	1DIS, 3DFR , 3M08, 3SA2, 3SRQ, 3TQB, 4ELE, 4FGG, 4LAE	druggability, 0.69; volume, 1183; hydrophobicity, 37.2	yes
DNA gyrase B	1AJ6 , 1KZN, 4DUH	druggability, 0.71; volume, 872; hydrophobicity, 22.9	no
estrogen receptor α	1GWR, 1QKM, 1YIN, 2AYR, 2IOG, 2IOK, 2QE4, 2R6Y, 3ERD, 3ERT	druggability, 0.95; volume, 1480; hydrophobicity, 51.6	yes
fibroblast growth factor receptor 1	1AGW , 1FGI, 2FGL, 3C4F, 3JS2, 3RHX, 3TT0, 4F64, 4F65, 4NK9	druggability, 0.70; volume, 1045; hydrophobicity, 21.4	yes
glucocorticoid receptor	1M2Z , 1NHZ, 3BQD, 3E7C, 3K22, 3K23, 4CSJ, 4LSJ, 4P6W	druggability, 0.82; volume, 785; hydrophobicity, 45.8	yes
glycogen phosphorylase b	1P4G , 3MRT, 3NP9, 3T3G	druggability, 0.02; volume, 447; hydrophobicity, 13.7	yes
hen egg white lysozyme	1LJN, 1LZB	druggability, 0.65; volume, 539; hydrophobicity, 28.8	no
integrin α L	1RD4 , 1XDG, 1XUO, 2O7N, 3BQM, 3BQN, 3E2M, 3M6F	druggability, 0.92; volume, 880; hydrophobicity, 44.7	no
map kinase p38 α	1A9U , 2YIX, 2ZAZ, 2ZB0, 3D7Z, 3IWS	druggability, 0.69; ^c volume, 3613; ^c hydrophobicity, 20.4 ^c	yes
peroxisome proliferator activated receptor γ	1FM6 , 1FM9, 1K74	druggability, 0.84; volume, 1955; hydrophobicity, 35.1	yes
trypsin	1K1J, 1K1M , 1OYQ, 1QBO, 1V2R, 1XUJ	druggability, 0.06; volume, 615; hydrophobicity, 16.2	yes

^aThe reference structure used in MD simulations is highlighted in bold. ^bObtained with Fpocket.³⁴ ^cThe default Fpocket parameters were modified in order to get a single pocket for the binding site, thus obtaining rather big cavities.

of interactions are also able to reveal protein–ligand interaction hot spots,^{15–18} in a sort of *in silico* version of the MSCS approach. However, these simulations are much faster and cheaper than their experimental sibling, and thus of wider applicability. To have a high concentration of probe molecules and thus maximize their capacity of revealing protein–ligand interaction sites, small water miscible molecules, like ethanol or acetonitrile, are commonly used. The apoprotein of interest is immersed in these solvent mixtures, and plain MD simulations are performed to allow the probe molecules to interact properly with the protein. As shown by us and others^{19–22} for pure water, due to the presence of specific protein–solvent interactions, probe molecules are not placed randomly on the macromolecule surface, but instead tend to occupy specific positions and orientations thereby revealing relevant interaction sites. This is particularly evident in regions such as protein active sites and ligand binding regions.^{23–26} Moreover, these probe revealed “solvent sites” can be characterized structurally and thermodynamically, using for example the inhomogeneous fluid solvation theory,²⁷ a knowledge that can be later used for the design of new, more active, ligands.²⁸

Given its potential, several works have used solvent mixture MD simulations to study protein ligand complexes in a variety of proteins using a bunch of different probes.^{29,30} However, to our knowledge, most of them focused on a small set of proteins or

solvents and a systematic comparison of different probes and their real predictive power is still lacking. Also, most studies looked at the relation between protein–ligand complexes and the sites revealed by the probes only qualitatively, excluding, except for some particular cases,^{31,32} a statistical or thermodynamic type of analysis. Therefore, to analyze further the potential of this approach to reveal protein–ligand interaction hot spots, in the present work we have performed MD simulations of 18 different proteins in pure water as well as water mixtures of ethanol, acetamide, acetonitrile and methylammonium acetate, leading to a total of 5.4 μ s simulation time. We analyzed the resulting protein–solvent interactions sites in relation to 121 different protein–ligand complex structures and showed how they can be used to identify protein hot spots, improve molecular docking and make accurate ligand binding energy predictions.

The proteins were selected in order to cover a wide variety of binding/active sites, as well as several well characterized protein targets from the DUD-E data set.³³ The solvent probes were chosen in order to cover the most important types of molecular interactions usually established between proteins and their ligands. Briefly, water has been shown to be a very good probe for polar interactions, especially those established between ligand hydroxyl/carbonyl groups and the protein. Ethanol was chosen because it is highly miscible in water and can probe both polar and nonpolar interactions, the later by means of the methyl end.

Table 2. Molecular Interaction Assignment for Solvent Sites

solvent probe\type of solvent site	hydrogen bond donor (HBD)	hydrogen bond acceptor (HBA)	aliphatic/aromatic hydrophobic (HS)	ionic + (PI)	ionic - (NI)
water	x	x			
ethanol (-OH)	x	x			
acetamide (-NH ₂)	x				
acetamide (=O)		x			
ethanol (-CH ₃)			x		
acetonitrile (-CH ₃)			x		
methylammonium				x	
acetate					x

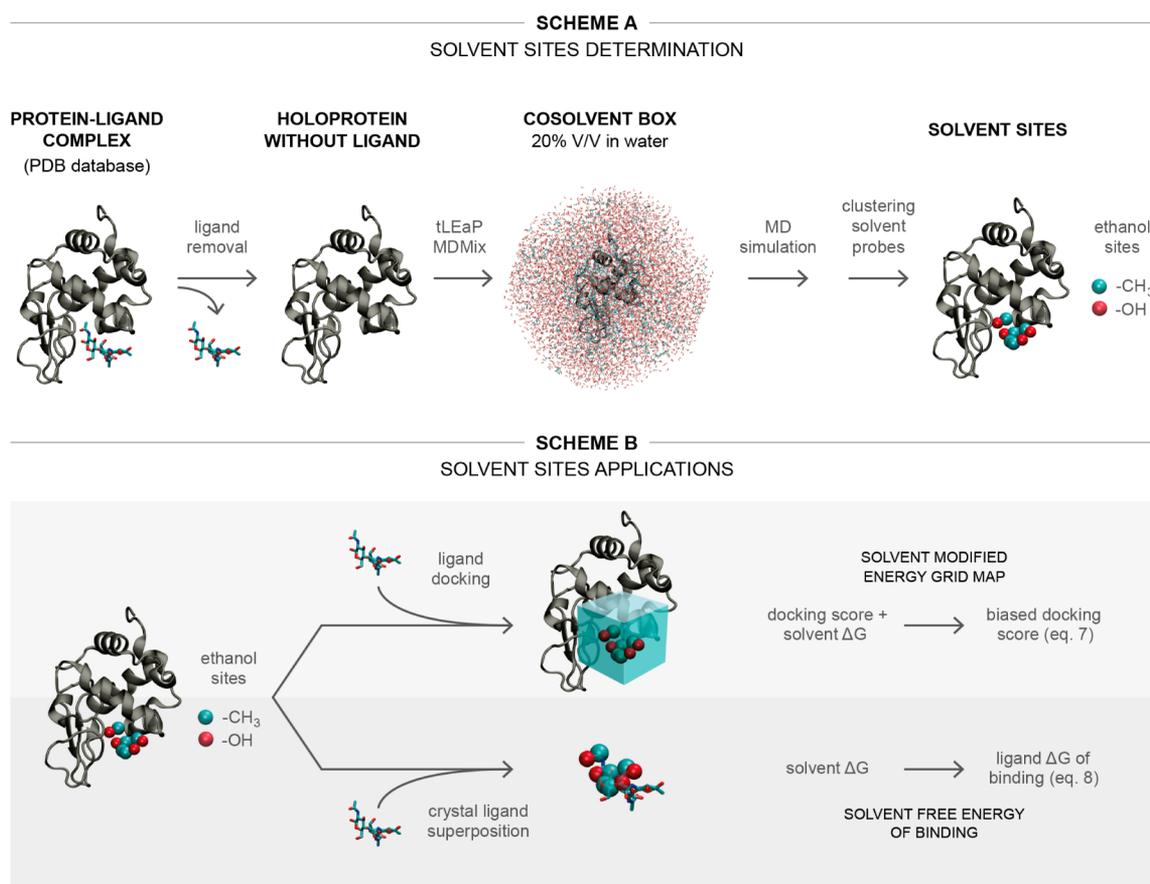


Figure 1. (Scheme A) Solvent site identification workflow example using ethyl alcohol as cosolvent. After removing the ligand and all non-protein atoms from the complex PDB structure, triplicates of explicit water–ethanol solvent MD simulations were performed (20 ns each). Ethanol hydrophilic and hydrophobic sites were obtained by post processing MD trajectories with the clustering algorithm described in the text. (Scheme B) Two possible applications for the detected solvent sites. Upper side: docking ligands to protein receptor using a modified scoring function that combine the conventional AutoDock4 scoring terms with the solvent sites free energy (see eq 7 and Figure S1). Lower side: Estimation of protein–ligand free energy of binding by superposition of the cocrystallized ligand with the solvent sites and summing the ΔG from the replaced solvent sites (see eq 8 and Figure S2).

Acetamide was selected to test its capacity for mimicking amide groups of ligands, whereas methylammonium acetate was used to probe ionic interactions. Finally, we also tested acetonitrile as an alternative hydrophobic probe (methyl end).

Our results provide a comprehensive view of how MD simulations in mixed solvents are able to reproduce protein–ligand interaction hot spots, showing which are the best solvent probes and what is their predictive power. We also show that combining just two types of sites (hydrophilic and hydrophobic) derived from water/ethanol simulations, significant improvements in docking calculations and accurate estimations of ligand binding free energies can be achieved.

COMPUTATIONAL METHODS

Protein Data Set. The protein–ligand data set used in the present work was defined in order to achieve a high diversity in size, shape and polarity of the binding sites. It includes selected proteins from the DUD-E database,³³ which are useful for benchmarking virtual screening procedures. For each protein, several complex structures (with up to 10 different ligands) were analyzed. Table 1 lists the protein complexes structures with their PDB ids and binding site characteristics.

Starting Structures and Setup for Molecular Dynamics Simulations. Starting structures for all studied protein–ligand complexes were downloaded from the Protein Data Bank (PDB)

database³⁵ (www.rcsb.org). For each protein, one of the complex structures was selected to perform the molecular dynamics (MD) simulations (the ones highlighted in bold in Table 1). All nonstructural ions, solvent and ligand molecules were removed *in silico* in order to define the starting protein structure. Missing side chains as well as hydrogen atoms were subsequently added using the LEaP module from Amber 12 package.³⁶ Standard protonation state at physiological pH was assigned to all ionizable residues unless otherwise explicitly stated. The resulting structures were immersed in a truncated octahedral box of solvation extending at least 10 Å from any protein atom using MDMix.²⁹ Besides pure water, the following water/solvent mixtures were tested: ethyl alcohol, acetamide, methylammonium acetate and acetonitrile using in all cases 20% v/v mixtures. TIP3P model was used for all water molecules. Force field parameters used for the other solvents can be found in Table S1 from the Supporting Information. For all protein residues, Amber ff99SB force field was used.³⁷

MD Equilibration and Production Procedure. All solvated systems were subjected first to a geometry optimization procedure to adjust solvent orientation, eliminate local clashes and stereochemical inaccuracies. The following equilibration protocol consisted in 0.8 ns of constant volume MD, where temperature was slowly raised from 100 to 300 K, after which 1 ns of constant pressure and temperature MD was performed (1 bar, 300 K) to allow the system to reach proper density. In all MD simulations, temperature and pressure control was achieved using the Langevin thermostat and Berendsen barostat, respectively. Systems were simulated using periodic boundary conditions and Ewald sums (grid spacing of 1 Å) for treating long-range electrostatic interactions with a 9 Å cutoff for direct interactions. The SHAKE algorithm was used to keep bonds involving H atoms at their equilibrium length, allowing the employment of a 2 fs time step for the integration of Newton's equations. Simulations were performed with the PMEMD implementation of SANDER from the Amber 12 program package.^{36,38,39} Finally, the system was subjected to 20 ns MD production run. For each protein, three independent MD simulations starting from the minimization (geometry optimization) step were done for each solvent mixture and pure water to enhance sampling. No restraints were applied to protein atoms. To verify protein stability during the MD simulations, the backbone rmsd was monitored over time for each simulation, using as reference the corresponding crystal structure. Results presented in Supporting Information (Table S2) show that average rmsd is below 2 Å in all cases, having the same magnitude in mixed solvents simulations when compared with those of pure water, confirming that proteins remain properly folded and in a stable conformation.

Determination of Solvent Sites. The method used to detect protein–solvent interaction sites is derived and inspired in previous works from our group^{19,40} and others⁴¹ that developed ways for obtaining the so-called water sites (WS). First, each type of solvent site is defined as being probed by a functional group or atom of the solvent molecules. Thus, for example, water molecules define polar sites that can be further assigned as hydrogen bond donor or acceptors, ethanol methyl group defines hydrophobic, and also aromatic, sites, whereas charged solvents define ionic interactions. Table 2 lists which solvent probes are used to define each type of sites. In the present work, in addition to the traditional WS we will define ethanol –OH based hydrophilic sites with both hydrogen bond donor and acceptor capabilities (HBD/HBA), ethanol –CH₃ based hydrophobic

sites (HS), acetamide –NH₂ based hydrogen bond donor (HBD) sites, acetamide =O based hydrogen bond acceptor (HBA) sites, methylammonium positive ionic (PI) sites and acetate negative ionic sites (NI). Assignment of both aliphatic and aromatic interactions to aliphatic probes was sustained by previous findings.²⁹

After we defined which probe atoms define each type of site, the solvent sites were determined from the corresponding explicit solvent mixture MD production simulations using a clustering algorithm.¹⁹ The solvent site identification workflow is schematically shown in Figure 1, Scheme A, for ethyl alcohol as an example. Solvent sites are defined as space regions adjacent to the protein surface where the probability of finding a probe-atom is higher than that in the bulk solvent. The strategy starts from a set of snapshots derived from the whole MD simulation. A structural alignment of all the protein conformations from the different snapshots is performed, based on a selected group of protein residues, usually those forming the binding/active site. This group of residues also defines the protein surface region near which the solvent sites will be determined. To identify the solvent site, a clustering algorithm is used. Only probe molecules interacting with the surface of interest along the whole MD simulation are considered. The positions of the probe atom of interest (e.g., C1 from ethanol for hydrophobic sites, see Table S1 of the SI) are the ones considered for the clustering. The algorithm starts with the first probe atom from the first snapshot and finds all other probe atoms that are closer than a threshold distance in all subsequent snapshots. This procedure is performed iteratively and a new round of search is performed for each of these newly found probe atoms until no more of them are found within the threshold in any snapshot. Then, all probe atoms found are clustered together. If the number of probe atoms belonging to the cluster is larger than that expected for a piece of bulk solvent bearing the same volume, the cluster defines a solvent site. Each snapshot can contribute with no more than one probe atom to the cluster. All the probe atoms from the cluster (whether or not they are stored as a solvent site) are then removed from further evaluation and the process starts again for the remaining probe atoms. Once there are no more probe atoms left, the process is completed and all solvent sites have been identified. The site center coordinates correspond to the center of mass of all probe atoms that are found inside the site along the MD simulation. Further applications for the determined solvent sites are shown in Figure 1, Scheme B, and explained below.

Calculation of Statistical Parameters Associated with Each Probe Predictive Power. The following parameters were defined to establish the capacity of each type of solvent site for predicting protein–ligand interactions:

True positives (TP) correspond to solvent sites that match a corresponding type of protein–ligand interaction (according to Table 2).

False positives (FP) correspond to solvent sites that fail to match a corresponding type of protein–ligand interaction.

True negatives (TN) are those protein–ligand interactions that are not reproduced by a non corresponding probe.

False negatives (FN) are those protein–ligand interactions that are not reproduced by a corresponding probe.

We define that a solvent site and a protein–ligand interaction coincide whenever the distance between the solvent site center and the ligand atom responsible for the interaction is less than 1.5 Å apart.

For the pharmacophore centered analysis, solvent sites were limited to only those matching (or not) a pharmacophoric

protein–ligand interaction of any type. Thus, FP are redefined as those sites that match a pharmacophoric protein–ligand interaction of inadequate type.

Sensitivity or true positive rate (TPR) is defined as the ratio between true positives and the total number of real positives:

$$\text{sensitivity} = \text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (1)$$

Precision or positive predictive value (PPV) is defined as the ratio between true positives and the total number of predicted positives:

$$\text{precision} = \text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

Specificity or true negative rate (TNR) is defined as the ratio between true negatives and the total number of real negatives:

$$\text{specificity} = \text{TNR} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (3)$$

Accuracy is defined as the ratio between correct results and the total number of results:

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (4)$$

Calculation of Solvent Site Properties. Once identified, for all the solvent sites we compute the following parameters:

- (i) Probe finding probability, PFP, corresponding to the probability of finding a probe atom in the region defined by the solvent site, using the arbitrary volume value of 1 Å³ and normalized with respect to that of the bulk solvent.

The PFP is directly related to the probability density function and is obtained from a radial 3D histogram of the occurrence of probe atoms in each space region using the described clustering algorithm. The PFP is calculated by evaluating this occurrence between the center of the solvent site and a radius value of approximately 0.6 Å (volume of 1 Å³) and dividing by the occurrence obtained for the probe in the bulk solvent (eq 5). Thus, the probe finding probability is not simply a probability, but corresponds to the relative probability of finding a probe atom in the solvent site with respect to that of the bulk solvent. Therefore, it can be interpreted as a molecular approximation to an equilibrium constant between the solvent site and the bulk solvent, as shown in eq 5.

$$\text{PFP} = \frac{N_{\text{solvent_site}}}{N_{\text{bulk}}} = K \quad (5)$$

where $N_{\text{solvent_site}}$ corresponds to the number of frames from the MD simulation in which a probe atom is found inside the solvent site, and N_{bulk} corresponds to the number of frames in which a probe atom is found in the bulk, inside a sphere whose volume equals that of the solvent site (1 Å³). By using the inverse Boltzmann relationship, we obtain eq 6 and the estimated ΔG thus corresponds to the difference in free energy of a probe atom being inside the solvent site with respect to that of being in the bulk solvent.

$$\Delta G = -RT \ln(\text{PFP}) \quad (6)$$

- (ii) R_{90} , corresponding to the radius the solvent site should have in order to contain 90% of the probe atoms that

define the site. This value is a measure of the solvent site dispersion.

- (iii) The nearest distance of the solvent site to any protein atom, NPD.
- (iv) The site solvent accessible surface area (siSASA).

Molecular Docking. To analyze the improvement of docking results using the information derived from the solvent sites we compared, as in our previous works,^{26,42} the performance of the Autodock4 program using its usual parameters, which we call conventional autodock docking method (CADM), with that of a modified protocol that includes information derived from the solvent sites, the solvent site biased docking method (SSBDM).

Conventional Autodock Docking Method (CADM). The CADM was performed with the AutoDock 4.2 program⁴³ using the default parameters. Briefly, based solely on the protein structure, the energy maps were computed as usual. The grid size and position was chosen considering the reference ligand coordinates: it was centered in the ligand geometric center and extended to reach a cubic box with an edge distance equal to 2 times the maximum distance between any two ligand atoms. The spacing between grid points was set at 0.375 Å. 100 docking runs were performed and the resulting poses were clustered according to the ligand heavy atom rmsd using a cutoff of 2 Å, thus defining a population for each cluster. The genetic algorithm was kept at their default values.

Solvent Site Biased Docking Method (SSBDM). The SSBDM is based on the protocol developed and thoroughly tested in our previous works for water sites and carbohydrate type of ligands^{19,26,42} and extended here to include hydrophobic sites derived from ethanol/water MD simulations and any type of ligand. To take advantage of the fact that ligand polar groups tend to occupy the same positions than the MD derived water or ethanol hydrophilic sites and that ethanol hydrophobic sites tend to probe hydrophobic interactions, especially those involving aromatic rings, we modified the corresponding AutoDock4 grids. For this sake, for each ligand heavy atom that could be involved in hydrogen bond interactions (OA/NA/N atom types), an additional energy term was added to the original autodock function, according to eq 7.

$$\Delta G_{\text{SSBDM}} = \Delta G_{\text{CADM}} - RT \sum_{i=1}^N [\ln(\text{PFP}_i)] \times e^{\sqrt{(x-x_i)^2 + (y-y_i)^2 + (z-z_i)^2} / R_{90,i}} \quad (7)$$

where ΔG_{SSBDM} corresponds to the resulting modified scoring function, ΔG_{CADM} is the original AutoDock4 scoring function, R is the gas constant, T is the temperature (298 K), the sum extends over the total number N of water or ethanol–OH sites interacting with the protein in the binding site, PFP is the probe finding probability of the solvent site, (x, y, z) are the grid points coordinates, (x_i, y_i, z_i) are the solvent site coordinates and $R_{90,i}$ is the solvent site dispersion parameter. Thus, for each identified water or ethanol–OH site an energy well is created in the OA, NA and N grid maps. The well deepness (i.e., energy reward) increases with the solvent site PFP and its width (i.e., extension through space away from the solvent site center) increases with the solvent site R_{90} . Figure S1 schematically shows how the modifications to the grid maps are applied.

To include the hydrophobic bias, a slightly different scheme is applied. We first created for each ligand aromatic ring a new dummy atom located in the ring center. We then created a new

grid map for the dummy atom whose energy is calculated by eq 7, with $\Delta G_{\text{CADM}} = 0$ and using ethanol determined hydrophobic sites. We tried also other possibilities that involved using directly the aromatic carbon map, but the performance was less satisfactory. In summary, the biased docking potential will favor those ligand poses where hydrogen bond donor/acceptor atoms of the ligand replace water or ethanol–OH sites and where aromatic rings replace ethanol hydrophobic sites, with a strength that is directly related to the solvent site PFP.

The SSBDM was then employed in the same manner as the CADM but using the modified grid maps. For strict comparison purposes, all other docking parameters were the same as those used in the CADM. All scripts to analyze crystal structures, determine and characterize the solvent sites and modify the Autodock4 grid maps are available upon request.

Docking Experiments and Analysis of Results. We performed three different molecular docking experiments for two representative proteins from our set: FGF receptor 1 and AmpC β -lactamase. The first experiment consisted in a redocking experiment in which a cocrystallized ligand (with randomized position and torsions) was docked against its own crystal receptor. The second strategy used different receptor structures taken from MD simulations of the protein in explicit water. Finally, we also performed cross-docking experiments, in which a representative cocrystallized receptor was chosen and a set of different ligands were docked against it. In particular, for the FGF receptor we docked all the ligands from the original data set (Table 1), whereas for β -lactamase we only kept noncovalent ligands and added all ligands from the PDB with reported K_i below 100 μM to enrich the final set (Table S5). All three experiments were performed independently with the conventional and biased docking methods. To compare both methods, we considered two main issues. First, the equivalence between the reference ligand structure and the ligand pose predicted by the docking method, computed by the ligand heavy atoms rmsd. Second, the method capacity to discriminate the correct ligand pose (i.e., that with rmsd < 2 Å when compared with the reference structure) from wrong predictions, by using the predicted binding free energy score (ΔG) and the population, which is the number of individual docking results that predicted the same ligand pose in the whole docking experiment (100 runs). A successful method should find the correct pose with low binding energy and high population, while also significantly differing in these both parameters from the other poses. As shown in the Results section, this can be easily analyzed by plotting population vs binding free energy score for all the predicted complexes in a given docking calculation. When several ligands are docked against the same receptor, the method discriminating capability for each ligand is tested by calculating the differences in the predicted binding free energy score ($\Delta\Delta G$) and in the population ($\Delta\text{population}$) between the correct ligand pose, that with rmsd < 2 Å, and the best ranked, that with lowest energy, of the remaining poses. A negative $\Delta\Delta G$ value implies that the correct complex is found and has better binding energy score than any other predicted pose, whereas a positive $\Delta\Delta G$ means that the correct pose is less tightly bound than other predicted complexes or is not found at all. In the same way, a successful method should obtain positive $\Delta\text{population}$ values, meaning that the correct complex has the most populated cluster. Thus, results located in the upper left quadrant of the plot correspond to successful predictions where the correct complex has the lowest energy and highest population, whereas results in the lower right quadrant represent wrongly predicted poses.

Solvent Site Contribution to Ligand Binding Free Energy. Binding free energy calculations were performed on human coagulation factor Xa (Uniprot ID P00742) and *Escherichia coli* AmpC β -lactamase (Uniprot ID P00811), due to the large number of ligands with available high resolution complex X-ray structures and thermodynamic binding data (K_i). The complete set of PDB IDs that were analyzed, 75 for factor Xa and 28 for β -lactamase are reported in Tables S6 and S7 in the SI. The predicted binding free energy was obtained using the above-mentioned X-ray protein–ligand complexes and the MD derived solvent sites as described in eq 8, which is similar to that used for the modified docking protocol but without the weight related to the R_{90} (because it did not provide any improvement). The use of X-ray structures for free energy estimations avoids introducing possible bias and permits the evaluation of the solvent site potential for this task, independently of their performance for docking, which was evaluated in the previous section.

$$\Delta G_{\text{pred}} = -RT \sum_{i=1}^N \ln \text{PFP}_i \quad (8)$$

where PFP_i is the probe finding probability of the i -th solvent site replaced by a ligand group capable of establishing the same type of interactions as the solvent site and N is the total number of replaced solvent sites. A hydrophilic solvent site is considered replaced when its position coincide with a nitrogen or oxygen atom from hydrogen bond donor/acceptor groups of the ligand, whereas hydrophobic solvent sites were considered replaced when overlaying aliphatic or aromatic carbon atoms from the ligand. The distance threshold was set at 1.5 Å. All complex structures were aligned according to the active site with the reference structure used to obtain the solvent sites. A schematic representation of the calculations done with eq 8 is shown in Figure S2.

To analyze the performance of our free energy estimates, we built predicted vs experimental free energy plots, computed the linear regressions and calculated the resulting determination coefficient (R^2) and root-mean-square error (rmse). For the ligand group based analysis, a Tanimoto index based cutoff of 0.7 was used to cluster similar ligands either in pairs for computing $\Delta\Delta G_{\text{pred}}$ between them or in groups of more than three members for comparing their absolute ΔG_{pred} values. The similarity groups/sets are presented in Tables S6 and S7 of the Supporting Information.

RESULTS

Solvent Interaction Sites Predict Protein–Ligand Interactions. We begin our analysis by looking at how many protein–ligand interaction sites, of each type, are sampled by the solvent probes inside the ligand binding site of all 18 analyzed proteins. For this sake we first analyzed, for each studied protein, a set of selected protein–ligand structures with wide range of molecular interactions and K_i values, which together add up to 121 different complexes (Table 1 in Computational Methods), and determined the corresponding protein–ligand interaction sites, classifying them by type as shown in Table 3. As expected, most observed interactions correspond to hydrogen bonds, followed by hydrophobic (aliphatic and aromatic) interactions, whereas only a minor number of ionic interactions are present in the current data set.

In parallel, we performed, for each protein (in absence of ligand), a series of three independent 20 ns long MD simulations in pure water as well as the following 20% v/v water mixtures:

Table 3. Number of Protein–Ligand Interaction Sites Identified in all 121 Protein–Ligand Complexes Classified by Their Interaction Type

type of interaction site ^a	# total interaction sites	# pharmacophoric interaction sites ^b
hydrogen bond donor	47	18
hydrogen bond acceptor	71	22
aliphatic hydrophobic	51	19
aromatic hydrophobic	28	9
positive ionic	5	1
negative ionic	5	0

^aThe type of protein–ligand interaction site is described from the ligand point of view. ^bPharmacophoric interaction sites are defined as those protein–ligand interaction sites present in half or more of the analyzed complexes for each protein.

ethanol, acetamide, methylammonium acetate and acetonitrile. Altogether, we performed over 5.4 μ s of production MD simulations. Subsequently, for each case, and with the same approach used to define water sites (WS) in our previous works^{40,42} (see [Computational Methods](#) for details), the following solvent sites as defined by the corresponding probe atoms were determined: hydrogen bond donor (HBD) sites, hydrogen bond acceptor (HBA) sites, hydrophobic sites (HS), positive ionic (PI) sites and negative ionic (NI) sites. Analysis of each probe performance for predicting both general and pharmacophoric protein–ligand interaction sites, classified by their type, is shown in [Table 4](#), whereas particular examples illustrating the relation between probed identified sites and ligand structure are shown in [Figure 2](#).

To perform a comparative statistical analysis (see [Computational Methods](#) for details), we defined true positives (TP) as those solvent sites that reproduce protein–ligand interactions of their own type (according to [Table 2](#) in [Computational Methods](#)), whereas false positives (FP) correspond to solvent sites that fail to do so. In this context, the precision (or positive predictive value, PPV) measures how many of the probe detected sites are able to predict adequately protein–ligand interactions. On the contrary, false negatives (FN) are those protein–ligand interaction sites that fail to be reproduced by a solvent probe bearing the same type of interaction. Thus, the sensitivity (or true positive rate, TPR) measures how many of the observed protein–ligand interactions are actually sampled by a corresponding probe. True negatives (TN) are protein–ligand interaction sites not reproduced by a mismatching probe (e.g., a protein–ligand hydrogen bond that is not detected by

acetonitrile $-\text{CH}_3$ probe). Thus, the specificity (or true negative rate, TNR) shows each probe capacity to avoid detecting sites corresponding to other types of interactions. Finally, the accuracy is a summary measure that relates the correct results (true positives and true negatives) with the total number of results (either true or false). When analyzing only pharmacophoric interaction sites (values in parentheses in [Table 4](#)), positives and negatives were redefined accordingly.

Looking at all the revealed protein–ligand interactions, data from [Table 4](#) shows that most probes display a reasonable accuracy, with most values close or above 0.5. Water remains somewhat behind the rest (accuracy = 0.39) mainly due to its increased number of false positives. Also, for most probes sensitivity is higher than precision because more sites are sampled by the probes than those corresponding to protein–ligand interactions. The best probes are able to detect about two out of every three protein–ligand interactions of their own type (sensitivity > 0.64), whereas about half of the detected solvent sites correspond to these protein–ligand interactions (precision > 0.48).

Comparative analysis of polar probes shows that water, through the identification of water sites (WS), which can be HBD, HBA or both, is significantly better at predicting hydrogen bond interactions than ethanol or acetamide (sensitivity of 0.64 vs 0.36 or less). Moreover, there is significant redundancy in the detection of WS with respect to these other probes, thus little predictive power is added by combining all of them. However, its lower specificity show that WS also significantly overlap with other type of protein–ligand interaction sites. Ethanol–OH, on the contrary, is a less sensitive probe, but its solvent sites rarely appear in regions where no apparent protein–ligand hydrogen bond interaction can be established, thus significantly enhancing the probe specificity. Moreover, its higher precision value indicates that ethanol–OH sites are enriched in those that actually reproduce hydrogen bond interactions from the ligand set. Finally, for acetamide, both the low sensitivity and precision diminish its predictive capacity. Thus, although water or ethanol–OH do not allow direct, *a priori*, differentiation of donor and acceptor sites, both outperform the acetamide probes.

Concerning nonpolar interactions, both ethanol– CH_3 and acetonitrile– CH_3 show similar and very good performance, being able to detect more than 70% of the protein–ligand hydrophobic interactions, while also achieving high precision and specificity. The results are similar if nonpolar interactions are divided in aliphatic or aromatic, although detailed analysis of which ligand functional groups coincide with the identified solvent hydro-

Table 4. Performance of Solvent Probes for the Prediction of Protein–Ligand Interactions

solvent probe	type of solvent site ^a	sensitivity	precision	specificity	accuracy
water	HBD/HBA	0.64 (0.73)	0.33 (0.64)	0.24 (0.43)	0.39 (0.60)
ethanol (–OH)	HBD/HBA	0.36 (0.48)	0.48 (0.63)	0.56 (0.61)	0.45 (0.53)
acetamide (–NH ₂)	HBD	0.34 (0.39)	0.18 (0.29)	0.62 (0.66)	0.56 (0.59)
acetamide (=O)	HBA	0.27 (0.36)	0.25 (0.38)	0.62 (0.72)	0.51 (0.60)
ethanol (–CH ₃)	HS	0.72 (0.86)	0.58 (0.96)	0.72 (0.98)	0.72 (0.93)
acetonitrile (–CH ₃)	HS	0.77 (0.89)	0.54 (0.83)	0.68 (0.88)	0.71 (0.88)
all ^b	HBD/HBA/HS	0.53 (0.64)	0.39 (0.64)	0.56 (0.73)	0.55 (0.69)
all cosolvents ^c	HBD/HBA/HS	0.49 (0.61)	0.42 (0.64)	0.65 (0.77)	0.59 (0.71)

^aHBD = hydrogen bond donor, HBA = hydrogen bond acceptor, HS = hydrophobic site. ^b“All” includes water, ethanol, acetamide and acetonitrile probes. ^c“All cosolvents” include ethanol, acetamide and acetonitrile probes (it excludes water). Values in parentheses are computed considering only protein–ligand pharmacophoric interaction sites (see text for details).

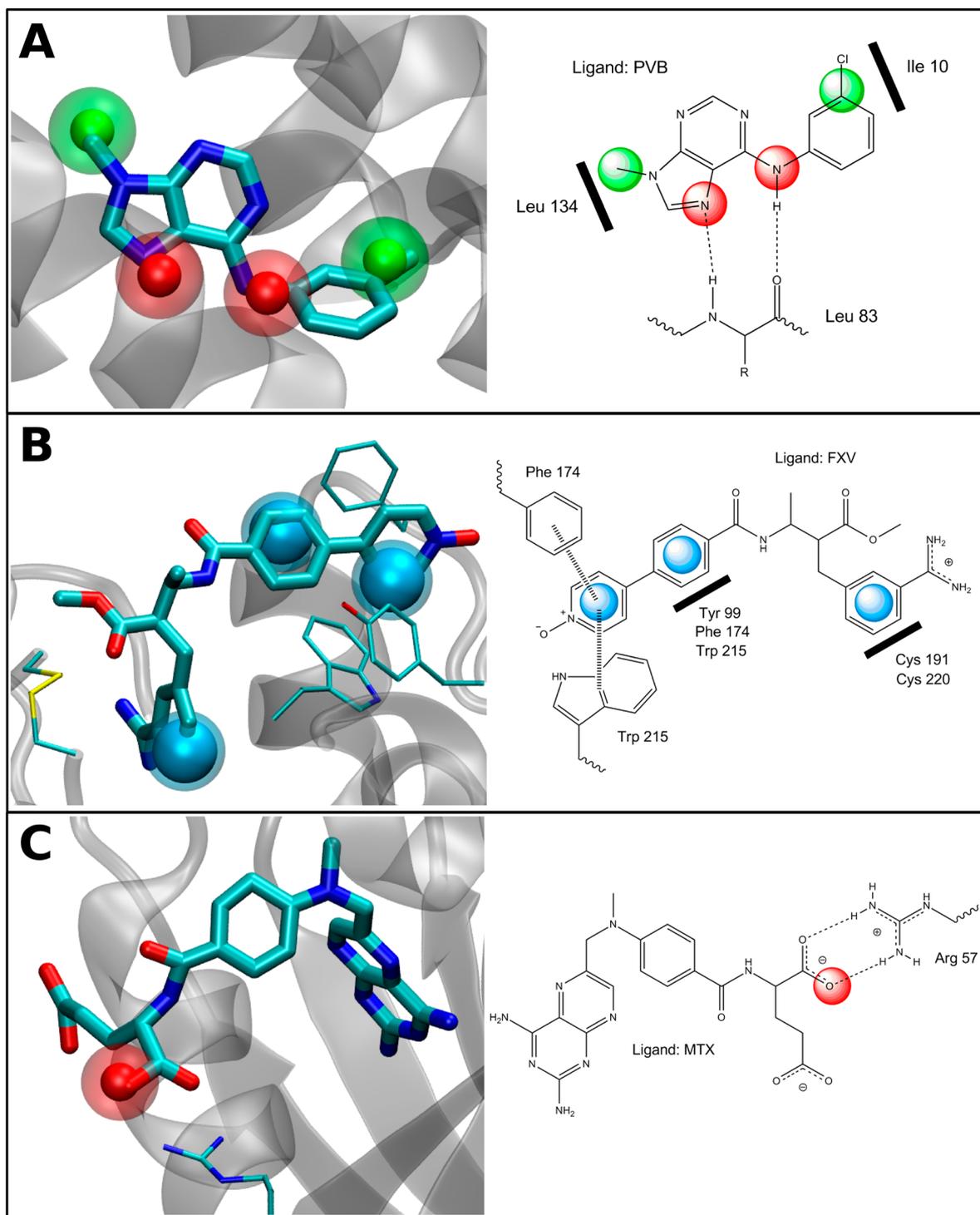


Figure 2. Examples of solvent predicted protein–ligand interaction sites. For each panel, the left side shows the predicted solvent sites superimposed to the protein–ligand structure, whereas the right side shows a simplified interaction scheme according to the detected solvent sites: hydrogen bonds are indicated as dashed lines, hydrophobic aliphatic interactions are shown as thick lines and aromatic π – π interactions are depicted as consecutive parallel bars. (A) Water sites for CDK2 superimposed on PVB ligand (PDB ID 1ckp). Red and green spheres correspond to true positive and false positive water sites, respectively. (B) True positive ethanol hydrophobic sites (light blue spheres) for coagulation factor Xa superimposed on FXV ligand (PDB ID 1ksn). (C) True positive negative ionic site from acetate (red sphere) in DHFR active site superimposed on MTX ligand (PDB ID 3dfr).

phobic sites shows that there is slight preference for aromatic rings vs aliphatic rings or side chains.

Finally, although the number of observed ionic protein–ligand interactions is small (Table 3) and no statistic parameters can be derived, about half of both positive and negative ionic interactions are revealed by the corresponding solvent sites

(Table S8 in the Supporting Information). It is worth mentioning that due to their hydrogen bond capabilities, ionic probes also reveal HBA or HBD sites, although they are highly redundant because all but one of these protein–ligand interactions (94%) were already detected by water.

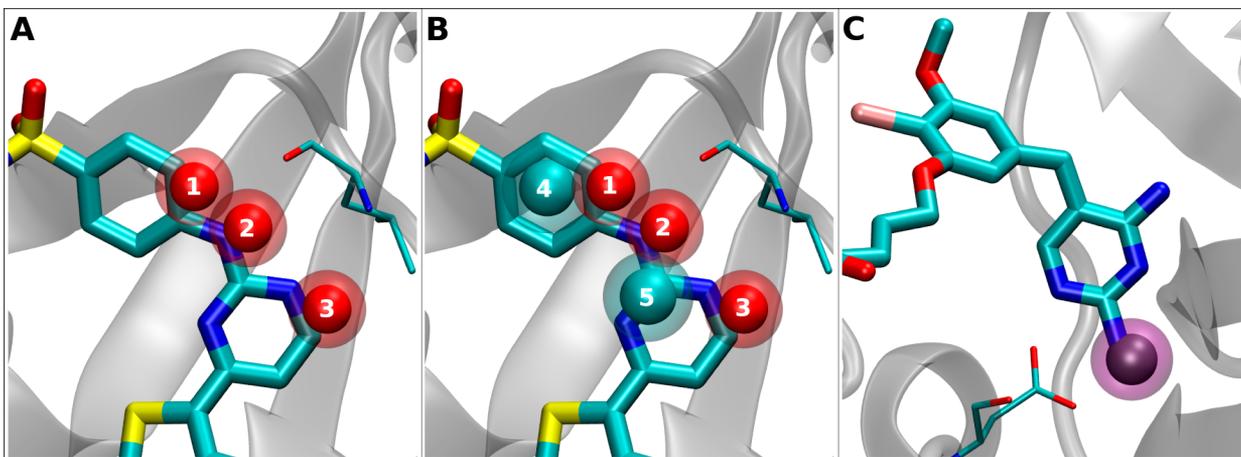


Figure 3. (A,B) CDK2 binding site (PDB ID 2xmy) highlighting the hydrogen bond pharmacophoric interactions between Leu83 backbone and the ligand (shown as sticks). Ethanol hydrophilic and hydrophobic sites, depicted as red and cyan spheres respectively, are superimposed to the ligand structure. (C) DHFR active site (PDB ID 1dis) highlighting the hydrogen bond pharmacophoric interaction between the ligand and Asp26 carboxylate (shown as sticks). Acetonitrile hydrophobic site depicted as a purple sphere is superimposed to the ligand structure.

Figure 2 illustrates the above-described results for particular cases. For example, the WS found in the active site of CDK2 (Figure 2A) replace polar interacting ligand groups such as the amine hydrogen bond donor and the purine acceptor N7, both interacting with Leu83 backbone (these WS are depicted in red). Also, as inferred from the low specificity value for water (Table 4), some WS such as those shown in green in Figure 2A seem to be false positives. Figure 2B shows the ethanol revealed hydrophobic sites and how they are able to sample adequately the position of aromatic rings in the Factor Xa–FXV complex. Finally, in Figure 2C we show how a negative ionic site is able to reproduce a ionic interaction between a ligand carboxylate group and an arginine side chain from DHFR active site.

We now focus not on all possible interactions, but only on those that are known to be critical for the protein–ligand binding. These interactions, which are present in half or more of the analyzed complexes for each protein, define thus the basic pharmacophoric framework for those ligands that bind the target tightly. Table 3 enumerates these pharmacophoric interactions classified by type for the present protein–ligand set. To analyze the probe performance in this context, we redefined positives and negatives only considering the match/mismatch between solvent sites and protein–ligand pharmacophoric sites. The results (presented in Table 4 as values in parentheses) show, as could be expected, that all probes increase their overall performance.

For hydrogen donor and acceptor interactions, water still achieves the highest sensitivity, being able to detect over 70% of the protein–ligand sites. Also important is the fact that sensitivity is increased for all the probes, being ethanol–OH the one achieving the highest increase (12%). For nonpolar interactions, ethanol and acetonitrile show again similar outstanding performance, being able to sample ca. 90% of the key protein–ligand interactions. This means that all the probes are better reproducing pharmacophoric protein–ligand interactions and thus generating a relative decrease in false negatives when compared with general protein–ligand interaction points. In addition, it is clear that when analyzing pharmacophoric sites, a high precision is achieved by all probes except acetamide, showing that more than 60% of the polar solvent sites and more than 80% of the apolar solvent sites reproduce an interaction of adequate type. Finally, cosolvent probes also reveal as highly

specific. These facts are important for their applications in prospective analysis as will be shown in later sections.

When Do Probes Fail? As already mentioned, it is clear that there is a number of solvent sites that do not correspond to potential ligand interactions (as revealed by low/moderate precision values). To examine if there was any systematic behavior related to this observation, we analyzed in detail the corresponding false positives. The analysis showed three main causes. First, when the protein displays, for example, an hydrogen bond acceptor such as a carbonyl group, sometimes the immediate environment allows the presence of two polar solvent sites interacting from different orientations with the same carbonyl group. That is the case, for example, of CDK2 Leu83 backbone as shown by ethanol polar sites 1 and 2 depicted as red spheres in Figure 3A. However, it is important to note that the ligand structural framework is usually unable to accommodate both sites simultaneously, and thus only either one of them is replaced (site 2 in this particular case). Therefore, although the probes correctly predict these HBD interaction sites, only one of the sites is computed as TP, the other being a FP (site 1 in Figure 3A). An interesting point is that the ligand places an aromatic C with some H bond donor character over solvent site 1, although this type of interactions were not taken into consideration in our analysis. We conclude that, as general rule, when two solvent sites are too close and making the same interaction with the protein, only one of them will be functional in the ligand.

The second cause for wrong predictions is related to the presence of apolar sites in amphiphilic environments. As shown in Figure 3B for the same CDK2 binding site, the environment allows the formation of not only the above-mentioned hydrophilic sites but also hydrophobic sites (cyan spheres). Because the ligand pyrimidine ring is mainly establishing a hydrogen bond with the protein, as reproduced by the polar site 3, the hydrophobic site 5 is considered a false positive, although it actually samples the hydrophobic part of the ring. It is worth mentioning that this hydrophobic site (false positive) is established aided by probe coupling with the nearby hydrophilic sites 2 and 3 (true positives), as a consequence of using a dual hydrophilic–hydrophobic cosolvent molecule.

Finally, there are some false positives due to sampling artifacts. An acetonitrile hydrophobic site was found inside DHFR active site located in an evident polar region where all the analyzed

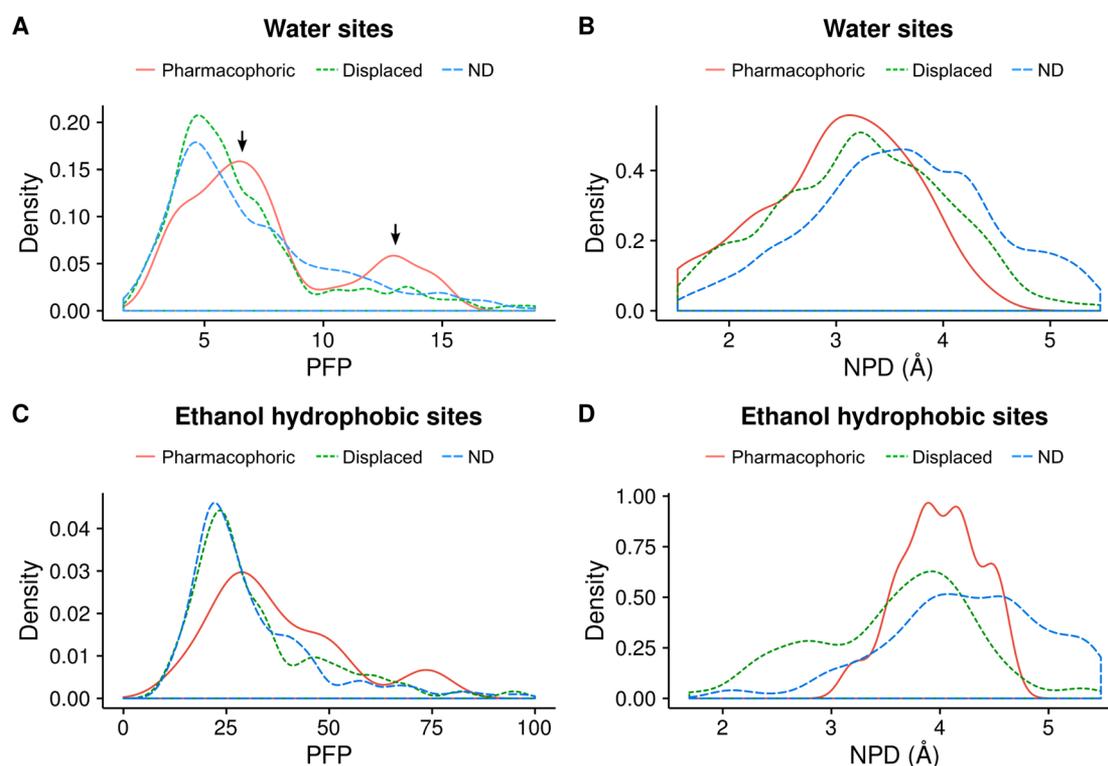


Figure 4. Probability density function for different probe types and solvent site properties. (A) Water sites probe finding probability (PFP). (B) Water sites nearest protein distance (NPD). (C) Ethanol hydrophobic sites PFP. (D) Ethanol hydrophobic sites NPD. Red full line, green dots and blue dashed line correspond to pharmacophoric, displaced and other (nondisplaced) solvent sites, respectively.

ligands place an hydrogen bond donor moiety interacting with a carboxylate group from Asp26 (Figure 3C). Analyzing the MD simulations in detail revealed that in one of the replicas, an acetonitrile molecule remains trapped during several nanoseconds due to steric reasons in the region where this hydrophobic site was later established. This means that protein flexibility is not always enough to allow free diffusion of solvent probe molecules and thus careful inspection of the initial position of the probes should be made.

Lastly, concerning false negatives or missing interactions that account for the observed low sensitivities, the main cause is related to the inadequate sampling capacity of solvent probes in occluded binding regions. For example, glycogen phosphorylase has a relatively buried active site and the lack of solvent accessibility is evident in the MD simulations because the region was practically unexplored by most probes. In these cases, water has a clear advantage being able to penetrate faster and deeper. It is worth mentioning that this active site is predicted as totally non druggable by fpocket,³⁴ with a practically null druggability score (Table 1 in the Computational Methods). Less frequently, false negatives were found when protein–ligand interactions are established in regions highly exposed to the solvent, representing secondary interaction sites that may only appear when the ligand is already fixed in the cavity.

In conclusion, solvent mixture MD is able to predict protein–ligand interaction sites, especially those defining the protein’s pharmacophore, both with good sensitivity and specificity. Interestingly, hydrophobic probes achieve better performance than polar ones, thus it seems the directionality of hydrogen bonds does not improve probe accuracy. Taken together, the above-described results suggest that performing pure water and water/ethanol MD simulations are jointly powerful enough to

determine both polar, combining WS and ethanol–OH sites, and nonpolar, as revealed by ethanol HS interactions sites, and thus predict potential protein–ligand interactions, the other probes being mostly redundant in terms of their predictive power. The election of water or ethanol probes for obtaining hydrophilic sites will depend on the desired application because water, although being a more sensible probe even capable of reaching occluded binding sites, presents an increased number of false positive sites, hence making ethanol a more suitable option when specificity is required. The second key point is that solvent probes are able to find most pharmacophore defining interactions, as well as more than half of all observed possible protein–ligand interactions. The analysis also shows that usually more solvent sites are found than those later mimicked by the ligand functional groups. This fact could, however, reflect the incomplete sampling of the potential interactions in the protein active site by the present set of ligands or the presence of tightly bound waters capable of bridging protein–ligand interactions. Leaving this aside, to improve the predictive power of the solvent sites, we analyze several of their properties in relation to the chance that they define a pharmacophoric site in the following section.

As a final remark, to compare our method performance with other established and simpler mapping strategy we also computed protein hot spots with FTMap⁴⁴ for coagulation factor Xa and FGF receptor 1 kinase domain (a more thorough comparison between MD based and direct mapping methods can be found in ref 29). Briefly, the FTMap method docks 16 different probes to the protein surface, minimizes the energy and generates different consensus clusters of probes to identify druggable cavities in the structure of interest. As shown in Figure S3, even though both methods are able to detect key active site regions defining pharmacophoric interactions, those provided by

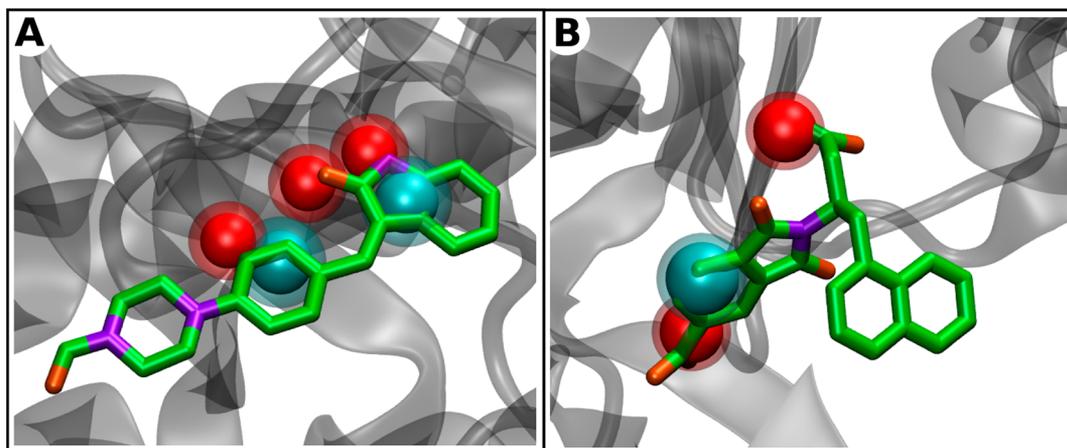


Figure 5. (A) FGF receptor 1-SU2 cocystal binding site structure (PDB ID 1agw) with displaced ethanol hydrophilic (red spheres) and hydrophobic (light blue spheres) sites superimposed. (B) AmpC β -lactamase-WH6 cocystal active site structure (PDB ID 2r9x) with selected displaced water (red spheres) and hydrophobic ethanol (light blue spheres) sites superimposed.

the MD derived solvent sites are better defined in terms of position and type of interactions. Although FTMap has the clear advantage of jointly using a wide variety of probes, the relative orientation of the different probe molecules from the same hot spot are sometimes ambiguous mainly due to a high dispersion of the clusters (as in the hinge region of FGF receptor 1 displayed in Figure S3, lower panel). The defined localization of MD derived solvent sites is probably related to the fact that explicit solvation permits water competition which reduces the appearance of weak interacting sites. Therefore, it is possible that although being very useful for the assessment of pocket bindability and protein druggability, FTMap is not precise enough to allow its sites to be used in docking and ΔG estimations. However, more work is needed to assess properly its potential for these applications. Finally, although FTMap is faster in terms of computational cost, performing ca. 100 ns of pure water and water/ethanol MD simulations, as required for solvent site identification, can be done in ca. 72 h using standard GPU implementation of AMBER, and represents a minor cost compared to, for example, the performance of Virtual Screening in which solvent information may be used in order to improve results.

Detailed Analysis of Solvent Sites Properties. We determined for all identified water and ethanol hydrophilic and hydrophobic sites the following properties (defined in Computational Methods): (i) probe finding probability (PFP), which is a direct measure of the protein–solvent site interaction free energy, (ii) R_{90} , which is a measure of the solvent site dispersion and entropy, (iii) nearest protein to solvent site distance (NPD) and (iv) site solvent accessible surface area (siSASA). Second, we classified the solvent sites in three groups, based on the information presented above: (a) those solvent sites reproducing protein–ligand pharmacophoric interactions of their own type (pharmacophoric solvent sites), (b) those solvent sites that although overlapping with the ligands do not correspond to pharmacophoric interactions (displaced solvent sites) and (c) those identified sites that do not overlap with any of the analyzed ligands.

Figure 4 presents the corresponding probability density functions for those properties where pharmacophoric solvent sites show a skewed distribution. The WS PFP (Figure 4A), for example, shows Gaussian-like distributions with a long tail to large PFP values (>10) that reveals the presence of tightly bound waters. The plot also shows that the distribution for the

pharmacophoric solvent sites is slightly skewed to larger values (maximum is ca. 7 compared to 5 for the others, first arrow in Figure 4A) and even displays second peak at very high PFP values (second arrow, Figure 4A). The minimal distance between WS and the protein (Figure 4B) shows that there is a slight increase of the proportion of pharmacophoric solvent sites at short distances (<3.5 Å), thus in direct contact with the protein surface, when compared to the others. On the other hand, both R_{90} and siSASA show similar distribution for the three types of sites and thus do not have capacity to discriminate among them. Similar results are obtained for ethanol–OH sites (data not shown). Finally, hydrophobic sites determined by ethanol methyl group (Figure 4C,D) also show that pharmacophoric sites tend to have larger PFP values and there is a slight increase in the fraction of these sites between 3.5 and 4.5 Å from the protein surface, whereas no significant differences are shown by other properties.

Summarizing, the analysis of the solvent site properties shows that although there is not a large difference in the properties of the pharmacophoric solvent sites with respect to other sites, a good rule of thumb for selecting potential solvent sites reproducing pharmacophoric interactions is to choose the ones with high finding probability and close, i.e., in direct contact, with the protein surface.

Finally, before moving forward to the potential use of the determined solvent sites to improve ligand docking and estimate ligand binding free energies, we need to determine the sampling efficiency and statistical uncertainty related to PFP (and ΔG , eq 6) values of the solvent sites. We first determined the mean value (x) and standard deviation (sd) in both parameters for each solvent site in the three independent MD simulations. The resulting coefficient of variation values, computed as sd/x , were averaged for all sites corresponding to the same probe. The results show that PFP values for water and ethanol sites, which are used for docking and free energy estimates, have variation values around 30% of the PFP, representing less than 16% of ΔG (Table S3 shows ΔG coefficient of variation for all solvents). Second, we determined the average number of times that a probe molecule is exchanged in a given solvent site. The results, presented in Table S4, show that although on average water molecules are exchanged nearly 3 times in each nanosecond, other solvents are exchanged between one and two times, reaching more than 20 molecule exchanges per solvent site when considering the whole 20 ns MD simulation. Both these results

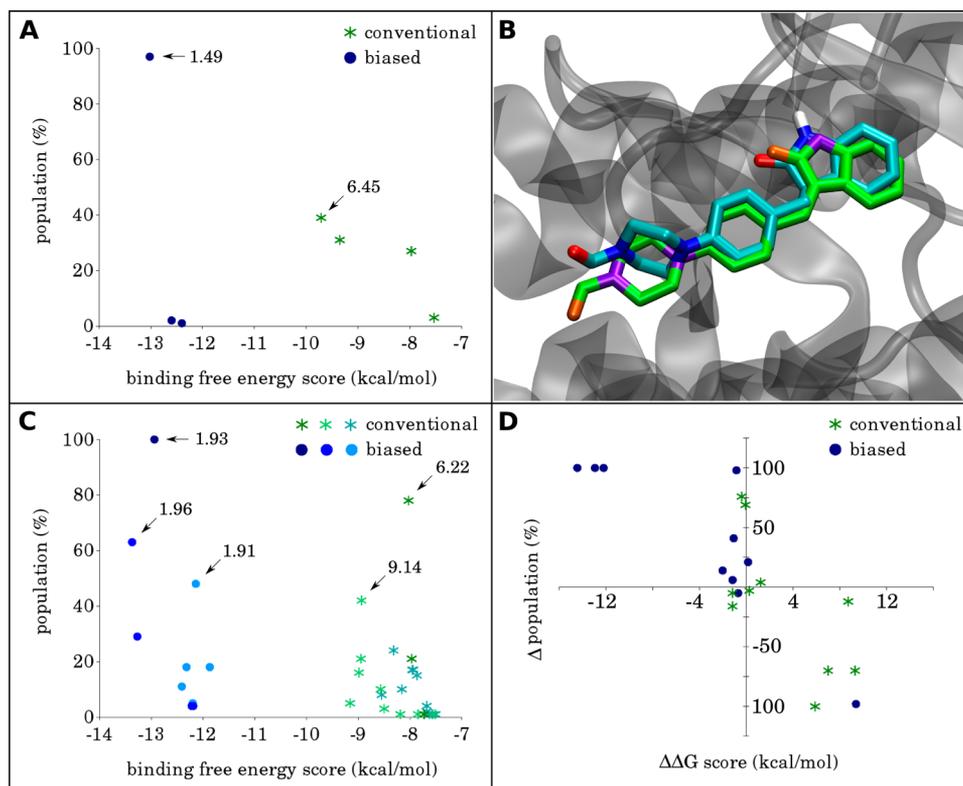


Figure 6. Population vs binding free energy score (panels A and C) and Δ population vs $\Delta\Delta G$ score (panel D) plots using the conventional (CADM) and biased (SSBDM) docking methods with ethanol hydrophilic and hydrophobic sites for FGF receptor 1. (A) Redocking of SU2 ligand (PDB ID 1agw). (B) Structure of the best ranked ligand pose obtained with the biased method superimposed to the reference crystal structure for the redocking experiment shown in panel A. (C) Docking of SU2 ligand against three different protein structures obtained from explicit water MD simulations. (D) Cross-docking experiments for all ligands from the data set against protein structure from PDB ID 4f64. The values indicated with arrows represent the ligand heavy atom rmsd, in Å, between the predicted complex pose and the reference complex structure.

show that sampling efficiency and statistical uncertainty of the solvent sites and their associated values are adequate for the desired applications.

Using Solvent Site Derived Information To Improve Molecular Docking. Previous works from our group showed that it is possible to use water sites to improve significantly docking of carbohydrates to lectins^{26,42} and also possibly common drugs to target proteins, as shown by the docking of HTC to AmpC β -lactamase used as a representative example.¹⁹ The rationale behind the improvement lies in the fact that WS, found in the ligand free protein, mimic the positions of the carbohydrate hydroxyl groups or, in general, ligand functional groups with hydrogen bonding capabilities in the corresponding protein–ligand complexes. Furthermore, WS PFP is a good predictor of its chance of replacement. Therefore, and on the basis of the present results, we decided to analyze whether we could improve the docking of any type of ligand using both the hydrophilic and hydrophobic sites determined by water and ethanol. For this sake, and as performed in previous works, we compared the performance of the conventional (or unmodified) Autodock4 docking method (CADM) with the same docking protocol biasing the potential with the determined solvent sites. We will call this method solvent site biased docking method (SSBDM). Details on the formulation of the biased potential can be found in the [Computational Methods](#).

Because we already analyzed the potential of the WS in previous works, we decided to focus on ethanol amphiphilic properties. Therefore, alongside with the use of ethanol–OH sites to bias the position of ligand groups with hydrogen bonding

capabilities, we used ethanol derived hydrophobic sites to bias the location of ligand aromatic rings. As a test case, we selected FGF receptor 1 and first performed a typical redocking experiment of the cocrystallized ligand (PDB ID 1agw) using both docking methods. [Figure 5A](#) shows all the ethanol sites in the FGFR1 binding site displaced by SU2 ligand, superimposed to the complex crystal structure. The two hydrophobic sites, depicted as light blue spheres, clearly overlay both ligand aromatic rings, whereas two of the three hydrophilic sites (red spheres) are well replaced by N/O ligand atoms.

When Autodock is used, it is common practice to perform 100 docking runs for each ligand and to cluster the resulting poses into similar sets according to a heavy atom rmsd threshold between them, thus defining a population value for each set or cluster, which is also characterized by the lowest binding energy of its members. [Figure 6A](#) shows the population vs binding energy score plot for the conventional and biased redocking methods. For the conventional method, the plot displays no clear outlier, and the best ranked complex (best energy/highest population) shows a ligand rmsd of 6.45 Å against the reference structure. The origin of this deviation ([Figure S4](#), panel A) is that the predicted pose is displaced and deeply buried inside the pocket when compared to the reference complex. Interestingly, the best predicted pose for the CADM (rmsd of 1.60 Å) bears poor binding energy score and population, therefore being the worst ranked structure. The SSBDM significantly improves the results. A clear outlier is now identified in the population vs energy plot ([Figure 6A](#), blue dots), which corresponds to a pose displaying a rmsd of only 1.49 Å compared to the reference

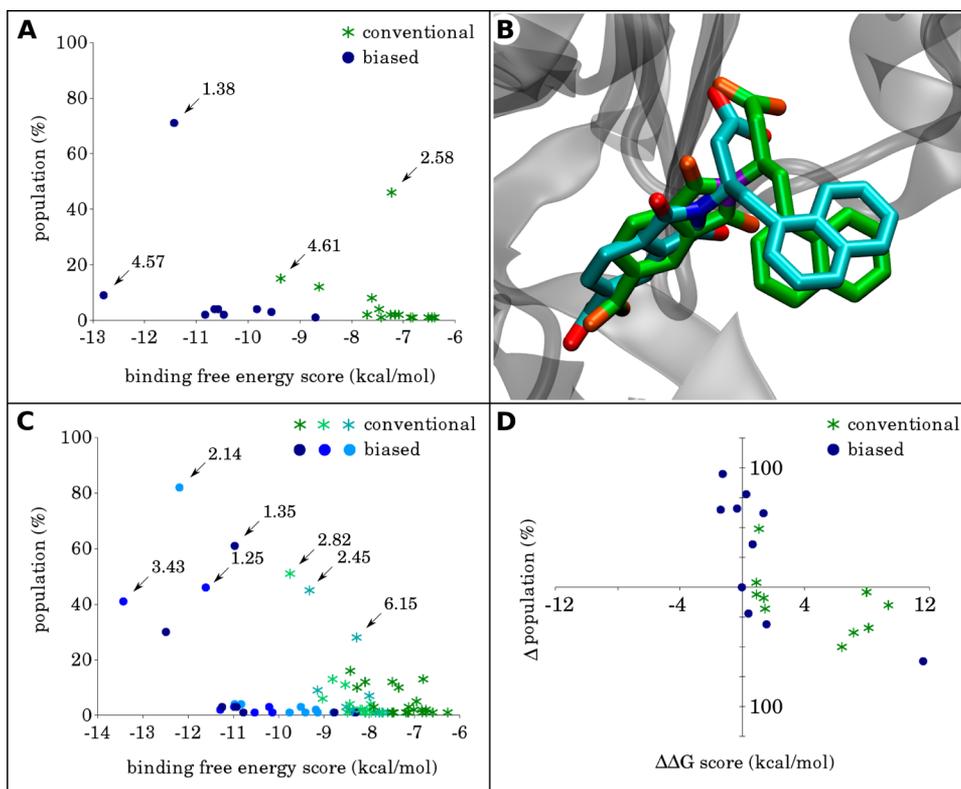


Figure 7. Population vs binding free energy score (panels A and C) and Δ population vs $\Delta\Delta G$ score (panel D) plots using the conventional (CADM) and biased (SSBDM) docking methods with water and ethanol hydrophobic sites for AmpC β -lactamase. (A) Redocking of WH6 ligand (PDB ID 2r9x). (B) Structure of the highest population ligand pose obtained with the biased method superimposed to the reference crystal structure for the redocking experiment shown in panel A. (C) Docking of WH6 ligand against three different protein structures obtained from explicit water MD simulations. (D) Cross-docking experiments for 10 ligands (Table S5) against protein structure from PDB ID 2r9x. The values indicated with arrows represent the ligand heavy atom rmsd, in Å, between the predicted complex pose and the reference PDB complex structure.

complex. Superposition of predicted and reference complexes (Figure 6B) shows that both ligands share the same molecular interactions and are identically placed, except for a small rotation of the outermost aliphatic ring, which is exposed to the solvent and does not alter the overall binding mode. Figure 6C shows similar improvement using the biased method when docking is performed using different receptor structures derived from MD simulations. These results are not quite unexpected because the solvent sites already capture some of the receptor dynamics. Finally, we cross-docked FGFR1 ligands from the data set (Table 1) to a single crystal structure using both methods. Figure 6D plots the differences in predicted binding ΔG score ($\Delta\Delta G$) and population (Δ population) between the correct ligand pose (that with rmsd < 2 Å) and the best ranked of the remaining poses. The graphic shows that biased docking significantly improves the results, finding the correct poses with the best energy and population (upper left quadrant) in most cases. There is only one incorrect prediction (lower right quadrant) for the biased method corresponding to a relatively small ligand, extracted from PDB ID 3js2, that bears no more than two molecular interactions with the protein and whose ligand pose is also wrongly predicted by the conventional method. Overall, these results nicely demonstrate the potential improvement of solvent site biased docking using ethanol as a probe.

To complete the picture, we also tested the SSBDM using the WS to bias hydrophilic interactions mixed with ethanol hydrophobic sites to bias the position of aromatic rings. We used as a test case the β -lactamase–WH6 complex (PDB ID 2r9x). Figure 5B shows the WS (red spheres) and ethanol HS

(light blue sphere) displaced by WH6 ligand, superimposed to the complex crystal structure. As for the FGF receptor, we performed redocking, MD receptor based and cross docking experiments. The results presented in Figure 7 show that in most cases SSBDM performance is better, although the improvement is not as good as for FGF. For most ligands, the conventional docking completely fails to identify the correct pose, whereas the biased method is able to prioritize it either due to a lower binding energy or a higher population. As an example, for WH6 ligand, in the conventional redocking the lowest energy pose is completely wrongly predicted (rmsd 4.61 Å) whereas the highest population, which is the best predicted pose, displays an rmsd of 2.58 Å (the ligand is adequately oriented but slightly shifted, see Figure S4, panel B). The results for the biased docking method considering all WS and ethanol HS formed in the active site, on the contrary, shows a clear outlier with high population above 70% and good relative binding energy score, and its pose nicely resembles the crystal pose (rmsd = 1.38 Å), as shown in Figure 7B. Finally, Figure 7C shows that only the biased method is able to find the correct ligand pose when docking against MD derived protein structures and Figure 7D shows that the biased method used in cross docking experiments increases the number of correct predictions with either energy or population outliers from 10% to 60% of the assayed ligands.

Summarizing, the above-described results show that biasing the docking using information on the potential interaction sites derived from explicit mixed solvent MD simulations, significantly improves the prediction of protein–ligand complex structures in terms of both accuracy (obtained poses show lower rmsd against

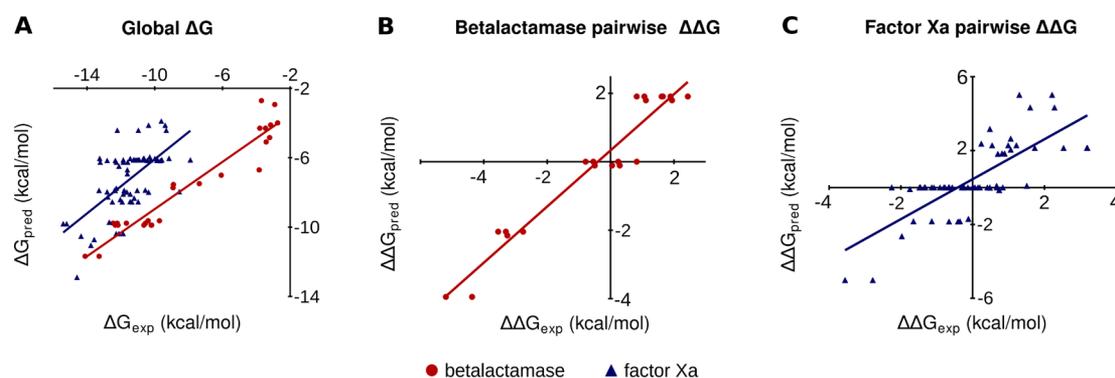


Figure 8. Predicted binding free energy from replaced ethanol hydrophilic and hydrophobic sites vs experimental binding free energy plots for human coagulation factor Xa and *E. coli* AmpC β -lactamase. (A) Predicted vs experimental binding ΔG plots for all the ligands from the set (see Tables S6 and S7 in the Supporting Information): β -lactamase depicted as red dots and factor Xa as blue triangles. (B and C) Predicted vs experimental binding $\Delta\Delta G$ plots for similar ligand pairs for β -lactamase (B) and factor Xa (C). Linear regression curves are shown for all plots. R^2 and other meaningful parameter values for these linear regressions are shown in Table 5.

Table 5. Linear Regression Results for Different Ligand sets of β -Lactamase and Factor Xa

correlated variable	ligand set (protein) ^a	# ligands	R^2	slope	rmse (kcal/mol) ^b	$\Delta\Delta G$ sign prediction ^c
absolute (ΔG) correlation	all ligands (β -lactamase)	26	0.92	0.68	1.53	na
	all ligands (factor Xa)	72	0.38	0.78	4.49	na
	similarity set 1 (factor Xa)	5	0.80	2.01	2.10	na
	similarity set 2 (factor Xa)	9	0.68	1.29	3.78	na
	similarity set 3 (factor Xa)	7	0.73	0.77	4.16	na
	similarity set 4 (factor Xa)	4	0.91	2.07	3.50	na
relative ($\Delta\Delta G$) correlation	similar pairs (β -lactamase)	7 ^d	0.94	0.82	0.76	88%
	similar pairs (factor Xa)	43 ^d	0.58	1.08	1.31	97%

^aSee Tables S6 and S7 in SI for ligand set composition. ^bRoot mean squared error (rmse) is calculated from predicted and experimental ΔG (or $\Delta\Delta G$) values. ^cFor $\Delta\Delta G$ correlations, sign prediction is also calculated and corresponds to the percentage of times that predicted and experimental $\Delta\Delta G$ values had the same sign, taking into account the cases in which predicted $\Delta\Delta G$ values were different from 0 ($\Delta\Delta G_{\text{pred}} \neq 0$ in 81% of the cases for β -lactamase and 52% for factor Xa); na = not applicable. ^dThe number of points involved in the regression analysis is 21 for β -lactamase and 63 for factor Xa, reflecting the total number of similar pairs.

the reference) and predictive capacity (best predicted poses show higher population and lower binding energies), thus establishing a proof of concept that encourages future studies toward its application in virtual screening campaigns.

Relation of Solvent Site Properties with Ligand Binding Thermodynamics. Our final analysis concerns the use of solvent sites to estimate ligand binding free energies, building at the same time a solvent site based scoring function to rank a set of ligands according to their binding affinity. To test this idea, we selected two proteins, human coagulation factor Xa and *E. coli* AmpC β -lactamase, for which there is a wide set of known ligands whose binding mode and affinity have been experimentally characterized (see the list of complexes in the Supporting Information, Tables S6 and S7). To compute the binding free energy (ΔG_{pred}) of any ligand, starting from the protein–ligand complex crystal structure, we added the energy contribution of all solvent sites overlapping ligand atoms capable of establishing the same type of interactions (see Computational Methods). We used the hydrophobic sites determined from ethanol methyl group and the hydrophilic sites derived from either ethanol–OH or water. We will show the results using the ethanol derived hydrophilic sites because they are slightly better than those obtained using the WS. These are shown in Figure 8 and summarized in Table 5.

Figure 8A shows ΔG_{pred} vs the experimentally determined binding free energy (ΔG_{exp}) for both proteins, including all available ligands in the set. The results show that performance is highly dependent on the protein system. The correlation for

AmpC β -lactamase is surprisingly good ($R^2 = 0.92$, $\text{rmse} = 1.53$ kcal/mol) and shows the potential of the solvent sites to compute free energies from protein–ligand crystal complexes. Although for factor Xa the results are not as well correlated ($R^2 = 0.38$, $\text{rmse} = 4.49$ kcal/mol), a closer look shows some interesting points that allow further improvement. First, when analyzing the ligands chemical diversity it is evident that whereas for β -lactamase most ligands are small and of similar size, factor Xa ligands are diverse in size and chemical composition, many being larger with several torsions, and so the entropy loss due to ligand binding, which is not accounted by the present method, would be variable between ligands. Second, because we are only considering the replacement of solvent sites to compute ΔG , those ligands that replace the same sites have the same ΔG_{pred} despite being dissimilar and thus showing different affinities. This behavior is evidenced in the plot as ligands lying on the same horizontal line. A final point of notice concerning this plot is that although ΔG values are slightly underestimated (20–50%), the accuracy is far better than that observed for MM-GB(PB)SA based estimations.⁴⁵

To analyze further the method predictive power, for each protein we separated the ligands in groups of more than three members according to their chemical similarity and computed the correlation between ΔG_{exp} and ΔG_{pred} separately for each group. The results for factor Xa, also presented in Table 5, show that correlations are significantly improved, although mean error values are still large and there is significant variation in the performance depending on the set under study. For β -lactamase,

there is only one set of similar ligands and the initial correlation is kept ($R^2 = 0.91$).

As a further attempt to improve the performance of our method, we decided to compute relative binding free energies ($\Delta\Delta G_{\text{pred}}$) between pairs of similar compounds from all the similarity sets. Figure 8B,C shows $\Delta\Delta G_{\text{pred}}$ results for all the similar pairs. The β -lactamase (Figure 8B) preserves its excellent performance. Table 5 shows that not only the good correlation measured by R^2 and the approximate unitary slope are maintained, but also a low rmse value of 0.76 kcal/mol is obtained and, most important, a correct sign prediction is attained. Out of all the cases in which $\Delta\Delta G_{\text{pred}}$ is different from 0, predicted and experimental values have the same sign in 88% of them. This is of great relevance when evaluating whether a ligand is to be more (or less) potent than another one.

Switching to factor Xa (Figure 8C), it becomes evident that the correlation (R^2 is 0.58) is better than that for the absolute values. Moreover, the slope is practically 1 and the rmse value is now acceptable (1.31 kcal/mol). Most important, $\Delta\Delta G$ sign prediction is again well accomplished because out of 33 non-null $\Delta\Delta G$ predictions, 32 had the same sign as $\Delta\Delta G_{\text{exp}}$, so there is only one wrongly predicted case ($\Delta\Delta G_{\text{exp}} = 0.9$ kcal/mol and $\Delta\Delta G_{\text{pred}} = -0.1$ kcal/mol). The results thus evidence that the important information the method is able to yield when complex structural information is available, concerns its high confidence in predicting the relative potency of binding ligands when $\Delta\Delta G_{\text{pred}}$ has nonzero value (i.e., ligands replacing at least one different solvent site). On the contrary, a weakness of the method is that it is not always able to provide this information, i.e., when $\Delta\Delta G_{\text{pred}} = 0$ (19% of the cases for β -lactamase and 48% for factor Xa).

Concerning the fact that ligands replacing the same solvent sites will display the same ΔG_{pred} , it is clear that other contributions, ligand binding entropy, size, etc., should be included to prevent it. Although, this is out of the scope of the present work, and despite the limited sample, we can measure how much of the overall variance in the ΔG_{exp} is captured by the solvent sites, and how much is left out. Our results show that although ΔG_{exp} shows an overall mean variance of 1.0 kcal/mol for the sets of similar ligands analyzed in factor Xa (2.7 kcal/mol for β -lactamase), the average variance of the sets replacing exactly the same solvent sites is only 0.3 kcal/mol for both proteins. In other words, the solvent sites are able to represent the same range of affinities observed experimentally, but there is still ca. 30% variation in the binding free energies that results from other properties that contribute to the ligand binding free energy.

DISCUSSION

Given the potential of MD simulations in explicit solvent, with the addition of molecular probes to reveal ligand interaction sites, in the past decade several groups used this kind of approach to characterize protein binding sites and relate them to the known protein–ligand complexes.^{29–31,46} Most of them focused on small sets of protein–ligand complexes and a few probes. They were usually directed toward the detection and characterization of the binding site in terms of their potential interactions, and suggested, but felt short, on further applications of the probe derived information, such as binding free energy estimations. The first part of our work extends (to 121 complexes in 18 different proteins) and confirms previous observations, related to the probes capabilities to unravel future protein–ligand interactions. Most important, having tested a variety of water miscible solvent probes (water, ethanol, acetamide, methylammonium acetate and acetonitrile) in the same proteins, allows

us to conclude that pure water and water/ethanol simulations are jointly powerful enough to determine both polar and non polar interaction sites, the other probes being mostly redundant. Moreover, even if no highly hydrophobic probe (such as benzene or propane as in the SILCS method^{30,47}) was used, more than 85% of the aliphatic and aromatic hydrophobic interaction sites could be determined combining both ethanol and acetonitrile methyl probes. This could be important because highly hydrophobic probes (butane or phenol) have been shown to result partially in protein denaturation.⁴⁸ Also, in line with previous reports,²⁹ analysis of the solvent site properties shows that those sites with higher finding probability, which is translated in higher affinity (see eq 6 in the Computational Methods), tend to be those that correspond to pharmacophoric sites.

Moving toward potential applications, we first showed how solvent sites can be used to improve molecular docking calculations. In line with our previous works, focused on lectin–carbohydrate complexes, we used the PFP and R_{90} to modify the grid maps computed by Autodock4 scoring function, favoring those ligand poses where polar and/or nonpolar functional groups replace the corresponding hydrophilic and/or hydrophobic solvent sites, respectively. Docking results are significantly improved in terms of accuracy, i.e., the method capacity to reproduce the structure of the protein–ligand reference complex, and precision, with right results displaying better energy and population than the false positives. Detailed analysis of how the solvent site biased docking method improves Autodock performance shows that key to success is a better scoring function and not an increased sampling. Having thus established the proof of concept for using the solvent sites to improve docking, the next logical step is to implement them in a virtual screening scheme.

The second application of the solvent sites concerns their use for estimating ligand binding free energies. Employing a similar approach as that used for docking, we were able to estimate ΔG values showing in some cases, as for AmpC β -lactamase, exceptional agreement with experimentally derived data. In general, predicted binding free energy values slightly underestimate the experimental ones and results are better when ΔG for similar sets of ligands are compared or if relative, $\Delta\Delta G$, values between pairs of similar ligands are computed. As expected, and in agreement with other works on the subject,²⁰ the predictive power depends strongly on the set of ligands that are evaluated.

The main limitations for predicting binding free energies by the present method arise from the presence of set of ligands that either replace the exact same solvent sites or are too small (fragments) and thus replace no more than one solvent site. For the first type of ligands, the problem is of moderate relevance because the variance in experimental ΔG values is rather small when compared to the overall range of ΔG_{exp} evaluated for the two assayed proteins. In any case, although the variation inside these ligand sets cannot be predicted, the ligands can be properly compared with ligands from other sets and be well discriminated in terms of their absolute binding free energy.

It is interesting to compare the present method with that developed by Abel's group,²⁰ where the energy, entropy and relative position of water sites were used to build a scoring function useful to predict binding $\Delta\Delta G$ between congeneric ligand pairs that differ by only small chemical modifications. Our implementation is simpler because it only considers in a binary fashion whether the ligand replaces (or not) a given solvent site. Although, this clear has limitations, it nonetheless shows good

performance in a comparative way. In any case, to assess these methods predictive power, as well as other alternatives, further systematic work is needed.

A last point of notice concerns the fact that albeit surely failing to consider several contributions to the binding free energy, both ΔG_{pred} and the biased docking scoring function are able to reveal the relevant aspects of the underlying thermodynamics of the binding process. In other words, the solvent site related free energy (derived from eq 6) must be a very good estimate of the free energy contribution of the corresponding ligand fragment (or functional group) that replaces it to the overall ligand affinity.

Two questions then arise. The first one is related to what contributions we are missing and how could they be added in order to be able to estimate different ΔG values for different ligands that nonetheless replace the same sites. Detailed analysis of our results suggests that one key factor is to account for those interactions that are not represented by the solvent sites. Although this seems obvious, it is not straightforward to implement in order to avoid double counting of interactions. Also important seems to be the desolvation of the ligand beyond the functional groups that replace the sites, and mostly related to ligand size. Third, loss of conformational entropy of both ligand and receptor should be considered, specially for large and flexible ligands. Finally, the presented approach also assumes additivity and independence between solvent sites, ignores molecular orientation and group identity and neglects the difference in the entropic cost of fixing many solvent molecules compared to only one ligand. The second point concerns comparison with other methods such as MM-GB(PB)SA that are known to have poor performance, especially for ligands bearing charged groups.⁴⁵ Our results strongly suggest that solvent probes correctly describe the free energy of transferring the corresponding group from the bulk solvent to the binding site. This implies that within a given force field, and if proper sampling is achieved, probe interactions between bulk water and the protein binding site are well balanced resulting in accurate free energy differences. On the contrary, MM-GB(PB)SA methods usually combine direct explicit evaluation of the ligand–protein interaction energy with a continuum model to determine solvation energies and the contributions are not well balanced. The fact that force fields in explicit solvent are able to balance correctly protein–ligand and protein–solvent interactions is also supported by more recent works showing that accurate absolute binding free energies can be obtained by directly sampling the ligand binding and release processes in explicit solvent MD simulations⁴⁹ or by using free energy perturbation methods.⁵⁰ Although, our solvent site method is not as accurate and does not yield the detail achieved by this type of simulations, its computational cost is significantly smaller, especially when the same receptor is evaluated against many ligands.

CONCLUSION

By performing MD simulations in explicit water and water–cosolvent (ethanol, acetamide, methylammonium acetate, acetonitrile) mixtures for a large set of proteins, we show that combining water and ethanol identified hot spots (solvent sites) allows to sample over 70% of all possible protein–ligand interactions, especially those that are most important and represent the protein pharmacophore. We also show that information derived from the other probes is redundant. Finally, and most relevant, we demonstrate that the identified solvent sites can be used to improve significantly ligand docking in terms of accuracy and precision, and that accurate predictions of ligand

binding free energies, together with a relative ranking of ligands affinity, can be performed.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.6b00678.

Additional information on solvents force field parameters, sampling, ligand sets for docking and binding free energy prediction and ionic probe results (PDF)

AUTHOR INFORMATION

Corresponding Authors

*A. G. Turjanski. E-mail: adrian@qi.fcen.uba.ar.

*M. A. Martí. E-mail: marti.marcelo@gmail.com.

ORCID

Juan Pablo Arcon: 0000-0003-3350-1576

Lucas A. Defelipe: 0000-0001-7859-7300

Marcelo A. Martí: 0000-0002-7911-9340

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was financed by grants from Agencia Nacional de Promoción Científica y Tecnológica (PICTO-2012-0057), Consejo Nacional de Investigaciones Científicas y Técnicas (PIP 112 201101 0085) and Universidad de Buenos Aires (UBACyT012) to M.A.M., and Spanish Ministerio de Economía grant SAF2015-68749-R to X.B. J.P.A. acknowledges Consejo Nacional de Investigaciones Científicas y Técnicas for doctoral fellowship and BEC.AR program (Ministerio de Modernización de la República Argentina) for short term scholarship, and is deeply in debt to ALN for enriching discussions.

REFERENCES

- (1) Bissantz, C.; Kuhn, B.; Stahl, M. *A Medicinal Chemist's Guide to Molecular Interactions*. *J. Med. Chem.* **2010**, *53*, 5061–5084.
- (2) Medina-Franco, J. L.; Mendez-Lucio, O.; Martínez-Mayorga, K. The Interplay between Molecular Modeling and Chemoinformatics to Characterize Protein–Ligand and Protein–Protein Interactions Landscapes for Drug Discovery. In *Advances in Protein Chemistry and Structural Biology*, first ed.; Karabencheva-Chistova, T., Ed.; Academic Press: London, United Kingdom, 2014; Vol. 16, pp 1–37.
- (3) Drews, J. Drug Discovery: A Historical Perspective. *Science* **2000**, *287*, 1960–1964.
- (4) Loding, W.; Rodríguez-Esteban, R.; Hill, J.; Freeman, T.; Miglietta, J. Cheminformatic/bioinformatic Analysis of Large Corporate Databases: Application to Drug Repurposing. *Drug Discovery Today: Ther. Strategies* **2011**, *8*, 109–116.
- (5) Fadda, E.; Woods, R. J. Molecular Simulations of Carbohydrates and Protein–Carbohydrate Interactions: Motivation, Issues and Prospects. *Drug Discovery Today* **2010**, *15*, 596–609.
- (6) Powlesland, A. S.; Quintero-Martínez, A.; Lim, P. G.; Pipirou, Z.; Taylor, M. E.; Drickamer, K. Engineered Carbohydrate-Recognition Domains for Glycoproteomic Analysis of Cell Surface Glycosylation and Ligands for Glycan-Binding Receptors. *Methods Enzymol.* **2010**, *480*, 165–179.
- (7) Du, X.; Li, Y.; Xia, Y.; Ai, S.; Liang, J.; Sang, P.; Ji, X.; Liu, S.-Q. Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods. *Int. J. Mol. Sci.* **2016**, *17*, 144.
- (8) Konc, J.; Lesnik, S.; Janežic, D. Modeling Enzyme–Ligand Binding in Drug Discovery. *J. Cheminf.* **2015**, *7*, 1–8.
- (9) Mattos, C.; Bellamacina, C. R.; Peisach, E.; Pereira, A.; Vitkup, D.; Petsko, G. A.; Ringe, D. Multiple Solvent Crystal Structures: Probing

Binding Sites, Plasticity and Hydration. *J. Mol. Biol.* **2006**, *357*, 1471–1482.

(10) English, A. C.; Groom, C. R.; Hubbard, R. E. Experimental and Computational Mapping of the Binding Surface of a Crystalline Protein. *Protein Eng., Des. Sel.* **2001**, *14*, 47–59.

(11) Liepinsh, E.; Otting, G. Organic Solvents Identify Specific Ligand Binding Sites on Protein Surfaces. *Nat. Biotechnol.* **1997**, *15*, 264–268.

(12) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* **1996**, *274*, 1531–1534.

(13) Dechene, M.; Wink, G.; Smith, M.; Swartz, P.; Mattos, C. Multiple Solvent Crystal Structures of Ribonuclease A: An Assessment of the Method. *Proteins: Struct., Funct., Genet.* **2009**, *76*, 861–881.

(14) Davies, D. R.; Begley, D. W.; Hartley, R. C.; Staker, B. L.; Stewart, L. J. Predicting the Success of Fragment Screening by X-Ray Crystallography. *Methods Enzymol.* **2011**, *493*, 91–114.

(15) Seco, J.; Luque, F. J.; Barril, X. Binding Site Detection and Druggability Index from First Principles. *J. Med. Chem.* **2009**, *52*, 2363–2371.

(16) Guvench, O.; MacKerell, A. D. J. Computational Fragment-Based Binding Site Identification by Ligand Competitive Saturation. *PLoS Comput. Biol.* **2009**, *5*, e1000435.

(17) Bakan, A.; Nevins, N.; Lakdawala, A. S.; Bahar, I. Druggability Assessment of Allosteric Proteins by Dynamics Simulations in the Presence of Probe Molecules. *J. Chem. Theory Comput.* **2012**, *8*, 2435–2447.

(18) Lexa, K. W.; Carlson, H. A. Full Protein Flexibility Is Essential for Proper Hot-Spot Mapping. *J. Am. Chem. Soc.* **2011**, *133*, 200–202.

(19) Lopez, E. D.; Arcon, J. P.; Gauto, D. F.; Petruk, A. A.; Modenutti, C. P.; Dumas, V. G.; Marti, M. A.; Turjanski, A. G. WATCLUST: A Tool for Improving the Design of Drugs Based on Protein-Water Interactions. *Bioinformatics* **2015**, *31*, 3697–3699.

(20) 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*, 2817–2831.

(21) Hu, B.; Lill, M. A. Protein Pharmacophore Selection Using Hydration-Site Analysis. *J. Chem. Inf. Model.* **2012**, *52*, 1046–1060.

(22) 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*, 2769–2780.

(23) Wong, S. E.; Lightstone, F. C. Accounting for Water Molecules in Drug Design. *Expert Opin. Drug Discovery* **2011**, *6*, 65–74.

(24) Ladbury, J. E. Just Add Water! The Effect of Water on the Specificity of Protein-Ligand Binding Sites and Its Potential Application to Drug Design. *Chem. Biol.* **1996**, *3*, 973–980.

(25) Li, Z.; Lazaridis, T. Water at Biomolecular Binding Interfaces. *Phys. Chem. Chem. Phys.* **2007**, *9*, 573–581.

(26) Modenutti, C.; Gauto, D.; Radosky, L.; Blanco, J.; Turjanski, A.; Hajos, S.; Marti, M. A. Using Crystallographic Water Properties for the Analysis and Prediction of Lectin-Carbohydrate Complex Structures. *Glycobiology* **2015**, *25*, 181–196.

(27) Li, Z.; Lazaridis, T. Computing the Thermodynamic Contributions of Interfacial Water. In *Computational Drug Discovery and Design*; Baron, R., Ed.; Springer New York: New York, NY, 2012; Vol. 819, pp 393–404.

(28) Ichihara, O.; Shimada, Y.; Yoshidome, D. The Importance of Hydration Thermodynamics in Fragment-to-Lead Optimization. *ChemMedChem* **2014**, *9*, 2708–2717.

(29) Alvarez-Garcia, D.; Barril, X. Molecular Simulations with Solvent Competition Quantify Water Displaceability and Provide Accurate Interaction Maps of Protein Binding Sites. *J. Med. Chem.* **2014**, *57*, 8530–8539.

(30) Raman, E. P.; Yu, W.; Lakkaraju, S. K.; Mackerell, A. D. Inclusion of Multiple Fragment Types in the Site Identification by Ligand Competitive Saturation (SILCS) Approach. *J. Chem. Inf. Model.* **2013**, *53*, 3384–3398.

(31) Yang, C. Y.; Wang, S. Computational Analysis of Protein Hotspots. *ACS Med. Chem. Lett.* **2010**, *1*, 125–129.

(32) Lakkaraju, S. K.; Raman, E. P.; Yu, W.; Mackerell, A. D. Sampling of Organic Solutes in Aqueous and Heterogeneous Environments Using Oscillating Excess Chemical Potentials in Grand Canonical-like Monte Carlo-Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 2281–2290.

(33) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55*, 6582–6594.

(34) Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: An Open Source Platform for Ligand Pocket Detection. *BMC Bioinf.* **2009**, *10*, 168.

(35) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(36) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B. P.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, L.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *Amber 12*; University of California: San Francisco, CA, 2012.

(37) 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*, 712–725.

(38) Goetz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER - Part 1: Generalized Born. *J. Chem. Theory Comput.* **2012**, *8*, 1542–1555.

(39) Salomon-Ferrer, R.; Goetz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs - Part 2: Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878–3888.

(40) Gauto, D. F.; Di Lella, S.; Guardia, C. M. A.; Estrin, D. A.; Marti, M. A. Carbohydrate-Binding Proteins: Dissecting Ligand Structures through Solvent Environment Occupancy. *J. Phys. Chem. B* **2009**, *113*, 8717–8724.

(41) 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*, 808–813.

(42) Gauto, D. F.; Petruk, A. A.; Modenutti, C. P.; Blanco, J. L.; Di Lella, S.; Marti, M. A. Solvent Structure Improves Docking Prediction in Lectin-Carbohydrate Complexes. *Glycobiology* **2013**, *23*, 241–258.

(43) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.

(44) Kozakov, D.; Grove, L. E.; Hall, D. R.; Bohnuud, T.; Mottarella, S. E.; Luo, L.; Xia, B.; Beglov, D.; Vajda, S. The FTMap Family of Web Servers for Determining and Characterizing Ligand-Binding Hot Spots of Proteins. *Nat. Protoc.* **2015**, *10*, 733–755.

(45) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. I. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69–82.

(46) Ghanakota, P.; Carlson, H. A. Moving beyond Active-Site Detection: MixMD Applied to Allosteric Systems. *J. Phys. Chem. B* **2016**, *120*, 8685–8695.

(47) Yu, W.; Lakkaraju, S. K.; Raman, E. P.; Fang, L.; Mackerell, A. D. Pharmacophore Modeling Using Site-Identification by Ligand Competitive Saturation (SILCS) with Multiple Probe Molecules. *J. Chem. Inf. Model.* **2015**, *55*, 407–420.

(48) Foster, T. J.; MacKerell, A. D.; Guvench, O. Balancing Target Flexibility and Target Denaturation in Computational Fragment-Based Inhibitor Discovery. *J. Comput. Chem.* **2012**, *33*, 1880–1891.

(49) Limongelli, V.; Bonomi, M.; Parrinello, M. Funnel Metadynamics as Accurate Binding Free-Energy Method. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 6358–6363.

(50) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* **2015**, *137*, 2695–2703.