JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Role of Water in Ligand Binding to Maltose-Binding Protein: Insight from a New Docking Protocol Based on the 3D-RISM-KH Molecular Theory of Solvation

WenJuan Huang,[†,‡] Nikolay Blinov,[†,‡] David S. Wishart,[§,∥,‡] and Andriy Kovalenko*[,‡,†]

[†]Department of Mechanical Engineering, University of Alberta, Edmonton, AB T6G 2G8, Canada

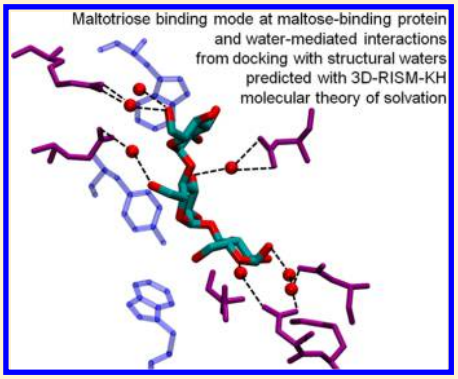
[‡]National Institute for Nanotechnology, 11421 Saskatchewan Dr., Edmonton, AB T6G 2M9, Canada

[§]Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2G8, Canada

[∥]Department of Computing Science, University of Alberta, Edmonton, AB T6G 2G8, Canada

Supporting Information

ABSTRACT: Maltose-binding protein is a periplasmic binding protein responsible for transport of maltooligosaccarides through the periplasmic space of Gram-negative bacteria, as a part of the ABC transport system. The molecular mechanisms of the initial ligand binding and induced large scale motion of the protein's domains still remain elusive. In this study, we use a new docking protocol that combines a recently proposed explicit water placement algorithm based on the 3D-RISM-KH molecular theory of solvation and conventional docking software (AutoDock Vina) to explain the mechanisms of maltotriose binding to the apo-open state of a maltose-binding protein. We confirm the predictions of previous NMR spectroscopic experiments on binding modes of the ligand. We provide the molecular details on the binding mode that was not previously observed in the X-ray experiments. We show that this mode, which is defined by the fine balance between the protein−ligand direct interactions and solvation effects, can trigger the protein's domain motion resulting in the holo-closed structure of the maltose-binding protein with the maltotriose ligand in excellent agreement with the experimental data. We also discuss the role of water in blocking unfavorable binding sites and water-mediated interactions contributing to the stability of observable binding modes of maltotriose.



1. INTRODUCTION

Maltose-binding protein (MBP) is a periplasmic binding protein[1] responsible for transport of maltose and maltodextrins with up to seven glucose rings through the periplasmic space in *E. coli* and other Gram-negative bacteria.[2,3] The structure of MBP has been resolved in X-ray and NMR spectroscopy experiments both for the liganded (holo) and unliganded (apo) states.[2−12] These studies revealed that MBP has two structural domains connected by the hinge region with the carbohydrate binding site located in the cleft between the domains (Figure 1). The domains can perform large scale motion, with the closed state of MBP characterized by small interdomain angles, stabilized by the physiological ligands (such as maltose and maltotriose[11]) that bridge interactions between the domains.

Many experimental and theoretical studies were dedicated to investigation of ligand binding and large scale conformational changes in MBP.[3,11,13−17] Originally, two mechanisms were proposed to explain ligand binding to MBP. The induced fit mechanism[4,11,14,16] suggests that the ligand initially binds to the open state of MBP, which in turn triggers the conformation changes resulting in the protein domains closure. According to the conformational selection mechanism,[16] the ligand can bind with higher affinity to less populated semi-closed conformations. The NMR experiments confirmed the existence of apo-closed (or semi- closed) states of MBP,[9] but the recent single-molecule fluorescence energy transfer measurements ruled out the conformation selection mechanism of binding.[14]

As demonstrated in the X-ray experiments, the holo-open state of MBP can be stabilized by maltotetraitol,[12] a nonphysiological ligand. For maltotriose, the holo-open state of MBP likely exists only transiently (because ligand binding triggers domains closure), which complicates experimental analysis of the initial MBP−ligand interactions and resulting conformational changes. Thus, further clarification of the initial stages of ligand binding to MBP is needed. It is worth noting that while the NMR studies suggest that two binding modes of maltotriose may exist in solution[7] only one binding mode has been characterized so far by the X-ray experiments.[3,18]

The role of water-mediated interactions in protein−carbohydrate complex structures has been well recognized.[10,19,20] The experimental data suggest that hydration effects play an important role in ligand recognition and binding, as well as in triggering conformational transitions upon ligand binding to MBP.[4,12] Expulsion of water molecules from the (hydrophobic) binding cleft by a ligand can provide a favorable

Received: August 27, 2014
Published: December 29, 2014

DOI: 10.1021/ci500520q
J. Chem. Inf. Model. 2015, 55, 317−328

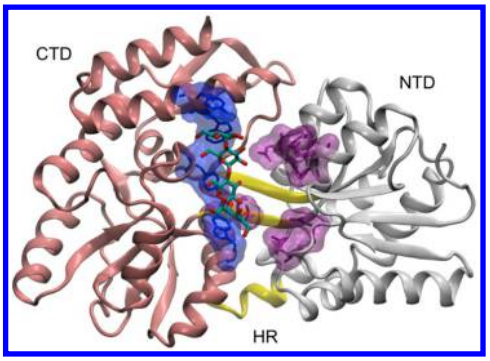


Figure 1. Holo-open structure of MBP with bound maltotetraitol (PDB accession code 1EZ9[12]). MBP is shown in cartoon representation with N-terminal domain (NTD) shown in gray, C-terminal domain (CTD) in pink, and the hinge region (HR) in yellow. Aromatic residues (Trp 230, Tyr 155, Trp 340, and Tyr 341) in the X-ray binding site are in blue, and polar residues (Asn 12, Asp 14, Glu 44, Glu 45, Asp 65, Arg 66, and Glu 111), which are implicated in protein–ligand interactions in the closed state of MBP, are shown in purple. Heavy atoms of maltotetraitol ligand are shown in stick representation and colored according to their atom types (oxygen in red and carbon in cyan).

entropic contribution to the free energy of binding and domain motion.[11,12,21] Conserved (structural) water molecules mediate the protein–ligand interactions through the network of hydrogen bonds[5,11,12] stabilizing the closed state of MBP, thus increasing the binding affinity. Water-mediated interactions could possibly contribute to initiating the domain closing as well.

In MBP, waters can occupy hydration sites around the hydrophilic residues in the ligand binding site close to the N-terminal domain (NTD) and form lower density pockets around the hydrophobic area at the C-terminal domain (CTD).[12] Water molecules that are tightly bound to the protein could efficiently block unfavorable ligand binding sites. Structural (tightly bound) water molecules play an important role in defining molecular recognition and mechanisms of protein–ligand binding.[22−24] Inability to properly describe such solvation effects is one of the most critical issues of many docking studies.[20,22,25,26]

Explicit solvent molecules can be included into a docking protocol by using different approaches. *In silico* docking experiments were performed with the crystallographic water[27] as well with water molecules predicted by different computational methods.[19,22,24,26,28−34] Molecular dynamics (MD) or Monte Carlo simulations with explicit solvent molecules are the most accurate approaches (within the limitations imposed by the choice of particular force field and by ignoring the quantum mechanical effects) to predict solvation structure around a receptor and to identify locations of tightly bound water molecules.[22,35] While such approaches can considerably improve the outcomes of the docking experiments,[19,22] they are computationally demanding and can be used only to model simple solvent conditions.[22,24] Many alternative computational methods have been developed in the past few years;[20,22−24,26,36] however, they do not provide sufficient information on solvation structure and thermodynamics and are not transferable to different solvent conditions/composition.

As an alternative to computationally demanding explicit solvent simulations, the 3D-RISM-KH molecular theory of solvation (three-dimensional reference interaction site model with the Kovalenko–Hirata closure relation)[37,38] can be used for fast and reliable prediction of the solvation structure around a macromolecule. The theory is based on the rigorous statistical–mechanical foundation and accounts for chemical specificity of solvent, including formation of hydrogen bonds between solute and solvent molecules, solvation entropic effects (hydrophobicity), and solution composition and thermodynamic state.[34,37−48] The 3D maps of water density distributions obtained with the 3D-RISM-KH molecular theory of solvation can be used to identify locations of structural water molecules for using in docking experiments.[49,50] This capability is implemented in the Molecular Operating Environment (MOE) software package.[51]

Recently, a new 3D-RISM-based protocol has been developed to predict the locations of water molecules around a protein.[49,50] In this study, we introduce a docking approach that combines this water placement protocol with the AutoDock Vina[52] docking capability. With this new approach, we investigate the binding modes of maltotriose to the open state of MBP and discuss the role of water in ligand binding and resulting conformation changes of MBP. We also compare the docking results obtained with this new approach with available experimental data as well as with outcomes of conventional docking experiments. We then discuss the role and functions of structural water in initial ligand binding and triggered large scale domain motions of MBP.

2. COMPUTATIONAL METHODS

In this study, we use the new protocol that employs the 3D-RISM-KH molecular theory of solvation[37,38] to predict locations of water molecules around the receptor, which then are used in docking experiments with the AutoDock Vina software.[52] In this section, we briefly overview the 3D-RISM-KH theory, following ref 37, and describe the protocols used to identify locations of water molecules and to predict binding modes of maltotriose at the open state of MBP.

2.1. 3D-RISM-KH Molecular Theory of Solvation. The solvation structure around a protein (or any other solute particle) can be described by the probability density $\rho_\gamma g_\gamma(\mathbf{r})$ of finding atomic site γ (oxygen or hydrogen for water) of a solvent molecule at 3D space position $\mathbf{r}$ around the solute. This density is determined by the average number density ρ_γ of bulk solvent multiplied by the 3D normalized density distribution function $g_\gamma(\mathbf{r})$. The latter describes the solvent site density enhancement when $g_\gamma(\mathbf{r}) > 1$ or depletion when $g_\gamma(\mathbf{r}) < 1$ relative to the average bulk density (where $g_\gamma(r) \to 1$).

The 3D distribution functions of solvent interaction sites are obtained from the 3D-RISM integral equation[37,38,53,54]

$$h_\gamma(\mathbf{r}) = \sum_\alpha \int d\mathbf{r}' c_\alpha(\mathbf{r} - \mathbf{r}')\chi_{\alpha\gamma}(r') \quad (1)$$

where $h_\gamma(\mathbf{r}) = g_\gamma(\mathbf{r}) - 1$ and $c_\gamma(\mathbf{r})$ are the 3D total and direct correlation functions of solvent, and the summation runs over all interaction sites of all solvent species. The direct correlation function is formally defined as a solution to eq 1. Outside the repulsive core, it has the asymptotics of the solute–solvent interaction potential $u_\gamma(\mathbf{r})$ scaled by the Boltzmann constant k_B times solvent temperature T: $c_\gamma(\mathbf{r}) \sim -u_\gamma(\mathbf{r})/(k_BT)$, where its value inside the repulsive core is related to the solvation free energy. The site–site susceptibility of solvent $\chi_{\alpha\gamma}(r)$ in eq 1 is an input to the 3D-RISM theory and is obtained from RISM theory[55] for solvent.

DOI: 10.1021/ci500520q
J. Chem. Inf. Model. 2015, 55, 317−328

To be solved, the 3D-RISM integral equation, eq 1, for the 3D total and direct correlation functions has to be complemented with another relation called a closure that also involves the interaction potential $u_\gamma(\mathbf{r})$ specified with a molecular force field. The exact closure has a nonlocal functional form that can be presented as an infinite diagrammatic series in terms of multiple integrals of the total correlation function.[55] However, it is computationally intractable and in practice is replaced with amenable approximations that should analytically ensure asymptotics of the correlation functions and features of the solvation structure and thermodynamics to properly represent the solvation physics.

We use the closure proposed by Kovalenko and Hirata (KH closure approximation).[37,38] It provides an accurate description of the solvation structure and thermodynamics in biomolecular systems as well as various association effects in complex liquids and electrolyte solutions.[34,37−48] The 3D KH closure that reads

$$g_\gamma(\mathbf{r}) = \begin{cases} \exp(-u_\gamma(\mathbf{r})/(k_B T) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})) & \text{for } g_\gamma(\mathbf{r}) \le 1 \\ 1 - u_\gamma(\mathbf{r})/(k_B T) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r}) & \text{for } g_\gamma(\mathbf{r}) > 1 \end{cases} \quad (2)$$

couples in a nontrivial way the hypernetted chain (HNC) and mean spherical approximation (MSA) closures.[55] The former is applied to the spatial regions of solvent density depletion $g_\gamma(\mathbf{r}) < 1$, including the repulsive core of the solute, and the latter to the regions of solvent density enrichment $g_\gamma(\mathbf{r}) > 1$, such as association peaks, while keeping the right asymptotics of $c_\gamma(\mathbf{r})$ peculiar in both HNC and MSA.[37,38]

The site−site susceptibility of pure solvent breaks up into the intra- and intermolecular terms

$$\chi_{\alpha\gamma}(r) = \omega_{\alpha\gamma}(r) + \rho_\alpha h_{\alpha\gamma}(r) \quad (3)$$

where the intramolecular correlation function $\omega_{\alpha\gamma}(r) = \delta_{\alpha\gamma}\delta(r) + (1 - \delta_{\alpha\gamma})\delta(r - l_{\alpha\gamma})/(4\pi l_{\alpha\gamma}^2)$ represents the geometry of solvent molecules with site−site separations $l_{\alpha\gamma}$ specified by the molecular force field, and $h_{\alpha\gamma}(r)$ is the radial total correlation function between solvent sites α and γ. Prior to 3D-RISM-KH calculations, $h_{\alpha\gamma}(r)$ values are obtained from the dielectrically consistent DRISM theory[56] with the KH closure (DRISM-KH approach)[38] for the solvent system (which may include ions, cosolvent, and ligands at a given concentration). To properly describe the electrostatic interactions in solution with polar molecular and ionic species, the electrostatic asymptotics of both the 3D and radial correlation functions are separated out and treated analytically[38,57] while calculating the convolution in the 3D-RISM-KH and DRISM-KH integral equations

2.2. Initial Structure Preparation for 3D-RISM-KH and Docking Simulations. The initial structures of the holo-closed (with maltotriose as a ligand), holo-open (MBP with bound maltotetraitol), and apo-open states of MBP were taken from the Protein Data Bank (PDB accession codes 3MBP,[11] 1EZ9,[12] and 1OMP,[4] respectively). All crystallographic water molecules were removed. The receptor structures for the bound systems were built by additionally deleting the maltotriose and maltotetraitol ligands from the corresponding X-ray structures. Target receptor conformations were parametrized, and all missing hydrogen atoms were added with the Amber 12 Molecular Dynamics Package[58] and the ff10 Amber force field.[59−61] The maltotriose and maltotetraitol structures were prepared with the GLYCAM Carbohydrate Builder.[62]

2.3. 3D-RISM-KH Calculations. 3D-RISM-KH calculations were performed in 0.00029 g/cm^3 sodium chloride aqueous solution at temperature 25 °C (≈298 K). The water density was set to 0.99983 g/cm^3 and the dielectric constant to 78.5. The solute−solvent interaction potential comprises the Coulomb and Lennard−Jones interactions. The potential was parametrized with the AMBER ff10 force field[59−61] for the protein and the SPC model for water molecules.[63]

The 3D-RISM-KH integral equations, eqs 1 and 2, were solved on a uniform 3D grid of 256 × 256 × 256 points in a cubic supercell of size 128 Å × 128 Å × 128 Å large enough to accommodate MBP along with sufficient solvation space around. The 3D-RISM-KH equations were converged to a relative root-mean-square accuracy of 10^{-5} by using the modified direct inversion in the iterative subspace (MDIIS) accelerated numerical solver.[38] It was proven that such protocol can provide hydration structure around proteins compatible with that obtained in explicit solvent MD simulations.[41]

2.4. Docking with Explicit Water: Water Placement Protocol. Water densities calculated with the 3D-RISM-KH molecular theory of solvation were used to predict locations of water molecules in and around the binding site of MBP by using the *Placevent* script of Sindhikara et al.[49]

The water placement protocol includes the following steps:

(1) Solute preparation: Receptor structures of MBP are prepared as described above.
(2) Solvent preparation: *rism1d* code with DRISM theory and KH closure[37] as implemented in the Amber Tools 12 package[64] is used to calculate solvent susceptibility from eq 1.
(3) Calculation of solvent densities with 3D-RISM-KH: Calculations are performed with the *3drism.snglpnt* code from the Amber Tools 12 package.[64]
(4) Water placement: The *Placevent* protocol[49] is used to predict location of water molecules around the receptor. Water oxygen positions were obtained based on the solvent density (which defines the volume occupied by one water molecule) of 0.54045 g/cm^3.
(5) Identification of displaceable water molecules: The hydropathy index,[65] which describes the hydrophilic and hydrophobic properties of 20 amino acid side-chains, is calculated for residues within 4 Å of water oxygen sites of all predicted water molecules. Water molecules with a total hydropathy index exceeding −1 are considered as displaceable and are excluded from docking experiments.
(6) Preparation of a receptor structure with explicit water for docking experiments: Coordinates of water molecules are merged with those of MBP, and resulting structures are prepared for docking experiments with AutoDockTools.[66]

2.5. Docking Protocol. Parameters and structures for docking experiments were set up with the AutoDockTools.[66,67] The Kollman's and the gasteiger charges were used for the receptor and ligand, respectively. Docking simulations were performed with AutoDock Vina[52] based on the standard docking protocol. The size of docking grid box was set to 30 Å × 30 Å × 30 Å in all docking experiments. The center of the box was near the side-chain nitrogen of Trp 340 for the apo-open and holo-open structures of MBP (this corresponds to the location with coordinates −5, −10, and 5 Å in the PDB ID 1OMP structure[4]). For the holo-closed structure, the grid box was centered near the side-chain oxygen of Glu 153 (box center coordinates of 12, 21, and 36 Å in the PDB ID 3MBP structure[11]). In AutoDock Vina simulations, the default value of docking exhaustiveness parameter of 8 was used, the adaptive

DOI: 10.1021/ci500520q
J. Chem. Inf. Model. 2015, 55, 317−328

number of steps in a run was set based on complexity of the problem. Binding modes were ranked with the standard AutoDock Vina scoring function[52] weight settings for each energy term. The five highest-ranked binding modes from each docking simulation were chosen for further analysis. These modes were first energy minimized and then rescored with the Amber 12 Molecular Dynamics Package[58] with Amber ff10 force field[59−61] for the receptor and GLYCAM_06h force field[68] for the ligands. In this procedure, the generalized Born implicit solvent model[69] was used. Minimization was performed with 2500 steepest descent steps followed by 2500 conjugate gradient minimization steps. Docking simulations were repeated three times for each system to check consistency of the results.

2.6. Molecular Dynamics Simulations. To test the dynamical stability of predicted bound conformations and to verify the possibility of triggering domains closure by ligand binding, molecular dynamics (MD) simulations were carried out for selected protein−ligand structures. The top-ranked binding modes from the current docking experiments as well as the apo-open structures with maltose transferred from the superimposed holo-closed structure were used as initial structures in these simulations. Note that explicit water molecules were excluded from the initial structures. The force field parameters and solvation model were the same as discussed in the previous section. In total, five 7.5 ns MD runs corresponding to the different initial velocity distributions were recorded and analyzed.

3. RESULTS AND DISCUSSION

We successfully redocked maltotriose into the experimental binding site of the receptor from the holo-closed state of MBP with the AutoDock Vina software by using a conventional docking protocol and without accounting for explicit water molecules (section S1, Supporting Information). This indicates that the AutoDock Vina scoring function[52] can provide an accurate description of the binding modes of maltotriose, at least for the native receptor conformations.

However, docking experiments performed with the same protocol for the open state of MBP were unsuccessful (see section S2, Supporting Information, for details). To provide an explanation for that, we first considered the following possibly limiting factors: (i) underestimation of the carbohydrate−protein π-stacking interactions by the AutoDock Vina scoring function, (ii) inaccurate representation of a local structure of the binding site by the apo-open experimental structure, and (iii) the lack of receptor flexibility in docking experiments. Relaxing the above limitations in docking experiments provides only marginal improvements of the results (section S3, Supporting Information).

Both the apo-open and holo-closed experimental structures of MBP have water molecules in the cleft between the protein's domains near the ligand binding site. At the same time, accounting for explicit water molecules in docking simulations with the closed state may be less crucial than for the open state of MBP. This can be explained by the desolvation of the polar residues at NTD during the domain closure triggered by ligand binding. In the holo-closed state, these residues form hydrogen bonds with the hydroxyl groups of sugar, which in turn forms π-stacking and hydrophobic interactions with the aromatic residues at CTD part of the binding site. Thus, the ligand bridges the interactions between domains, and desolvation of the polar residues occurs not during the initial binding event (which is mostly driven by the ligand interactions with the aromatic residues at CTD) but later on in the process of the MPB domain closing.

In the apo-state, the polar residues at NTD are fully solvent exposed (Figure 1) and can efficiently form hydrogen bonds with surrounding water molecules.[4] Such water molecules strongly bound to polar residues can provide steric restrictions on possible binding modes. If this effect is underestimated by a scoring function, the accuracy of docking may deteriorate. Other water molecules can be displaced from the binding sites, and some water molecules can bridge the interactions between the ligand and the receptor. The complex interplay between the solvation effects, including those due to the high desolvation penalty of structural waters and hydrogen bonding between water molecules and the receptor as well as ligand, may be difficult to accommodate by the phenomenological solvation part of the scoring function.[5,11,12] In some cases, these problems can be resolved by accounting for explicit water molecules in docking experiments.

It is worth noting that accounting for X-ray (structural) water may be not sufficient for improving the accuracy of docking experiments.[24] Thus, redocking of maltotetraitol to the native open X-ray structure of MBP[12] was not successful (section S2, Supporting Information). This can be possibly attributed to missing some of the structural water molecules in X-ray data due to the difficulties of identification of water at low resolution[24] and/or different hydration structure of the protein in solution and under conditions of X-ray crystallography.

There is much evidence that accounting for explicit (structural) water molecules can improve the accuracy of the docking experiments.[22−24,36] To explore the role of water in maltotriose binding to MBP and to assess accuracy of docking of maltotriose to the open state of MBP with explicit waters, we first predict the water distributions and location of water molecules around the MBP receptor with a newly developed protocol based on the 3D-RISM-KH molecular theory of solvation. Then, we use these water molecules to predict binding modes of maltotriose at the apo-open state of MBP and to analyze the role of water in the maltotriose interaction with MBP. To verify the consistency of the docking experiments, we employ the same protocol for redocking of maltotetraitol and maltotriose to the receptor structures from the X-ray holo-open and holo-closed structures of MBP.

3.1. Hydration Structure of Maltose-Binding Protein from the 3D-RISM-KH Molecular Theory of Solvation. The 3D-RISM-KH molecular theory of solvation[37,38] provides 3D maps of density distributions for each atomic site of solvent molecules around the macromolecule. High density peaks on these maps may correspond to tightly bound (structural) water molecules at the macromolecular surface. It was shown that there is a good agreement between the 3D-RISM-KH predictions on structural water and X-ray experimental data,[43] as well as between 3D-RISM-KH and MD results on solvation structure of proteins.[41,42]

We use the 3D-RISM-KH molecular theory of solvation to calculate 3D water density maps around the open state of MBP and to predict possible locations of tightly bound water molecules at the protein surface. These waters are used in docking experiments in the next section. For illustration, we show the high density peaks of water oxygen atomic site density distribution around the apo-open state of MBP in Figure 2. There is a good agreement between location of these peaks and location of X-ray waters (such as, for example, in the top right

DOI: 10.1021/ci500520q
J. Chem. Inf. Model. 2015, 55, 317−328

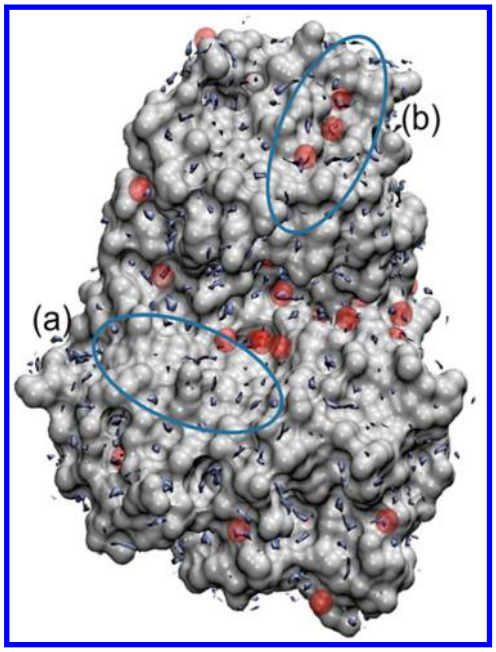


Figure 2. High density peaks of 3D water maps around the apo-open state of MBP predicted with the 3D-RISM-KH theory. The density of water oxygen is represented by the density isosurface with isolevel of 5 times the bulk water density (blue color). The X-ray waters from the apo-open state of MBP (PDB ID code 1OMP[4]) are shown with red van der Waals spheres of water oxygens. Solvent exposed surface of the protein is in gray. The region (a) corresponds to the experimental binding site of MBP. For illustration, the area with three X-ray water molecules overlapping with the high density peaks in water distribution is marked as (b).

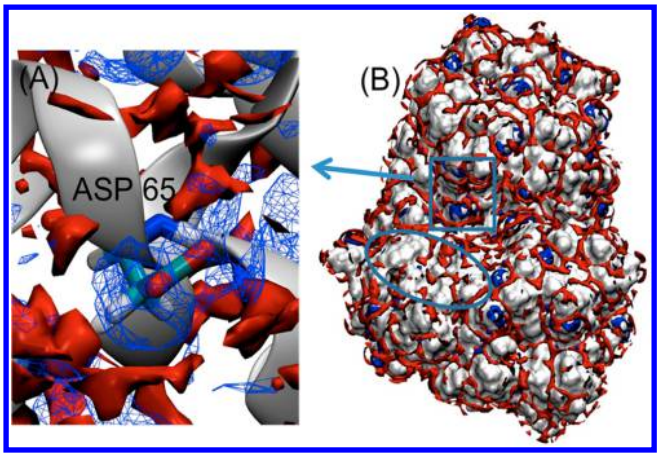


Figure 3. Water oxygen and hydrogen density distributions around the apo-open state of MBP predicted with the 3D-RISM-KH theory. The densities are shown by isosurfaces with isovalues of 4 for oxygen (red color) and 3 for hydrogen (blue color) times the bulk water density. Polarization of water is clearly seen around the polar residues. Panel A gives a zoomed view of the region around Asp 65 where a hydrogen bond forming water molecule was identified in X-ray experiments.[4,11] The water density polarization, which may be indicative of strong bonding between water and the protein, is clearly seen at the N-terminal domain in panel B above the maltotriose putative binding site (marked with the blue ellipse).

corner of the figure). There is a region with enhanced water density at the N-terminal domain right above the binding site marked by (a) in Figure 2.

It has been shown recently[22,24] that tightly bound water molecules that are not displaceable upon ligand binding are likely to form hydrogen bonding with a protein. The 3D-RISM-KH theory accounts for the specificity of solute−solvent and solvent−solvent interactions, including hydrogen bonding, by coupling the correlation functions of different solvent sites and by using site-specific parameters of the force field. In particular, in case of hydration, the formation of hydrogen bonds between water and a protein will result in polarization of distributions for water oxygen and hydrogen atomic sites, accounting for directionality of hydrogen bonding. Such polarization of the water site densities are clearly seen around the polar residues of MBP (Figure 3). For illustration, we show in panel A of Figure 3 a zoomed view of the region around Asp 65 where a hydrogen bond forming water molecule was identified by the X-ray experiments.[4,11] Other high density peaks of water oxygen and hydrogen characterized by the high degree of polarization can be found around the other polar residues of MBP, including Asn 12, Asp 14, Lys 15, Glu 44, Glu 45, Asp 65, Arg66, Gln 72, Glu 111, Glu 153, Ser 337, and Arg 344. Many of these residues are located in the cleft between CTD and NTD domains (Figure 3) above the maltotriose binding site at the C-terminal domain. The water density polarization at these locations may indicate strong (hydrogen) bonding between water and the protein. At the same time, the water densities in and around the CTD binding site lined with the aromatic residues (Tyr 155, Trp 230, Trp 340, and Tyr 341) show less polarization. This can be related to the lower desolvation penalty of removing water molecules from the mostly hydrophobic interface. These predictions are in qualitative agreement with the results of the free energy analysis of water molecules in binding sites of proteins which can be or cannot be displaced by ligand binding.[43,70]

3.2. Prediction of Structural Water in the Binding Site of the Open State of MBP. The 3D-RISM-KH solvent densities can be directly incorporated into the docking protocol to describe the desolvation effects upon ligand binding.[71] Also, locations of tightly bound water molecules can be predicted based on the 3D-RISM-KH densities,[49−51] and these explicit water molecules can be used in conventional docking simulations. In this section, we find locations of explicit water molecules at the apo-open state of MBP by using the protocol described in Computational Methods section.

The predicted water molecules in proximity of the (putative) binding site of MBP along with the 3D-RISM-KH water oxygen density distributions are shown in Figure 4. The locations of these water molecules coincide well with the high-density peaks of the water distributions (e.g., water oxygen site density distribution). For example, the high density peaks corresponding to the density isolevel of 4 overlaps with most of the predicted water molecules (panel A, Figure 4). Some of these water molecules will be displaced by the ligand to optimize the solvation entropic part of the binding free energy as well as the binding enthalpy. More tightly bound water molecules have higher binding affinity, and their locations may be associated with the higher peaks in the density distributions. Interestingly, these high density peaks are almost excluded from the experimental binding site of the apo-open state of MBP (panel B, Figure 4), but they persist above the binding site at the location of an alternative ligand binding site at NTD predicted by docking without explicit water molecules (Figure S2, Supporting Information).

DOI: 10.1021/ci500520q
J. Chem. Inf. Model. 2015, 55, 317−328

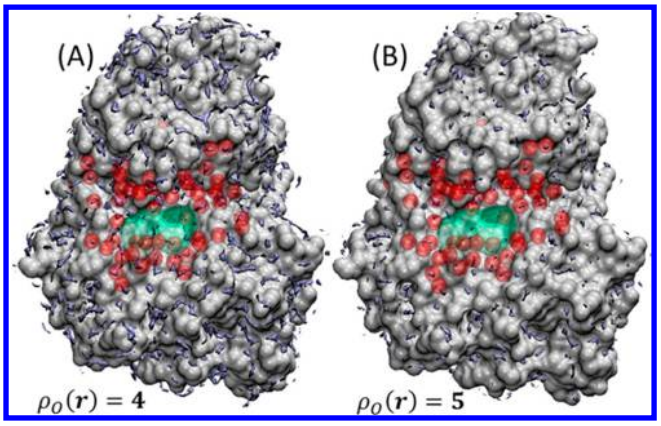


Figure 4. Water oxygen density distributions (shown in blue) around the apo-open state of MBP obtained with the 3D-RISM-KH theory. Structural water molecules predicted by the 3D-RISM-KH theory as discussed in section 3.3 are shown with red van der Waals spheres. Solvent exposed surface of the protein is in gray. The density maps are shown for isolevel of 4 (panel A) and 5 (panel B) in excess of the water bulk density. The ligand binding site is shown by the green surface.

There are approximately 20 water molecules at the interface between CTD and NTD of MBP. The distances between the protein heavy atoms with a substantial partial charge and water oxygen are between 3.0 and 4.0 Å. This further supports the fact that these water molecules can form a network of hydrogen bonds with MBP and can bridge the residues at CTD and NTD. For example, a hydrogen-bonding chain can form between NTD Trp 340 and CTD Asp 65, Arg 66 and Trp 62, with one water molecule bridging Trp 340 and Asp 65, and the other two bridging Asp 65 with Arg 66 and Trp 62. Another two water molecules can bridge CTD Ser 337 with NTD Arg 66 and Glu 45. Similar water mediated interactions can be seen between the polar residues from two domains at the edge of the experimental CTD binding pocket in the holo-open and holo-closed X-ray structures.[11,12] We note that the distances between the predicted locations of water oxygen at the edge of the experimental binding site can be as high as 11 Å (such as the distance between oxygen atoms of water molecules located near Asp 153 at NTD and Asp 14 at CTD).

The 3D-RISM-KH-based protocol for water placement also predicts water molecules near the aromatic residues (Tyr 155, Trp 230, and Trp 340) all within 4 Å distance from the maltose binding site at CTD. This agrees well with the results of X-ray experiments.[12] The shortest distance between the aromatic residue and water molecule is smaller than the distances between the residues and maltotetraitol in the experimental holo-open structure.[11] This can be explained by a smaller size of water molecules compared to the size of carbohydrate ligands, which allows them to access the protein surface easier.

3.3. Identification of Displaceable Water Molecules. Water molecules in and around a ligand binding site can be affected by binding. This includes both displacement of weakly bound water molecules by a ligand as well as rearrangement of solvation structure around the ligand to optimize protein−solvent and ligand solvent interactions. Ideally, docking experiments should account for such events, which in most cases currently are taken in consideration at the phenomenological level by introducing, for example, desolvation/hydrophobic terms in a scoring function, or by accounting for explicit structural water molecule in docking experiments. In the latter case, only tightly bound water molecules that cannot be displaced by a ligand should be kept in docking experiments.

Water molecules identified as described above have different binding affinities, depending on their local environments (side-chain orientations, partial charge distributions of the atomic sites of the receptor, etc). Some water molecules can be displaced by the ligand upon binding, while other may provide steric obstruction to binding or/and mediate favorable protein ligand interactions. Removing displaceable water molecules in docking with explicit water molecules can sufficiently improve the accuracy of docking results.[22,24,36] This can be done, for example, by ranking water molecules binding affinities obtained with the free-energy perturbation calculations[36] or by using the thermodynamic analysis with different solvation models based on explicit solvent MD simulations.[72]

While such free energy calculations can provide useful information on binding affinities of water molecules in and around a binding site, in many practical applications more computationally efficient and fast approaches are needed. In this study, we use a protocol based on the hydrophobicity scale[65] of residues in proximity of a putative binding site. The hydropathy index describes the hydrophobic/hydrophilic property of each amino acid side-chain.[65] Larger (positive) values correspond to residues characterized by more pronounced hydrophobic properties. Water molecules located in hydrophobic environments are easy to displace compared to those located around the hydrophilic residues. Thus, we assume, following refs 22 and 65, that the hydropathy index of a residue can be related to binding affinity (or desolvation penalty) of water molecules near that residue.

By summing up the hydropathy indices for all residues within 4 Å distances of the oxygen water atomic sites (Table S1, Supporting Information), we predicted the desolvation penalty of the water molecules in and around MBP binding site. After that, water molecules with a total hydropathy index larger than −1 were deleted because they are located in mostly hydrophobic environments and thus more likely to be displaced by a carbohydrate ligand. A similar procedure has been used recently as a part of the protocol for scoring water molecules in a protein binding sites for docking experiments.[22] The choice of a threshold for selecting displaceable water molecules can be justified by the fact that residues with −SH/−OH functional groups (such as Thr, Ser, and Tyr), with polarity similar to that of water molecules, have hydropathy index in the range from −0.7 to −0.9.[65] Increasing the threshold value (for example, up to 0) results in placement of water molecules near the aromatic residues in the binding site.

We note that water binding affinities can also be obtained from the 3D-RISM-KH theory coupled to the free energy calculations. While such an approach can provide more accurate estimates of binding affinities, it is more computationally demanding.

3.4. Docking Experiments with Explicit Water Molecules Predicted by 3D-RISM-KH: Maltotriose Binding to the Open State of MBP. Docking experiments without accounting for explicit water in hydration sites of MBP predict that maltotriose binding modes mostly located at the N-terminal domain of the MBP open state (see Supporting Information for details), which are not confirmed experimentally. Both the docking simulations with different MBP conformers and the structural analysis of the open structure of MBP suggest that the polar residues in and around the NTD binding site can be surrounded by tightly bound water

DOI: 10.1021/ci500520q
J. Chem. Inf. Model. 2015, 55, 317−328

molecules. These water molecules can either block unfavorable binding modes or provide favorable water-mediated interactions between the protein and the ligand. In this section, we discuss results of docking experiments with maltotriose at the open state of MBP, with the explicit (possibly tightly bound) water molecules predicted with the 3D-RISM-KH molecular theory of solvation.

Docking simulations were carried out with AutoDock Vina software.[52] The five highest-ranked binding modes were chosen for further analysis. All these modes are located close to the X-ray binding site of MBP at CTD. The top-ranked binding mode obtained after rescoring (see the docking procedure explanation in the Computational Methods section) is shown in Figure 5

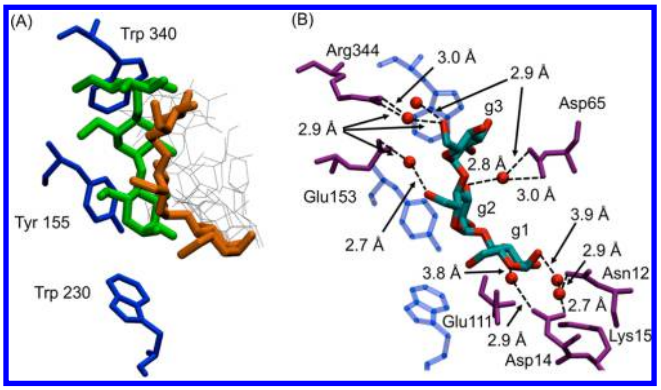


Figure 5. Binding modes of maltotriose at the open state of MBP from docking experiments with explicit water molecules predicted by the 3D-RISM-KH molecular theory of solvation. (A) Top-ranked maltotriose conformation from docking experiments (orange) compared to the conformation of maltotriose from the superimposed holo-closed X-ray structure[11] (green). The next four low-ranked binding modes are shown in gray. (B) Predicted water-mediated hydrogen bonds (dashed lines) between the top-ranked binding mode and the polar residues of MBP. The polar and aromatic residues are shown in purple and blue, respectively. Water oxygens are represented by red van der Waals spheres.

along with the experimental position of the maltotriose ligand.[11] It is worth noting that the improvement of docking results were achieved by accounting for the tightly bound, while removing potentially displaceable water molecules from the hydrophobic areas. Many of these potentially displaceable water molecules are located around the aromatic residues in the experimentally observable binding site (Figure 4B). By displacing these water molecules, the ligand can form favorable π-stacking interactions with the aromatic residues of the receptor (Figure 5). This also provides favorable entropic contribution due to displacement of these water molecules from mostly hydrophobic environments to the bulk solution.[12] We note that keeping these weakly bound water molecules located around the aromatic residues along with the structural waters in the docking experiments does not improve the accuracy of the docking (section S5, Supporting Information). The structural water molecules predicted with the 3D-RISM-KH theory provide steric obstruction thus directing the ligand in a proper location in the CTD binding cleft, as well as mediate favorable interactions between the hydroxyl groups of sugar and the polar residues located in proximity of the binding site (Figure 5B).

Interestingly, two binding modes of maltotriose at MBP were detected in the previous solution NMR and fluorescence spectroscopy studies.[7,73,74] The R mode (named after the red shift in fluorescence emission spectrum) is consistent with the X-ray binding mode,[5,18] where the ligand g1 sugar ring (Figure 5B for the explanation of notations) forms π-stacking with Tyr 155 and Trp 230, g2 ring with Trp 340, and g3 ring with Tyr 341. Our predicted binding mode of maltotriose at the open state of MBP is similar to the experimental B mode (after the blue shift in fluorescence emission spectrum), where the g1 sugar ring moves away from Trp 230[7] and closer to NTD. In the docked structure, the g2 sugar ring can form π-stacking interactions with Tyr 155 and g3 with Trp 340 (Figure 5). Due to the different Tyr 341 side-chain rotamer state in the apo-open state of MBP compared to the holo-closed state, no π-stacking interactions are formed with the g3 sugar ring. As discussed in the next section, such interactions are observed transiently in MD simulations starting from the docked conformation. Because both the R and B modes are observed experimentally (with the spectroscopic studies suggesting predominance of the R mode[18]), our docking results can provide valid predictions on the maltoriose binding modes for the MBP open state as well as on mechanisms of binding and triggered domains motion.

As discussed above, the water molecules not only provide a steric guidance for the ligand, but they also mediate favorable interactions between the ligand and the receptor. Thus, the g1 sugar ring of maltotriose forms water-mediated hydrogen bonds with the polar residues at NTD, which includes Asp 14, Asn 12, and Lys 15 (Figure 5). There is also a possibility of bridging the interactions between the g2 ring and Asp 65 at NTD by a water molecule. Similarly, the water molecules bridge the interactions between the polar residues at CTD (two water molecules form hydrogen bonds between the C6-OH group of the g3 ring of the ligand and Arg 344, and one water molecule forms a hydrogen bond between the C6-OH group of g2 and Glu 153). These water molecules are located close to the aromatic residues in the binding site, and they can push the ligand away from the CTD binding site toward to NTD. This could explain the difference between the experimental (from the superimposed holo-closed structure) and the predicted binding mode of maltotriose.

We also note that the ligand transferred from the superimposed holo-closed X-ray structure does not form hydrogen bonding interactions with any polar residue at NTD. The shortest distance between the ligand's hydroxyl groups and the NTD/HR polar residues is 5.4 Å (distance between the C3-OH group of the g2 sugar ring and Asp 65). In the holo-open X-ray structure with maltotetraitol, there are only three hydrogen bonds between the ligand and NTD with one bond formed between the linear sugar of maltotetraitol and Asp 14.[12] The interactions between the ligand and NTD for the predicted binding mode could provide more favorable conditions for triggering domain closure. This is partially confirmed by the results of MD simulations discussed in the next section.

To confirm the consistency of the docking experiments with explicit water predicted by the 3D-RISM-KH molecular theory of solvation, we performed redocking of native ligands to the receptor X-ray structures of the holo-open (MBP with bound maltotetraitol, PDB accession code 1EZ9[12]) and holo-closed (MBP with bound maltotriose, PDB accession codes 3MBP[11]) states of MBP. In both cases, the locations of possible tighly bound water molecules were indentified with the protocol discussed in section 3.2. In the case of the holo-open structure,

DOI: 10.1021/ci500520q
J. Chem. Inf. Model. 2015, 55, 317−328

there is a significant improvement in the ligand's placement compared to the docking without accounting for explicit solvent (section S2, Supporting Information). The top-ranked binding modes of both maltotriose and maltotetraitol are located in the experimental binding site, in good agreement with the available experimental data[12] (Figure 6).

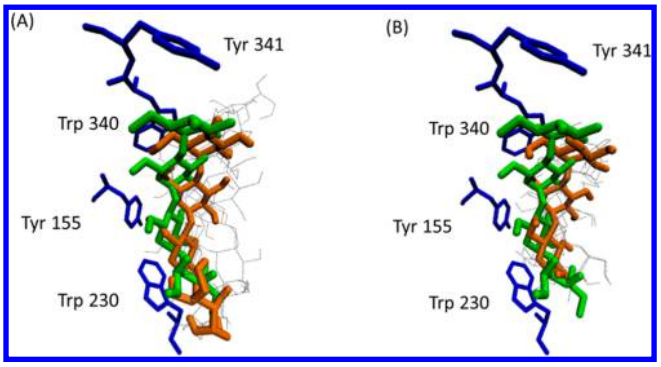


Figure 6. (A) Binding modes of maltotetraitol at the open state of MBP from docking experiments with explicit water molecules predicted by the 3D-RISM-KH molecular theory of solvation. The top-ranked ligand conformation from docking experiments (orange) is compared to the experimental structure (PDB accession code 1EZ9[12]) (green). The next four low-ranked binding modes are shown in gray. (B) Binding modes of maltotriose at the open state of MBP (receptor structure for docking was modeled after the X-ray holo-open structure of MBP, PDB accession code 1EZ9[12]). The notations are the same as in panel A.

We also predicted locations of tightly bound water molecules around the closed structure of MBP and carried out docking experiments with maltotriose (redocking experiments without account for explicit water molecules are described in section S1, Supporting Information). Interestingly, in this case, there are three water molecules located near the polar residues (Glu153 from CTD and Arg 66, Glu 44, and Glu 111 from NTD of MBP), which block the binding site (Figure 7). These water molecules are located in mostly hydrophilic environments, and thus, they are not listed as displaceable by the protocol based on the hydropathy index discussed in section 3.3. Two of these three water molecules are located at NTD, and thus, they are brought to the proximity of the binding site during the domain closure.

It was shown in the NMR experiments that 5% of apo MBP conformations correspond to the partially closed state and 95% to the open state.[9] At the same time, the recent single-molecule fluorescence resonance energy transfer experiments suggested that the ligand binds to the open state of MBP only, which triggers consequent closure of the protein domains.[14] The presence of tightly bound water molecules near/in the binding site in the closed state of MBP can explain the above-mentioned experimental finding. In this scenario, these water molecules are brought to the binding site of the apo-closed structure in the process of domain closure, and thus, they do not obstruct the binding site in the open state, as observed in docking experiments with the open structure of MBP. In the transition between the holo-open to holo-closed state, desolvation of the polar residues in the proximity of the binding site is driven by the domain closure, as discussed above.

3.5. Verification of Docking Predictions with MD Simulations. To verify our docking results and to test dynamical stability of the predicted maltotriose binding modes, we carried out MD simulations starting from different conformations of the maltotriose−MBP complex corresponding to (i) the top-ranked binding mode of maltotriose docked to the apo-open structure of MBP obtained with the new docking protocol, (ii) the X-ray maltotriose conformations transferred from the MBP holo-closed state after superimposition of the closed and open structures, and (iii) the NTD binding mode predicted by the conventional docking protocol without accounting for explicit water molecules (section S2, Supporting Information). For each system, MD simulations were repeated five times with different initial conditions, as explained in the Computational Methods section.

No significant interdomain angle changes were observed in four of out of five MD runs started from the NTD binding mode (Figure S3, Supporting Information). It was only in one MD run that domain closure was observed after the ligand moved toward the experimental binding site at CTD.

In the MD simulations with the CTD top-ranked binding mode obtained with the explicit water molecules shown in Figure 5, maltotriose quickly (at the time scale of simulations) moves toward the CTD aromatic residues to optimize the

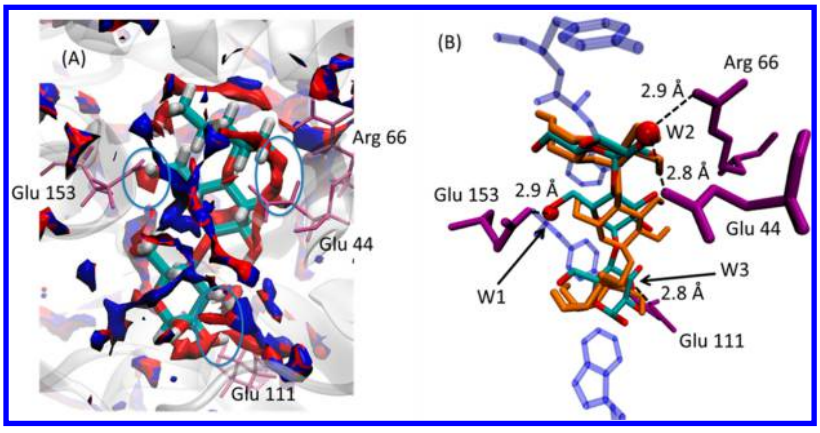


Figure 7. (A) Water oxygen density distributions around the maltotriose binding site in holo-closed (blue) and the apo-closed (red) states of MBP, as predicted with the 3D-RISM-KH theory. Density isosurfaces at isovalue 4. (Some −OH groups of maltotriose in the binding site overlap (in blue ellipse areas) with the high peaks of the water oxygen density distribution in the MBP unliganded state.) (B) X-ray structure of the holo-closed state of MBP. Water molecules blocking the binding site as predicted with the water placement protocol for the apo-closed structure are shown as red van der Waals spheres. Polar and aromatic residues are shown in purple and blue, respectively.

DOI: 10.1021/ci500520q
J. Chem. Inf. Model. 2015, 55, 317−328

aromatic and hydrophobic interactions. This movement and the following domain closure take place within first 2 ns of MD runs (Figure 8). Interestingly, the domain closure can occur

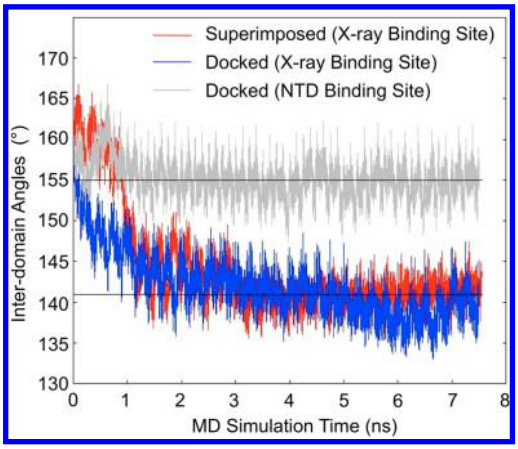


Figure 8. Evolution of the interdomain angle in MD simulations started from different initial structures of the maltotriose–MBP complex. The results are shown for the top-ranked maltotriose bound conformation at the CTD binding site (blue), false NTD binding site (gray), and ligand conformation from the superimposed X-ray closed structure (red).

faster for the predicted bound conformation compared to the case of the experimental ligand conformation from the superimposed holo-closed X-ray structure (Figure 8). After the domains get closed, maltotriose stays in the experimentally observable binding cleft corresponding to the holo-closed state of MBP, with the conformations of maltotriose resembling the experimentally observable structure from the holo-closed state of MBP. For illustration, we show in Figure 9 a representative snapshot from the MD trajectory started with the maltotriose top-ranked binding mode. There is excellent agreement between the X-ray structure of the complex[11] and our results

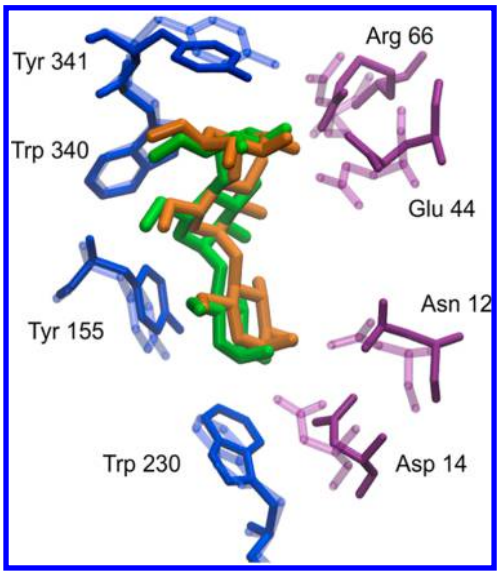


Figure 9. Snapshot of MD simulation started from the top-ranked maltotriose conformation from docking experiments with explicit water (orange) compared to the conformation of maltotriose in the holo-closed X-ray structure (green).[11] Aromatic residues in the binding site are shown in blue, and polar residues are shown in purple.

in which all the major interactions (hydrophobic, aromatic, and hydrogen bonding) between the receptor and the ligand are established. Thus, the docking protocol that uses the 3D-RISM-KH assignment of possible locations of structural water molecules provides a more accurate description of the binding modes of MBP compared to the conventional docking methods.

Because the initial structure for the MD simulations obtained in docking experiments with the explicit water resembles the maltotriose NMR B mode,[7,73,74] a plausible scenario for maltotriose binding to the open state of MBP can be as follows: First, maltotriose binds to the CTD binding site as shown in Figure 5, and then, in the course of the conformational changes triggered by binding (transition from the open to closed state), the ligand acquires experimentally observable conformation corresponding to the holo-closed state of MBP (Figure 9). This provides a new prospective on the mechanisms of induced fit-type binding of maltotriose to MBP.

4. CONCLUSIONS

The 3D-RISM-KH molecular theory of solvation[37,38] accurately and efficiently predicts the solvation structure around MBP, including locations of high density water pockets. This information can be used to place explicit water molecules in most probable hydration sites by using the protocol recently developed by Sindhikara et al.[49,50] Some of these water molecules can be displaced by ligand binding, especially from the hydrophobic interfaces, while the others characterized by the high desolvation penalty can impose steric restrictions on possible binding modes. The role of water molecules in ligands binding to MBP includes partial blocking of unfavorable binding sites, favorable hydrophobic interactions, including those with the aromatic residues at CTD, as well as mediation of the interactions between the polar residues at NTD and the ligand. The latter may be an important factor in the domains closure process triggered by ligand binding.

The current results also demonstrate the importance of accounting for structural water in docking experiments as discussed in many previous studies.[19,20,25,33] Importantly, the 3D-RISM-KH molecular theory of solvation constitutes a fast and accurate procedure for predicting locations of structural water molecules[41,50] at the protein surface, which can be used in docking experiments.

The new docking approach that combines the 3D-RISM-KH molecular theory of solvation to obtain solvent density distributions around a biomolecule and a protocol for prediction locations of explicit water molecules (such as the *Placevent* algorithm[49] used in the current study) with AutoDock Vina software can increase the accuracy of binding modes prediction, as we demonstrate for the binding modes of maltotriose at the open state of MBP. We also show that this approach provides a new insight into the conformational fitting mechanism of ligand binding, including binding-triggered large-scale conformation changes in MBP. In particular, we describe an alternative binding mode of maltotriose at the open state of MBP, the existence of which was previously suggested based on NMR spectroscopic experiments.[7] We show that such a binding event results in triggering the domains motion of MBP and evolution of the structure toward the holo-closed state, which is in excellent agreement with the experimental X-ray structure.[11]

Validating the ability of this new approach to provide systematic improvement of docking results for a general pool of biomolecular systems where structural solvation plays an

DOI: 10.1021/ci500520q
J. Chem. Inf. Model. 2015, 55, 317–328

important role requires large scale simulations with multiple crystal structures as well as assessment of different protocols for selections of displaceable water molecules and possible refinement of the current protocol parameters (such as the threshold value for hydropathy index). More advanced methods for ranking binding affinity of structural water molecules based, for example, on free energy calculations can be used instead of the hydropathy index-based approach of the current study, which will provide transferability of the proposed protocol and its applicability to systems with complex solvation environments.

■ ASSOCIATED CONTENT

Supporting Information

Results of redocking of maltotriose to the closed state of the maltose-binding protein (positive control docking experiments), docking of maltotriose to the open state of the maltose-binding protein (MBP) without accounting for structural water molecules, flexible docking and docking with alternative conformations of residues in the binding site, docking of maltotriose to the open state of MBP with X-ray crystallographic water, and docking of maltotriose to the open state of MBP with water molecules predicted by the 3D-RISM-KH molecular theory of solvation. Table S1 provides information on the hydropathy index used to identify displaceable (by the ligand) water molecules in and around the binding site of MBP. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: andriy.kovalenko@nrc-cnrc.gc.ca.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the financial support from the Bill and Melinda Gates Foundation Grand Challenges Explorations Grant, University of Alberta, and National Institute for Nanotechnology. The high performance computing resources were provided by the WestGrid − Compute Canada national advanced computing platform.

■ REFERENCES

(1) Oliver, D. B. Periplasm. In *Escherichia Coli and Salmonella: Cellular and Molecular Biology*; Neidhardt, F. C., Ed.; ASM Press: Washington, DC, 1996; Vol. *1*, pp 88−100.

(2) Sharff, A. J.; Rodseth, L. E.; Quiocho, F. A. Refined 1.8-Å Structure Reveals the Mode of Binding of Beta-Cyclodextrin to the Maltodextrin Binding Brotein. *Biochemistry* **1993**, *32*, 10553−10559.

(3) Schäfer, K.; Magnusson, U.; Scheffel, F.; Schiefner, A.; Sandgren, M. O. J.; Diederichs, K.; Welte, W.; Hülsmann, A.; Schneider, E.; Mowbray, S. L. X-ray Structures of the Maltose-Maltodextrin-Binding Protein of the Thermoacidophilic Bacterium *Alicyclobacillus acidocaldarius* Provide Insight into Acid Stability of Proteins. *J. Mol. Biol.* **2004**, *335*, 261−274.

(4) Sharff, A. J.; Rodseth, L. E.; Spurlino, J. C.; Quiocho, F. A. Crystallographic Evidence of a Large Ligand-Induced Hinge-Twist Motion Between the Two Domains of the Maltodextrin Binding Protein Involved in Active Transport and Chemotaxis. *Biochemistry* **1992**, *31*, 10657−10663.

(5) Evenäs, J.; Tugarinov, V.; Skrynnikov, N. R.; Goto, N. K.; Muhandiram, R.; Kay, L. E. Ligand-Induced Structural Changes to Maltodextrin-Binding Protein as Studied by Solution NMR Spectroscopy. *J. Mol. Biol.* **2001**, *309*, 961−974.

(6) Rubin, S. M.; Lee, S. Y.; Ruiz, E. J.; Pines, A.; Wemmer, D. E. Detection and Characterization of Xenon-Binding Sites in Proteins by 129Xe NMR Spectroscopy. *J. Mol. Biol.* **2002**, *322*, 425−440.

(7) Gehring, K.; Williams, P. G.; Pelton, J. G.; Morimoto, H.; Wemmer, D. E. Tritium NMR Spectroscopy of Ligand Binding to Maltose-Binding Protein. *Biochemistry* **1991**, *30*, 5524−5531.

(8) Xu, Y.; Zheng, Y.; Fan, J.-S.; Yang, D. A New Strategy for Structure Determination of Large Proteins in Solution Without Deuteration. *Nat. Methods* **2006**, *3*, 931−937.

(9) Tang, C.; Schwieters, C. D.; Clore, G. M. Open-to-Closed Transition in Apo Maltose-Binding Protein Observed by Paramagnetic NMR. *Nature* **2007**, *449*, 1078−1082.

(10) Quiocho, F. A. Carbohydrate-Binding Proteins: Tertiary Structures and Protein-Sugar Interactions. *Annu. Rev. Biochem.* **1986**, *55*, 287−315.

(11) Quiocho, F. A.; Spurlino, J. C.; Rodseth, L. E. Extensive Features of Tight Oligosaccharide Binding Revealed in High-Resolution Structures of the Maltodextrin Transport/Chemosensory Receptor. *Structure* **1997**, *5*, 997−1015.

(12) Duan, X.; Quiocho, F. A. Structrual Evidence for a Dominant Role of Nonpolar Interactions in the Binding of a Transport/Chemosensory Receptor to its Highly Polar Ligands. *Biochemistry* **2002**, *41*, 706−712.

(13) Millet, O.; Hudson, R. P.; Kay, L. E. The Energetic Cost of Domain Reorientation in Maltose-Binding Protein as Studied by NMR and Fluorescence Spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12700−12705.

(14) Kim, E.; Lee, S.; Jeon, A.; Choi, J. M.; Lee, H.-S.; Hohng, S.; Kim, H.-S. A Single-Molecule Dissection of Ligand Binding to a Protein with Intrinsic Dynamics. *Nat. Chem. Biol.* **2013**, *9*, 313−318.

(15) Kondo, H. X.; Okimoto, N.; Morimoto, G.; Taiji, M. Free-Energy Landscapes of Protein Domain Movements upon Ligand Binding. *J. Phys, Chem. B* **2011**, *115*, 7629−7636.

(16) Bucher, D.; Grant, B. J.; McCammon, J. A. Induced Fit or Conformational Selection? The Role of the Semi-Closed State in the Maltose Binding Protein. *Biochemistry* **2011**, *50*, 10530−10539.

(17) Stockner, T.; Vogel, H. J.; Tieleman, D. P.; Salt-Bridge Motif, A. Involved in Ligand Binding and Large-Scale Domain Motions of the Maltose-Binding Protein. *Biophys. J.* **2005**, *89*, 3362−3371.

(18) Duan, X.; Hall, J. A.; Nikaido, H.; Quiocho, F. A. Crystal Structure of the Maltodextrin/Maltose-Binding Protein Complexed with Reduced Oligosaccharides: Flexibility of Tertiary Structure and Ligand Binding. *J. Mol. Biol.* **2001**, *306*, 1115−1126.

(19) Gauto, D. F.; Petruk, A. A.; Modenutti, C. P.; Blanco, J. I.; Di Lella, S.; Marti, M. A. Solvent Structure Improves Docking Prediction in Lectin−Carbohydrate Complexes. *Glycobiology* **2012**, *23*, 241−258.

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

(21) Thomson, J.; Liu, Y.; Sturtevant, J. M.; Quiocho, F. A. A Thermodynamic Study of the Binding of Linear and Cyclic Oligosaccharides to the Maltodextrin-Binding Protein of *Escherichia Coli*. *Biophys. Chem.* **1998**, *70*, 101−108.

(22) Ross, G. A.; Morris, G. M.; Biggin, P. C. Rapid and Accurate Prediction and Scoring of Water Molecules in Protein Binding Sites. *PLoS One* **2012**, *7*, e32036.

(23) Roberts, B. C.; Mancera, R. L. Ligand−Protein Docking with Water Molecules. *J. Chem. Inf. Model.* **2008**, *48*, 397−408.

(24) Kumar, A.; Zhang, K. Y. J. Investigation on the Effect of Key Water Molecules on Docking Performance in CSARdock Exercise. *J. Chem. Inf. Model.* **2013**, *53*, 1880−1892.

(25) Moitessier, N.; Englebienne, P.; Lee, D.; Lawandi, J.; Corbeil, C. R. Towards the Development of Universal, Fast and Highly Accurate Docking/Scoring Methods: A Long Way to Go. *Br. J. Pharmacol.* **2008**, *153* (Suppl), S7−S26.

(26) Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Prediction of the Water Content in Protein Binding Sites. *J. Phys, Chem. B* **2009**, *113*, 13337−13346.

DOI: 10.1021/ci500520q
J. Chem. Inf. Model. 2015, 55, 317−328

(27) Huang, N.; Shoichet, B. K. Exploiting Ordered Waters in Molecular Docking. *J. Med. Chem.* **2008**, *51*, 4862−4865.
(28) Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **1985**, *28*, 849−857.
(29) Miranker, A.; Karplus, M. Functionality Maps of Binding Sites: A Multiple Copy Simultaneous Search Method. *Proteins: Struct., Funct., Genet.* **1991**, *11*, 29−34.
(30) Pitt, W. R.; Goodfellow, J. M. Modelling of Solvent Positions Around Polar Groups in Proteins. *Protein Eng.* **1991**, *4*, 531−537.
(31) Verdonk, M. L.; Cole, J. C.; Taylor, R. SuperStar: A Knowledge-Based Approach for Identifying Interaction Sites in Proteins. *J. Mol. Biol.* **1999**, *289*, 1093−1108.
(32) Kortvelyesi, T.; Dennis, S.; Silberstein, M.; Brown, L., III; Vajda, S. Algorithms for Computational Solvent Mapping of Proteins. *Proteins: Struct., Funct., Bioinf.* **2003**, *51*, 340−351.
(33) Barelier, S.; Boyce, S. E.; Fish, I.; Fischer, M.; Goodin, D. B.; Shoichet, B. K. Roles for Ordered and Bulk Solvent in Ligand Recognition and Docking in Two Related Cavities. *PLoS One* **2013**, *8*, e69153.
(34) Imai, T.; Harano, Y.; Kinoshita, M.; Kovalenko, A.; Hirata, F. Theoretical Analysis on Changes in Thermodynamic Quantities Upon Protein Folding: Essential Role of Hydration. *J. Chem. Phys.* **2007**, *126*, 225102.
(35) Henchman, R. H.; McCammon, J. A. Extracting Hydration Sites Around Proteins from Explicit Water Simulations. *J. Comput. Chem.* **2002**, *23*, 861−869.
(36) Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Energetics of Displacing Water Molecules from Protein Binding Sites: Consequences for Ligand Optimization. *J. Am. Chem. Soc.* **2009**, *131*, 15403−15411.
(37) Kovalenko, A.; Hirata, F. Self-Consistent Description of a Metal-Water Interface by the Kohn-Sham Density Functional Theory and the Three-Dimensional Reference Interaction Site Model. *J. Chem. Phys.* **1999**, *110*, 10095−10112.
(38) Kovalenko, A. Three-Dimensional RISM Theory for Molecular Liquids and Solid−Liquid Interfaces. In *Molecular Theory of Solvation*; Hirata, F., Ed.; Kluwer: Dordrecht, 2003; Vol. *24*, pp 169−275.
(39) Imai, T.; Harano, Y.; Kinoshita, M.; Kovalenko, A.; Hirata, F. A Theoretical Analysis on Hydration Thermodynamics of Proteins. *J. Chem. Phys.* **2006**, *125*, 024911.
(40) Kovalenko, A.; Blinov, N. Multiscale Methods for Nano-chemistry and Biophysics in Solution. *J. Mol. Liq.* **2011**, *164*, 101−112.
(41) Stumpe, M. C.; Blinov, N.; Wishart, D.; Kovalenko, A.; Pande, V. S. Calculation of Local Water Densities in Biological Systems − A Comparison of Molecular Dynamics Simulations and the 3D-RISM-KH Molecular Theory of Solvation. *J. Phys, Chem. B* **2011**, *115*, 319−328.
(42) Imai, T.; Hiraoka, R.; Kovalenko, A.; Hirata, F. Locating Missing Water Molecules in Protein Cavities by the Three-Dimensional Reference Interaction Site Model Theory of Molecular Solvation. *Proteins: Struct., Funct., Bioinf.* **2007**, *66*, 804−813.
(43) Imai, T.; Oda, K.; Kovalenko, A.; Hirata, F.; Kidera, A. Ligand Mapping on Protein Surfaces by the 3D-RISM Theory: Toward Computational Fragment-Based Drug Design. *J. Am. Chem. Soc.* **2009**, *131*, 12430−12440.
(44) Kovalenko, A.; Kobryn, A. E.; Gusarov, S.; Lyubimova, O.; Liu, X.; Blinov, N.; Yoshida, M. Molecular Theory of Solvation for Supramolecules and Soft Matter Structures: Application to Ligand Binding, Ion Channels, and Oligomeric Polyelectrolyte Gelators. *Soft Matter* **2012**, *8*, 1508−1520.
(45) Blinov, N.; Dorosh, D.; Wishart, D.; Kovalenko, A. 3D-RISM-KH Approach for Biomolecular Modeling at Nanoscale: Thermodynamics of Fibril Formation and Beyond. *Mol. Simul.* **2011**, *37*, 718−728.
(46) Harano, Y.; Imai, T.; Kovalenko, A.; Kinoshita, M.; Hirata, F. Theoretical Study for Partial Molar Volume of Amino Acids and Polypeptides by the Three Dimensional Reference Interaction Site Model. *J. Chem. Phys.* **2001**, *114*, 9506−9511.
(47) Blinov, N.; Dorosh, D.; Wishart, D.; Kovalenko, A. Association Thermodynamics and Conformational Stability of β-sheet Amyloid β(17−42) Oligomers: Effects of E22Q (Dutch) Mutation and Charge Neutralization. *Biophys. J.* **2010**, *98*, 282−296.
(48) Yamazaki, T.; Blinov, N.; Wishart, D.; Kovalenko, A. Hydration Effects on the HET-s Prion and Amyloid-β Fibrillous Aggregates, Studied with 3D Molecular Theory of Solvation. *Biophys. J.* **2008**, *95*, 4540−4548.
(49) Sindhikara, D. J.; Yoshida, K.; Hirata, F. Placevent: An Algorithm for Prediction of Explicit Solvent Atom Distribution-Application to HIV-1 Protease and F-ATP Synthase. *J. Comput. Chem.* **2012**, *33*, 1536−1543.
(50) Sindhikara, D. J.; Hirata, F. Analysis of Biomolecular Solvation Sites by 3D-RISM Theory. *J. Phys, Chem. B* **2013**, *117*, 6718−6723.
(51) *Molecular Operating Environment (MOE) 2013.08*; Chemical Computing Group Inc.: Montreal, 2013.
(52) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization and Multithreading. *J. Comput. Chem.* **2010**, *31*, 455−461 vina.scripps.edu.
(53) Chandler, D.; McCoy, J.; Singer, S. Density Functional Theory of Nonuniform Polyatomic Systems. I. General Formulation. *J. Chem. Phys.* **1986**, *85*, 5971−5976. Chandler, D.; McCoy, J.; Singer, S. Density Functional Theory of Nonuniform Polyatomic Systems. II. Rational Closures for Integral Equations. *J. Chem. Phys.* **1986**, *85*, 5977−5982.
(54) Beglov, D.; Roux, B. An Integral Equation to Describe the Solvation of Polar Molecules in Liquid Water. *J. Phys. Chem. B* **1997**, *101*, 7821−7826.
(55) Hansen, J.-P.; McDonald, I. R. *Theory of Simple Liquids*, 3rd ed.; Elsevier: Amsterdam, The Netherlands, 2006.
(56) Perkyns, J.; Pettitt, B. M. A Site-Site Theory for Finite Concentration Saline Solutions. *J. Chem. Phys.* **1992**, *97*, 7656−7666.
(57) Kaminski, J. W.; Gusarov, S.; Wesolowski, T. A.; Kovalenko, A. Modeling Solvatochromic Shifts Using the Orbital-Free Embedding Potential at Statistically Mechanically Averaged Solvent Density. *J. Phys, Chem. A* **2010**, *114*, 6082−6096.
(58) Case, D. A.; Darden, T. A.; Cheatham, T. E. I.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; 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, 2012.
(59) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(60) Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(61) 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., Bioinf.* **2006**, *65*, 712−725.
(62) *GLYCAM Web (2005−2014)*; Woods Group, Complex Carbohydrate Research Center, University of Georgia: Athens, GA,http://www.glycam.com (accessed September 1, 2014).
(63) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration. In *Intermolecular Forces*; Pullman, B., Ed.; D. Reidel Publishing Company: The Netherlands, 1981; pp 331−342.
(64) AmberTools12 Reference Manual. http://ambermd.org/doc12/AmberTools12.pdf (accessed September 1, 2014).

DOI: 10.1021/ci500520q
J. Chem. Inf. Model. 2015, 55, 317−328

(65) Kyte, J.; Doolittle, R. F. A Simple Method for Displaying the Hydropathic Character of a Protein. *J. Mol. Biol.* **1982**, *157*, 105−132.
(66) 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 Flexibity. *J. Comput. Chem.* **2009**, *30*, 2785−2791.
(67) Sanner, M. F. Python: A Programming Language for Software Integration and Development. *J. Mol. Graphics Modell.* **1999**, *17*, 57−61.
(68) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: A Generalizable Biomolecular Force Field. Carbohydrates. *J. Comput. Chem.* **2008**, *29*, 622−655.
(69) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(70) Barillari, C.; Taylor, J.; Viner, R.; Essex, J. W. Classification of Water Molecules in Protein Binding Sites. *J. Am. Chem. Soc.* **2007**, *129*, 2577−2587.
(71) Nikolić, D.; Blinov, N.; Wishart, D. S.; Kovalenko, A. 3D-RISM-Dock: a New Fragment-Based Drug Design Protocol. *J. Chem. Theory Comput.* **2012**, *8*, 3356−3372.
(72) Beuming, T.; Che, Y.; Abel, R.; Kim, B.; Shanmugasundaram, V.; Sherman, W. Thermodynamic Analysis of Water Molecules at the Surface of Proteins and Applications to Binding Site Prediction and Characterization. *Proteins: Struct., Funct., Bioinf.* **2012**, *80*, 871−833.
(73) Szmelcman, S.; Schwartz, M.; Silhavy, T. J.; Boos, W. Maltose Transport in Escherichia Coli K12. A Comparison of Transport Kinetics in Wild-Type and Lambda-Resistant Mutants as Measured by Fluorescence Quenching. *Eur. J. Biochem.* **1976**, *65*, 13−19.
(74) Miller, D. M., 3rd; Olson, J. S.; Pflugrath, J. W.; Quiocho, F. A. Rates of Ligand Binding to Periplasmic Proteins Involved in Bacterial Transport and Chemotaxis. *J. Biol. Chem.* **1983**, *258*, 13665−13672.

DOI: 10.1021/ci500520q
J. Chem. Inf. Model. 2015, 55, 317−328