Tailoring the Variational Implicit Solvent Method for New Challenges: Biomolecular Recognition and Assembly

Clarisse Gravina Ricci[1]*, Bo Li[2,3], Li-Tien Cheng[2], Joachim Dzubiella[4,5] and J. Andrew McCammon[1]

[1] Department of Pharmacology and Department of Chemistry and Biochemistry, National Biomedical Computation Resource, University of California, San Diego, La Jolla, CA, United States, [2] Department of Mathematics, University of California, San Diego, La Jolla, CA, United States, [3] Quantitative Biology Graduate Program, University of California, San Diego, La Jolla, CA, United States, [4] Institut für Physik, Humboldt-Universität zu Berlin, Berlin, Germany, [5] Soft Matter and Functional Materials, Helmholtz-Zentrum Berlin, Berlin, Germany

Predicting solvation free energies and describing the complex water behavior that plays an important role in essentially all biological processes is a major challenge from the computational standpoint. While an atomistic, explicit description of the solvent can turn out to be too expensive in large biomolecular systems, most implicit solvent methods fail to capture "dewetting" effects and heterogeneous hydration by relying on a pre-established (i.e., *guessed*) solvation interface. Here we focus on the Variational Implicit Solvent Method, an implicit solvent method that adds water "plasticity" back to the picture by formulating the solvation free energy as a functional of all possible solvation interfaces. We survey VISM's applications to the problem of molecular recognition and report some of the most recent efforts to tailor VISM for more challenging scenarios, with the ultimate goal of including thermal fluctuations into the framework. The advances reported herein pave the way to make VISM a uniquely successful approach to characterize complex solvation properties in the recognition and binding of large-scale biomolecular complexes.

Keywords: solvation, VISM, implicit solvation, solvation free energy, molecular recognition, binding, solvation free energy of binding, solvent model

INTRODUCTION

The omnipresence of water in all living tissues supports the notion of biochemistry being simply "chemistry in aqueous medium." As a highly polarizable solvent capable of forming a complex net of hydrogen bonds, water is essential in screening electrostatic forces and in forming specific enthalpic interactions with or between biomolecules (Davis and McCammon, 2000; Ball, 2008). Moreover, water is the inherent counter-player in hydrophobicity, which is perhaps the most important driving force behind self-assembly processes, including biomolecular association and binding (Chandler, 2005; Berne et al., 2009). Indeed, water is crucial for molecular recognition (Levy and Onuchic, 2006; Hummer, 2010; Baron and McCammon, 2013), to such an extent that the free energy of ligand-receptor binding could be dominated not by direct interaction between the ligand and its binding pocket, but by water contributions (Baron et al., 2010; Setny et al., 2010).

Due to intrinsic thermodynamic fluctuations and the many configurations that water molecules can adopt to respond to the perturbation imposed by a solute, hydration (or solvation, if one includes ions into the picture) has a flexible disposition (Ball, 2003). While water can increase its local density in the first solvation shells surrounding hydrophilic solutes, it can also "evaporate" at the vicinity of hydrophobic surfaces, as predicted by theoretical models (Parker et al., 1994; Lum et al., 1999), computer simulations (Huang et al., 2003, 2004, 2005; Choudhury and Pettitt, 2005, 2007), or inferred from experiments (Tyrrell and Attard, 2001; Jensen et al., 2003; Schwendel et al., 2003; Steitz et al., 2003; Poynor et al., 2006). Such dewetting transitions are speculated to speed up the hydrophobic collapse that takes place in folding (ten Wolde and Chandler, 2002; Zhou et al., 2004), self-assembly (Lum et al., 1999; Huang et al., 2003, 2004; Liu et al., 2005) and molecular recognition (Young et al., 2007; Ahmad et al., 2008). In some cases, water might also completely disappear from hydrophobic protein cavities in the unbound state, with important implications for ligand binding (Young et al., 2007; Qvist et al., 2008; Matthews and Liu, 2009; Wang et al., 2011; Krimmer et al., 2017).

The outcomes of water plastic behavior are especially hard to predict when the solute itself also displays a high level of complexity. Thus, while relatively simple models can be used to predict the hydration pattern between two approaching paraffin plates (Lum et al., 1999), such attempt becomes harder for amphiphilic plates (Hua et al., 2009) and practically impossible in face of protein interacting surfaces, whose charge distribution and corrugated topology result in a detailed hydrophobicity/hydrophilicity balance, not to mention their conformational flexibility. Therefore, there is a great interest in developing computational tools to simulate microscopic (water distribution) and thermodynamic properties (solvation free energies) arising from such complexity.

Explicit solvent molecular dynamics (MD) simulations have proved to be a useful approach to describe water plasticity in the context of molecular association and binding, as revealed by simulations capturing capillary evaporation between two hydrophobic surfaces (Huang et al., 2003, 2005; Choudhury and Pettitt, 2005, 2007) and dewetting transitions in the interaction of amphiphilic protein interfaces (Huang et al., 2004; Zhou et al., 2004; Liu et al., 2005). However, explicit solvent MD simulations of large biomolecular systems are expensive and may require an impractical sampling of water configurations in order to obtain converged solvation free energies, unless one relies on the additional use of enhanced sampling techniques and/or perturbation methods.

As an alternative to explicit solvation, there is a great advantage in modeling water as continuum medium, in which microscopic structure and fluctuations are reduced to macroscopically related properties such as dielectric constant, ϵ, and surface tension, γ. Normally referred to as "implicit solvation" or "continuum medium", methods based on such an approach have significantly lower computational cost and can be used to calculate solvation free energies while avoiding altogether the challenge of getting enough statistical sampling of water configurations.

Implicit solvation or continuum methods typically split the system into a solute region that is treated explicitly, Ω_m, and a solvent region that is treated as a continuum, Ω_w, separated by a dielectric boundary, Γ (**Figure 1A**). Within this framework, the electrostatic component of the solvation free energy can be modeled by a continuum electrostatic reaction field—obtained by solving the Poisson–Boltzmann equation (Davis and McCammon, 2000; Baker, 2005) or the more simplified Generalized Born model (Bashford and Case, 2000; Feig and Brooks, 2004). Additionally, the hydrophobic contribution arising from first-shell solvation effects can be empirically modeled by surface tension-like coefficients integrated over the solute-solvent surface area (Eisenberg and McLachlan, 1986). Some of the most popular implicit solvent methods estimate the complete solvation free energy by independently calculating and then adding the electrostatic energy to the non-polar solvation free energy, assuming them to be additive (Still et al., 1990; Roux and Simonson, 1999).

In this context, an often neglected question is: what should be the dielectric boundary used in these calculations? While the location and the shape of the solvation interface can significantly impact the results—PB calculations, for instance, are extremely sensitive to the chosen dielectric boundary—implicit solvation methods conventionally employ surfaces that are closely related to the van der Waals surface of the protein (**Figure 1B**). Therefore, they rely on pre-established solvation interfaces normally *guessed* as solvent accessible surfaces (SAS) and *fixed* during the calculations.

Such an approach is in evident conflict with water's aforementioned plastic behavior and, while successful in many cases, cannot capture dewetting effects or the existence of polymodal hydration. **Figure 1C** illustrates the discrepant scenario resulting from the use of a SAS as dielectric boundary for a protein whose hydrophobic pocket expels water. More than a didactic example, concave and hydrophobic cavities such as the one in **Figure 1C** are good representatives of druggable binding pockets. By expelling water, such pockets strongly enhance their affinity for apolar molecules, which do not need to competitively displace water molecules in order to bind (Young et al., 2007; Wang et al., 2011; Krimmer et al., 2017).

This review focuses on the Variational Implicit Solvent Method (VISM), an implicit solvent method that can add water "plasticity" back to the picture by formulating the solvation free energy as a functional of the solvation interface, Γ, and then relaxing the interface toward the solvation free energy minimum (Dzubiella et al., 2006a,b). As such, VISM produces equilibrium solute-solvent interfaces as *output* of the theory while also coupling electrostatic and apolar contributions through the solvation boundary. Noteworthy, our method fits into a longer tradition of functional-based variational methods in condensed matter physics (Lum et al., 1999; Wallqvist et al., 2001; Ramirez and Borgis, 2005; Chen et al., 2010; Zhao et al., 2011; Jeanmairet et al., 2013). We begin by introducing the VISM functional and its combination with a level-set framework (Cheng et al., 2007, 2009b) that allows to deal with the very complex topologies exhibited by proteins and other biomolecules. Next we survey relevant applications of VISM in the context of molecular
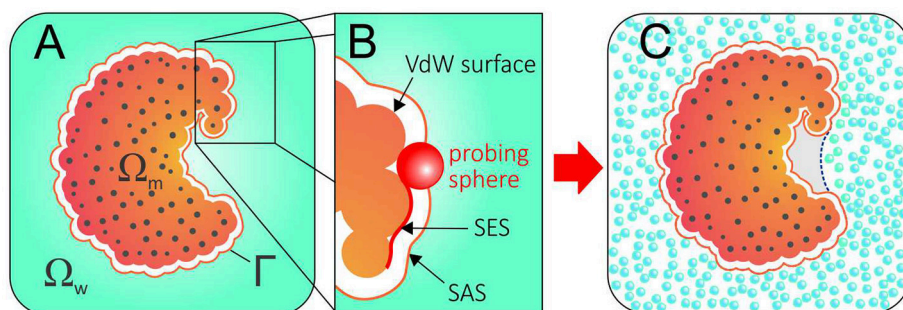
FIGURE 1 | (A) The geometry of a system described with an implicit solvent approach. The solute region, solvent region, and solute-solvent interface are denoted by Ω_m, Ω_w, and Γ, respectively. **(B)** Solvent Excluded Surface (SES) and Solvent Accessible Surface (SAS) are obtained by rolling a probing ball along the van der Waals surface. **(C)** Conflict between the SAS (solid orange line) and the explicit hydration displayed by a hydrophobic pocket that expels water. A correct implicit solvent description of such pocket would require an alternative solvation boundary (dotted blue line).
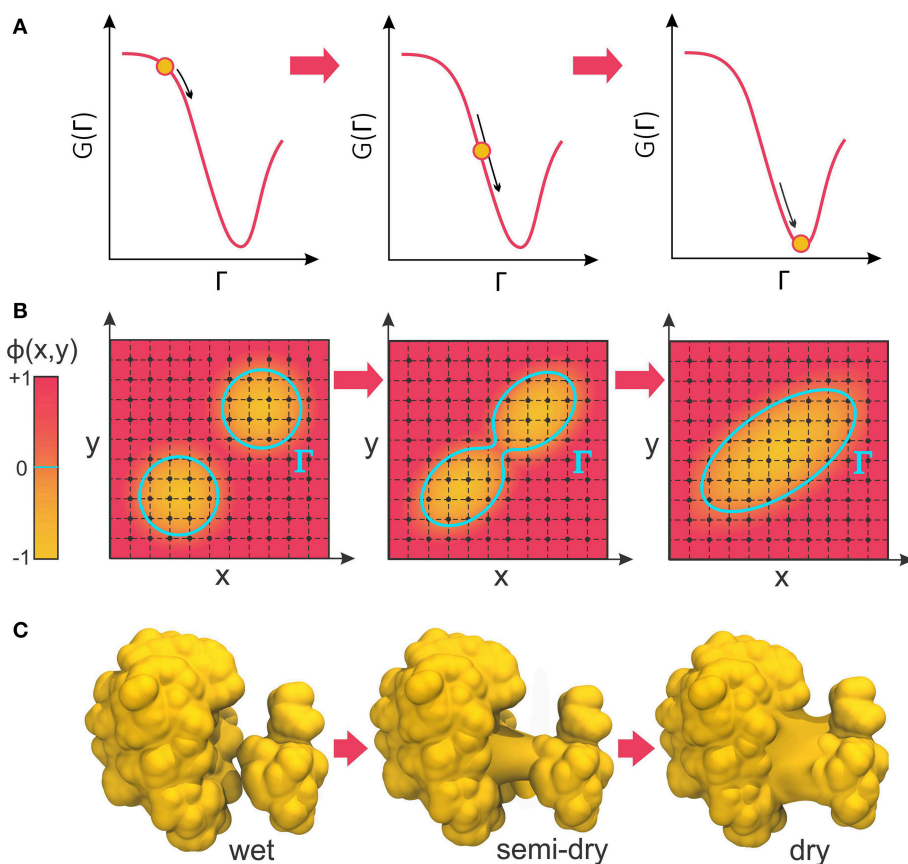

FIGURE 2 | (A) VISM calculations minimize the solvation free energy as a functional of the solvation interface, Γ, producing a final solvent-solute boundary that corresponds to a stable solvation state. **(B)** The level-set framework defines the solvation boundary, Γ, as the zero-level set of an auxiliary function, $\phi(x,y)$. By implicitly manipulating Γ through modification of the level-set function, $\phi(x,y)$, LS-VISM can easily track topological changes, as when two boundaries merge into one. **(C)** Corresponding example of two solvation interfaces merging as the proteins undergo a dewetting transition.

recognition and binding, highlighting the features that make VISM uniquely able to capture multiple states of solvation and the equilibrium between these states. Finally, we report how VISM can be combined with the MARTINI FF (Marrink et al., 2004, 2007; Monticelli et al., 2008; de Jong et al., 2013) to describe the solvation of coarse-grained proteins, with an example of how this method can be applied to study solvation in encounter complexes. We conclude by outlining the next steps to make

VISM an appealing tool not only to model, but also simulate the role of water in large biomolecular assemblies, hopefully in the near future.

THEORY

The VISM Functional

Consider the system displayed in **Figure 1A**. In VISM, as in other implicit solvation methods, the system is divided in three parts: (i) the solute region, Ω_m; (ii) the solvent region, Ω_w; and (iii) a solute-solvent interface, Γ. As usual, the solute region, Ω_m, contains all the N atoms belonging to the solute molecule, which are located at $x_1, \ldots, x_N$ inside Ω_m and carry point charges $Q_1, \ldots, Q_N$, while the solvent region, Ω_w, is treated as a continuum medium. However, instead of using a fixed solvation interface, VISM relaxes an initially guessed interface, Γ, toward the solvation free energy minimum by means of the functional:

$$G(\Gamma) = G_{geometric}(\Gamma) + G_{dispersion}(\Gamma) + G_{electrostatics}(\Gamma)$$

$$G(\Gamma) = \int_\Gamma \gamma \, dS + \rho_w \sum_{i=1}^N \int_{\Omega_w} U_i(|x - x_i|) dV + G_{elec}(\Gamma) \quad (1)$$

The first term in Equation (1) is purely geometric and accounts for first-shell solvation effects giving rise to hydrophobicity. In analogy to SASA-methods (Eisenberg and McLachlan, 1986), it consists of an integration of the surface tension, γ, over the solvation interface, Γ. However, the surface tension in VISM's formulation is sensitive to the local shape of the solvation boundary, being defined as:

$$\gamma = \gamma_0 (1 - 2\tau H) \quad (2)$$

where γ_0 is the constant macroscopic surface tension for a planar liquid-vapor interface, H is the mean curvature defined as the average of the two principal curvatures, and τ is a curvature correction coefficient (Dzubiella et al., 2006a). Such formulation models the higher surface tension displayed by concave geometries (negative H), thus increasing the hydrophobicity of deeply buried pockets as opposed to flat or convex protein surfaces. The curvature correction coefficient, τ, is a fitting parameter that accounts for the relative size of the solvent molecules with respect to the solute local curvature: the larger the size of the solvent molecules, the more sensitive solvent organization will be with respect to the solute curvature, leading to a more pronounced hydrophobic effect.

The second term in Equation (1) accounts for dispersion interactions between water and solute, which are modeled by 12–6 Lennard–Jones potentials, U_i:

$$U_i = 4\varepsilon_i \left[\left(\frac{\sigma_i}{|x - x_i|} \right)^{12} - \left(\frac{\sigma_i}{|x - x_i|} \right)^6 \right] \quad (3)$$

For the ith atom of the solute, located at x_i, the interaction potential is integrated over the entire solvent volume, which is represented by a grid with a numeric density pre-factor, ρ_w. We normally use sub-Å resolution grids, with cells of $\sim$0.3 to 0.5 Å. The solute and solvent Lennard-Jones parameters, ε_i and σ_i, are borrowed from molecular mechanics force fields (Best et al., 2012; Maier et al., 2015) and the TIP3P water model (Jorgensen et al., 1983).

Finally, the third term in Equation (1), $G_{elec}(\Gamma)$, is the electrostatic part of the solvation free energy. This term can be calculated at high level with PB theory (in its linearized or nonlinear forms) (Zhou et al., 2014, 2015), or with the advantageously faster but more approximate Coulomb Field Approximation (CFA) (Wang et al., 2012; Guo et al., 2013):

$$G_{elec}^{CFA}(\Gamma) = \frac{1}{32\pi^2\epsilon_0} \left(\frac{1}{\epsilon_w} - \frac{1}{\epsilon_m} \right) \int_{\Omega_w} \left| \sum_{i=1}^N \frac{Q_i(x - x_i)}{|x - x_i|^3} \right|^2 dV \quad (4)$$

where ϵ_0 is the vacuum permittivity, and ϵ_w and ϵ_m are the relative permittivity for the solvent and the solute, respectively. We highlight the CFA theory because it is common practice to employ CFA for most part of the boundary relaxation, and only switch to PB in the final steps of a VISM calculation (Zhou et al., 2015). Moreover, the studies reviewed in the next section employed CFA whenever electrostatics effects were included in the calculations.

An immediate consequence of the VISM functional is that the apolar ($G_{geometric} + G_{dispersion}$) and polar ($G_{electrostatics}$) components of the total solvation free energy are coupled together through the solvation boundary, Γ. A second consequence is that the VISM functional can be differentiated with respect to the local change of the solvation boundary:

$$\delta_\Gamma G(\Gamma) = 2\gamma_0 \left[H(x) - \tau K(x) \right] - \rho_w \sum_{i=1}^N U_i(|x - x_i|)$$
$$- \frac{1}{32\pi^2\epsilon_0} \left(\frac{1}{\epsilon_w} - \frac{1}{\epsilon_m} \right) \left| \sum_{i=1}^N \frac{Q_i(x - x_i)}{|x - x_i|^3} \right|^2 \quad (5)$$

where $K(x)$ is the Gaussian curvature, defined as the product of the two principal curvatures, at a point x on Γ.

The negative of this functional derivative, $-\delta_\Gamma G(\Gamma)$, can be considered an effective "boundary force" whose normal component, pointing from the solute toward the solvent, acts on the boundary interface. In VISM calculations, this force drives the boundary toward a final shape (or solvation state) that minimizes the solvation free energy (**Figure 2A**). While such minimization can be solved analytically for very simple geometries (Dzubiella et al., 2006a,b), proteins and other biomolecules are far from having topologically simple shapes. Therefore, Cheng et al. have combined VISM to a Level-Set approach, which can numerically find the free-energy minimizing solute-solvent interface for molecules of arbitrary shapes (Cheng et al., 2007, 2009b).

Level-Set VISM (LS-VISM)

In the level-set framework, the solvation boundary is defined as the zero-level set of an auxiliary function $\phi(x)$: $\Gamma = \{(x) | \phi(x) = 0\}$. The function $\phi(x)$ – called the level-set function of the surface Γ – is continuous and well-defined at all grid points of the finite computational box encompassing the system (**Figure 2B**). The solute region, Ω_m, is defined by points

where $\phi(\mathbf{x}) < 0$, while the solvent region, Ω_w, is defined by points where $\phi(\mathbf{x}) > 0$.

The unit vector $\mathbf{n}$ at the interface Γ, the mean curvature H, and the Gaussian curvature K can all be expressed in terms of the level-set function as:

$$\mathbf{n} = \frac{\nabla\phi}{|\nabla\phi|}, \ H = \frac{1}{2}\nabla\cdot\mathbf{n}, \ K = \mathbf{n}\cdot\text{adj}\,(\text{He}\,(\phi))\,\mathbf{n} \qquad (6)$$

where $\text{He}(\phi)$ is the 3×3 Hessian matrix of the function ϕ whose entries are all the second order derivatives $\partial_{ij}^2\phi$ of the level set function ϕ, and $\text{adj}(\text{He}(\phi))$ is the adjoint matrix of the Hessian $\text{He}(\phi)$. These properties are necessary for the calculation of the solvation force, $-\delta_\Gamma G\,(\Gamma)$ (see Equation 5), which is used in the level-set framework as the normal velocity, υ_n, in solving the level-set equation:

$$\partial_t\phi + \upsilon_n\,|\nabla\phi| = 0 \qquad (7)$$

Solution of the level-set equation can be used to determine, in a *steepest descent* manner, the evolution of the level-set function, $\phi(\mathbf{x},t)$, and the corresponding motion of the solvation interface, $\Gamma(t)$, defined as the zero-level set of $\phi(\mathbf{x},t)$ at each time t. Here, "time" corresponds to a minimization step.

The great advantage of the level-set method is that, by implicitly manipulating Γ through the underlying level-set function, $\phi(\mathbf{x})$, it can easily follow shapes that change topology, as for instance when two shapes merge in one (**Figure 2B**). Therefore, it is an appropriate approach to relax solvation boundaries that merge together as in a desolvated state (**Figure 2C**) or in the association between two binding proteins. By combining it with the level-set method, Cheng et al. promoted VISM from a toy solvation model to study spherical or cylindrical systems into a competitive solvation method able to deal with systems of more interesting geometries (Cheng et al., 2007, 2009b), as described in the next section.

LS-VISM APPLICATIONS TO MOLECULAR ASSOCIATION AND BINDING
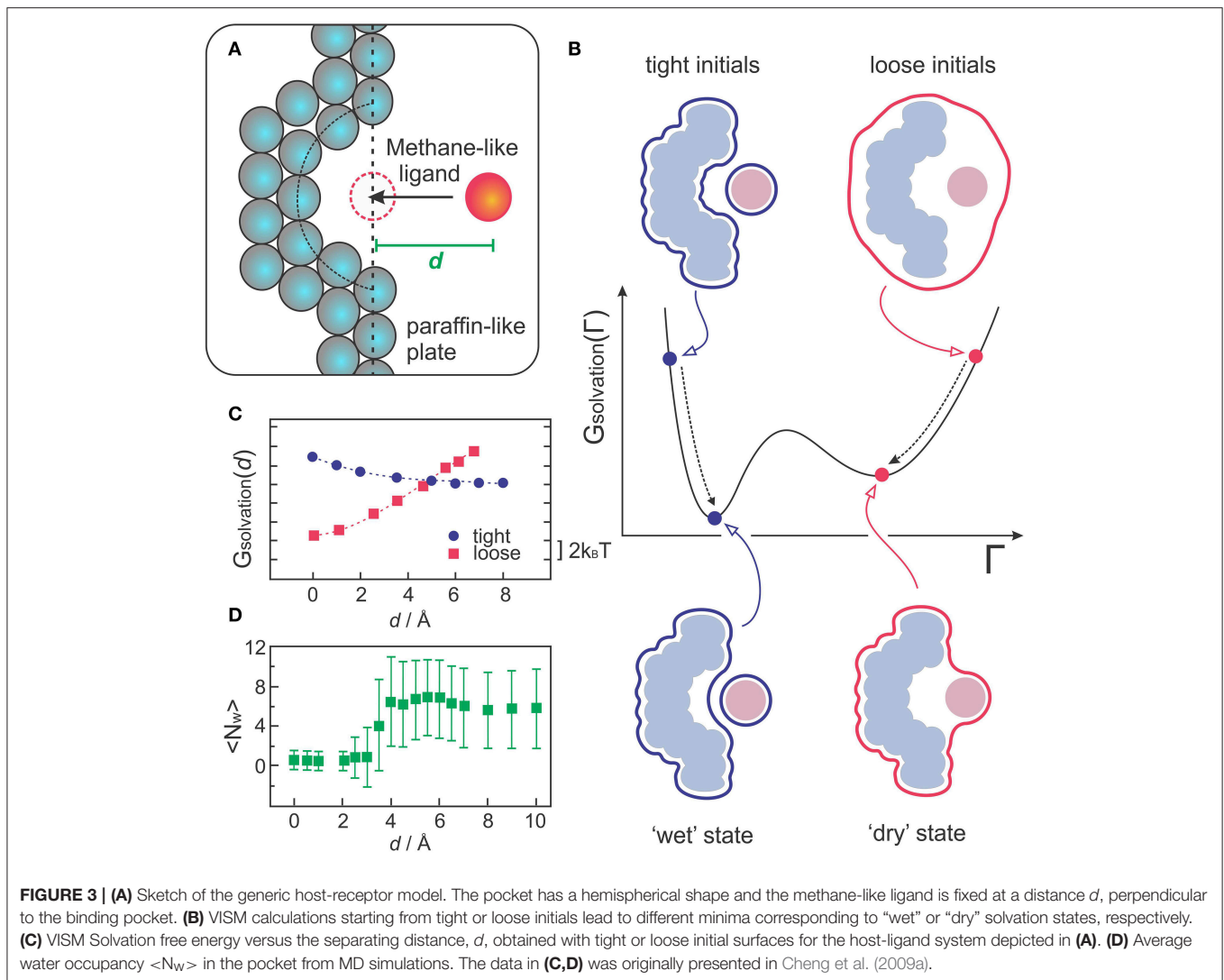
Receptor-Ligand Model Systems

VISM was the first implicit solvent method to capture multistate hydration in molecular recognition, as demonstrated with a simple but insightful model of a hydrophobic receptor-ligand system (Cheng et al., 2009a; Setny et al., 2009). The model consists of a paraffin-like plate containing a hemispherical nanoscopic pocket that binds a methane-like ligand (**Figure 3A**). Despite its apparent simplicity, this system can provide insight into the hypothesis that apolar pockets with concave geometries are particularly prone to capillary evaporation induced by an approaching ligand (Setny, 2007; Young et al., 2007; Ahmad et al., 2008). As such, the solvation properties of this generic host-ligand system have been studied in detail by explicit solvent MD simulations (Setny and Geller, 2006; Setny, 2007, 2008; Cheng et al., 2009a), which captured solvent fluctuations corresponding to more or less hydrated states of the pocket, depending on the proximity of the ligand. While highly hydrated (wet) states

predominate at large host-ligand distances, MD simulations revealed that poorly hydrated (dry) states start to become more favorable at a critical distance of ∼4 Å, finally predominating over wet states at close distances that do not yet sterically prevent water to occupy the binding interfaces (Cheng et al., 2009a).

Such heterogeneous hydration behavior poses a challenge for VISM, especially considering that the minimization of a free-energy functional with a given initial boundary can only provide one solution describing a static solvent distribution. To deal with this limitation, one can take advantage of the fact that VISM performs only a *local* minimization of the solvation free energy along the "solvation landscape." Therefore, in systems that allow for heterogeneous hydration patterns, calculations starting from different initial conditions can lead to different local minima, corresponding to distinct equilibrium solvation states (**Figure 3B**). Cheng et al. performed VISM calculations on this generic host-ligand system starting from several different initial surfaces, some of them tightly wrapping the solutes around their van der Waals surfaces (tight initials), others loosely wrapping the solutes in one single surface (loose initials) (Cheng et al., 2009a). This approach allowed them to capture the bimodal hydration displayed by this hydrophobic cavity and to reproduce the critical ligand distance at where dry states start to predominate over wet states (**Figure 3C**), followed by a complete dewetting transition, in good agreement with explicit solvent MD simulations (**Figure 3D**). Moreover, the curvature correction to the surface tension, which is normally absent in SASA-based methods, proved to be crucial to describe dewetting transitions in the concave pocket, with the curvature correction coefficient affecting the onset distance at where polymodal hydration starts (Cheng et al., 2009a; Setny et al., 2009).

Assembly of the BphC Monomer

Wang et al. applied VISM to study the collapse between two domains that form the BphC protein monomer (Wang et al., 2012). Such process can be interpreted as the last step in the folding mechanism of multidomain proteins, with water driving the assembly of individually formed domains into the final globular structure. VISM calculations starting from both tight and loose initials revealed the existence of two "solutions" for the solvation boundary when the domains lie at intermediate separation distances of ∼4–14 Å (Wang et al., 2012). These "solutions" correspond to "dry" and "wet" solvation states that indirectly reflect the fluctuating nature of the solvent in equilibrium. With wet states being energetically more favorable than dry ones, a Boltzmann averaging of such discrete states would roughly predict a decreased but not vanishing solvent density in the interdomain region. This prediction is in good agreement with explicit solvent MD simulations of separated BphC domains, showing that water density at separation distances of 6Å becomes 15% lower than in the bulk (Zhou et al., 2004). The fact that this system does not undergo complete dewetting is attributed to favorable electrostatic and van der Waals interactions between BphC and water at the interdomain region. In agreement with MD simulations (Zhou et al., 2004), VISM calculations predicted significant dewetting when the protein charges and dispersion interactions with the solvent were
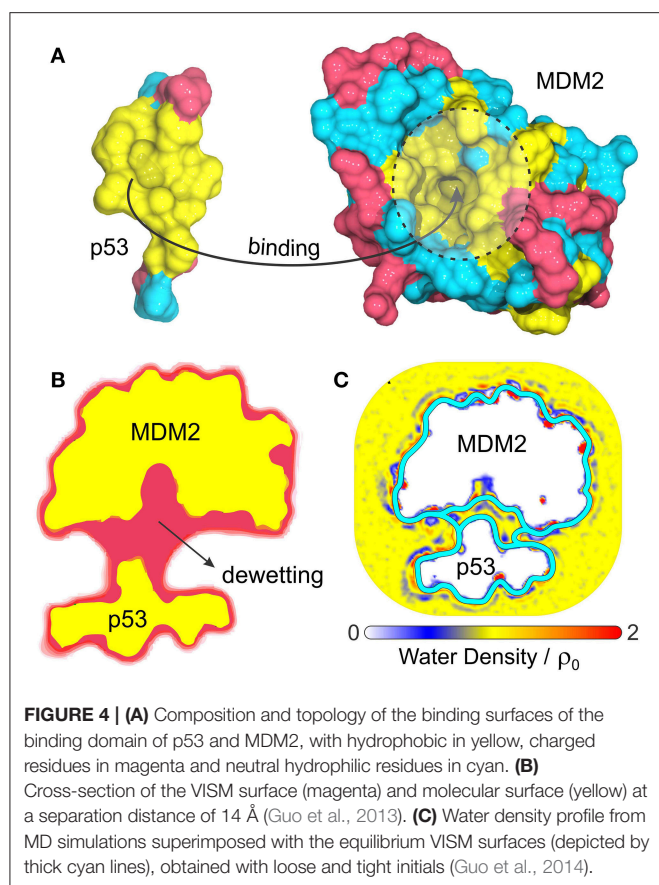
FIGURE 3 | (A) Sketch of the generic host-receptor model. The pocket has a hemispherical shape and the methane-like ligand is fixed at a distance d, perpendicular to the binding pocket. **(B)** VISM calculations starting from tight or loose initials lead to different minima corresponding to "wet" or "dry" solvation states, respectively. **(C)** VISM Solvation free energy versus the separating distance, d, obtained with tight or loose initial surfaces for the host-ligand system depicted in **(A)**. **(D)** Average water occupancy $<N_w>$ in the pocket from MD simulations. The data in **(C,D)** was originally presented in Cheng et al. (2009a).

turned off, revealing the importance of coupling polar and non-polar interactions when dealing with biological systems (Wang et al., 2012).

Binding of p53 to MDM2

Another interesting biological problem to which VISM has been successfully applied concerns the role of solvation in the binding of tumor-suppressing protein p53 to its repressor, MDM2 (Guo et al., 2013, 2014). p53 plays a vital role in suppressing tumors and disruption of its interaction with MDM2 is the guideline behind many anticancer therapies (Chène, 2003, 2004; Zhao et al., 2013). This system displays a particularly strong hydrophobic character at the p53/MDM2 binding interface (Kussie et al., 1996), with MDM2's binding pocket comprising a concave apolar patch with few hydrophilic residues lying at the edges (**Figure 4A**). As such, p53-MDM2 complex provides an excellent example of a biologically relevant system whose molecular association is likely to involve dewetting transitions or at least heterogeneous hydration.

Guo et al. have used VISM to estimate the solvation behavior arising from the delicate interplay between complicated geometry, hydrophobicity and polar interactions at the binding interfaces of p53 and MDM2 (Guo et al., 2013, 2014). As a result, VISM calculations of the two proteins separated by distances as large as 14 Å revealed significant water depletion inside the binding pocket and in the inter-domain region (**Figure 4B**) (Guo et al., 2013). This was found to be consistent with explicit solvent MD simulations of the two proteins separated by 12 Å, which revealed capillary evaporation at the bottom of MDM2's binding pocket, with the first solvation layers forming near its entrance (**Figure 4C**) (Guo et al., 2014). Additionally, the solvation free energy of binding estimated by VISM (237.3 kcal/mol) was in relatively good quantitative agreement with explicit water FEP calculations (306.7 kcal/mol), especially considering the large size (and associated errors) of this system (Guo et al., 2013). To conclude, dewetting transitions captured by VISM calculations are not only realistic, but they might also contribute for fast kinetics of binding between p53 and MDM2. Thus, by efficiently

FIGURE 4 | (A) Composition and topology of the binding surfaces of the binding domain of p53 and MDM2, with hydrophobic in yellow, charged residues in magenta and neutral hydrophilic residues in cyan. **(B)** Cross-section of the VISM surface (magenta) and molecular surface (yellow) at a separation distance of 14 Å (Guo et al., 2013). **(C)** Water density profile from MD simulations superimposed with the equilibrium VISM surfaces (depicted by thick cyan lines), obtained with loose and tight initials (Guo et al., 2014).

describing solvent behavior of this type of system, VISM could contribute to develop solvation-driven strategies of controlling the association of protein complexes with relevant impact in biology and health.

^{M}VISM and Solvation for Coarse-Grained Complexes

Combining VISM with a coarse-grained model for the solute is an interesting step to push VISM toward large-scale applications and eventually merge it with molecular dynamics simulations. Recently, we adapted VISM to produce solvation free energies for "martinized" proteins (Ricci et al., 2017). The MARTINI model is a well-established meso-scale force field for modeling large molecular systems, which replaces groups of atoms by interaction centers commonly referred to as "beads," based on an approximate 4-to-1 mapping (Marrink et al., 2004, 2007; Monticelli et al., 2008).

For this purpose, the main adaptation of the VISM functional consisted in replacing atomistic Lennard-Jones parameters, ε_i and σ_i, by the coarse-grained LJ parameters reported in Martini 2.1 (Marrink et al., 2007; Monticelli et al., 2008). Application of this new method, denominated ^{M}VISM, to estimate the solvation free energies of six different proteins revealed a good qualitative agreement between the fully coarse-grained approach and the atomistic version obtained with the original LS-VISM. We also
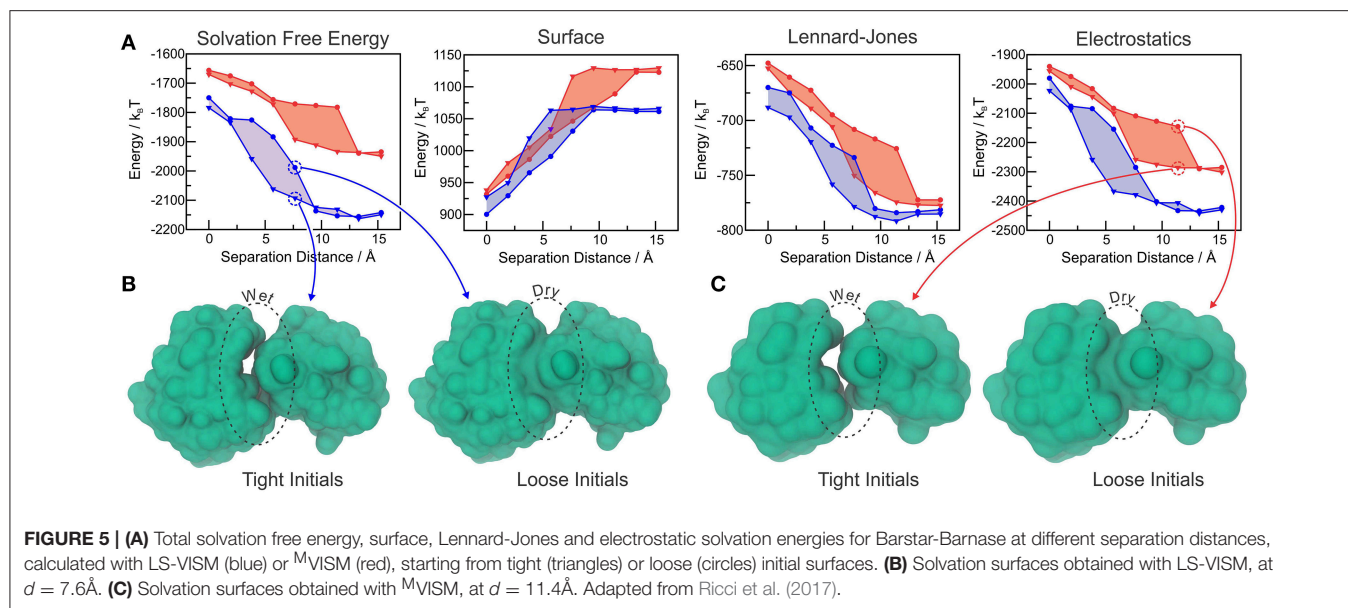
found the solvation free energies obtained with ^{M}VISM to be significantly underestimated due to un-optimized (too favorable) Lennard-Jones interaction energies between solute and water (Ricci et al., 2017). Overestimation of van der Waals interactions is characteristic of MARTINI, as reported in hydration of organic compounds (Marrink et al., 2007), protein-protein binding (Stark et al., 2013) and aggregation of polysaccharides (Schmalhorst et al., 2017). This trait stems from the fact that the force field was originally developed to model highly apolar environments such as lipid bilayers. In fact, increasing the strength of LJ interactions is a way of compensating for (i) the simplistic treatment given to electrostatic interactions in the Martini model, and (ii) the reduced number of degrees of freedom in Martini water, which can significantly reduce entropically-driven hydrophobic attraction between apolar molecules. In ^{M}VISM, overestimation of van der Waals interactions could be easily fixed by a simple downscaling of the Martini ε_i values, which brought solvation energies from the fully coarse-grained approach to fairly good quantitative agreement with the ones from the atomistic approach, while also providing a better partition of the solvation free energy among the coupled energy terms (hydrophobic, Lennard-Jones, and electrostatics) (Ricci et al., 2017).

We further tested ^{M}VISM's ability to describe hydration in the tight-binding barnase-barstar complex, which has achieved its extremely fast kinetics of binding by means of optimized electrostatic interactions (Lee and Tidor, 2001). **Figure 5A** shows the solvation free energies obtained with coarse-grained (^{M}VISM) or atomistic (LS-VISM) calculations of the barstar-barnase complex with separation distances ranging from 0 (native complex) to 15 Å. As compared to atomistic solvation energies, the coarse-grained energies are not as favorable, mainly due to the electrostatic term. Overestimation of the electrostatic solvation energy is very likely a consequence of (i) polar beads carrying no partial charges in Martini2.1, and (ii) the point charges carried by charged beads being buried deeper in the solute interface due to the large size of the beads.

The hysteresis in the energies obtained with loose or tight initial surfaces evidence the existence of multiple solvation states, some of which are illustrated in **Figures 5B,C**. As expected, tight initial conditions produce states that are systematically more hydrated as compared to the states obtained with loose initials, with larger discrepancies occurring for intermediate separation distances. Interestingly, both methods predict the wet encounter pathway to be significantly more stable than the dry one, as a result of favorable electrostatic interactions of the binding interfaces with water. This finding indicates that, despite the simplifications imposed by the Martini model in the treatment of electrostatic interactions, ^{M}VISM can successfully predict the highly hydrated nature of electrostatically-optimized association between barstar and barnase.

Final Considerations

In this review, we focus on the level-set VISM (LS-VISM) and its ability to capture the plastic behavior displayed by water in molecular binding and protein-protein association. Such ability distinguishes VISM from more traditional implicit

FIGURE 5 | (A) Total solvation free energy, surface, Lennard-Jones and electrostatic solvation energies for Barstar-Barnase at different separation distances, calculated with LS-VISM (blue) or ^{M}VISM (red), starting from tight (triangles) or loose (circles) initial surfaces. **(B)** Solvation surfaces obtained with LS-VISM, at $d = 7.6$Å. **(C)** Solvation surfaces obtained with ^{M}VISM, at $d = 11.4$Å. Adapted from Ricci et al. (2017).

solvent methods that rely on a fixed and pre-established solvation boundary, which are intrinsically incapable of capturing dewetting events.

We surveyed a few relevant applications of LS-VISM in the study of receptor-ligand hydrophobic binding, hydrophobic assembly of BphC protein, and dewetting effects that govern the association between p53 and MDM2. While these studies display good agreement with explicit solvent MD simulations, it is worth noting that they were performed with a relatively simple treatment of electrostatic interactions (CFA) or no electrostatic treatment at all in the case of hydrophobic ligand-host studies. More accurate results could be obtained by employing a Poisson-Boltzmann treatment of electrostatic interactions, as implemented in the current version of LS-VISM, also available in a public software package (Zhou et al., 2015).

In the last example, we showed how LS-VISM can work in combination with Martini2.1 to predict the hydration behavior in the association of a relevant biological complex (barnase-barstar). While encouraged by this success, we are aware that ^{M}VISM could be further improved by using newer and more sophisticated versions of Martini, which provide better electrostatic description and attempt to deal with the problem of superestimated LJ interactions and over aggregation. We are currently testing the more recent "polarizable" Martini2.2P (de Jong et al., 2013) and looking forward to test a new Martini that should be released in the near future (personal communication).

The use of a coarse-grained model for the solute is a sensible strategy to combine VISM with MD simulations, which typically require many thousands of integration steps if one desires

to sample rare events such as diffusion-controlled molecular binding while still keeping track of conformational dynamics and solvation plasticity. Unfortunately, relaxation of the solvation boundary with LS-VISM is still too slow (minutes to hours) and poses the last remaining obstacle preventing VISM from being efficiently combined with MD simulations. To deal with the problem of speed, we are currently working on a new boundary-relaxing algorithm capable of approximating the boundary solutions produced by the level-set framework with significantly lower computational cost. We expect this approach to facilitate the development of a hybrid MD-VISM method to not only model, statically, but actually *simulate* solvation in the assembly of large biomolecular complexes using the VISM framework.

AUTHOR CONTRIBUTIONS

All authors contributed to the design, organization, and review of the manuscript. CR took the lead in writing the manuscript and prepared the figures.

ACKNOWLEDGMENTS

REFERENCES

Ahmad, M., Gu, W., and Helms, V. (2008). Mechanism of fast peptide recognition by SH3 domains. *Angew. Chem.* 47, 7626–7630. doi: 10.1002/anie.200801856

Baker, N. A. (2005). Improving implicit solvent simulations: a poisson-centric view. *Curr. Opin. Struct. Biol.* 15, 137–143. doi: 10.1016/j.sbi.2005.02.001

Ball, P. (2003). How to keep dry in water. *Nature* 423, 25–26. doi: 10.1038/423025a

Ball, P. (2008). Water as an active constituent in the cell biology. *Chem. Rev.* 108, 74–108. doi: 10.1021/cr068037a

Baron, R., and McCammon, J. A. (2013). Molecular recognition and ligand association. *Annu. Rev. Phys. Chem.* 64, 151–175. doi: 10.1146/annurev-physchem-040412-110047

Baron, R., Setny, P., and McCammon, J. A. (2010). Water in cavity-ligand recognition. *J. Am. Chem. Soc.* 132, 12091–12097. doi: 10.1021/ja1050082

Bashford, D., and Case, D. A. (2000). Generalized born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.* 51, 129–152. doi: 10.1146/annurev.physchem.51.1.129

Berne, B. J., Weeks, J. D., and Zhou, R. (2009). Dewetting and hydrophobic effect in physical and biological systems. *Annu. Rev. Phys. Chem.* 60, 85–103. doi: 10.1146/annurev.physchem.58.032806.104445

Best, R. B., Zhu, X., Shim, J., Loper, P. E. M., Mittal, J., Feig, M., et al. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi1 and chi2 dihedral angles. *J. Chem. Theor. Comput.* 8, 3257–3273. doi: 10.1021/ct300400x

Chandler, D. (2005). Interfaces and the driving force of hydrophobic assembly. *Nature* 437, 640–647. doi: 10.1038/nature04162

Chène, P. (2003). Inhibiting the p53-MDM2 interaction: an important target for cancer therapy. *Nat. Rev. Cancer* 3, 102–109. doi: 10.1038/nrc991

Chène, P. (2004). Inhibition of the p53-MDM2 interaction: targeting a protein-protein interface. *Mol. Cancer Res.* 2, 20–28.

Chen, Z., Baker, N. A., and Wei, G. W. (2010). Differential geometry based solvation model i: eulerian formulation. *J. Comput. Phys.* 229, 8231–8258. doi: 10.1016/j.jcp.2010.06.036

Cheng, L. T., Dzubiella, J., McCammon, J. A., and Li, B. (2007). Application of the level-set method to the implicit solvation of nonpolar molecules. *J. Chem. Phys.* 127:084503. doi: 10.1063/1.2757169

Cheng, L. T., Wang, Z., Setny, P., Dzubiella, J., Li, B., and McCammon, J. A. (2009a). Interfaces and hydrophobic interactions in receptor-ligand systems: a level-set variational implicit solvent approach. *J. Chem. Phys.* 131:144102. doi: 10.1063/1.3242274

Cheng, L. T., Xie, Y., Dzubiella, J., McCammon, J. A., Che, J., and Li, B. (2009b). Coupling the level-set method with molecular mechanics for variational implicit solvation of nonpolar molecules. *J. Chem. Theory Comput.* 5, 257–266. doi: 10.1021/ct800297d

Choudhury, N., and Pettitt, B. M. (2005). On the mechanism of hydrophobic association of nanoscopic solutes. *J. Am. Chem. Soc.* 127, 3556–3567. doi: 10.1021/ja0441817

Choudhury, N., and Pettitt, B. M. (2007). The dewetting transition and the hydrophobic effect. *J. Am. Chem. Soc.* 129, 4847–4852. doi: 10.1021/ja069242a

Davis, M. E., and McCammon, J. A. (2000). Electrostatics in biomolecular structure and dynamics. *Chem. Rev.* 90, 509–521. doi: 10.1021/cr00101a005

de Jong, D. H., Singh, G., Bennett, W. F., Arnarez, C., Wassenaar, T. A., Schafer, L. V., et al. (2013). Improved parameters for the Martini coarse-grained protein force field. *J. Chem. Theor. Comput.* 9, 687–697. doi: 10.1021/ct300646g

Dzubiella, J., Swanson, J. M., and McCammon, J. A. (2006a). Coupling hydrophobicity, dispersion, and electrostatics in continuum solvent models. *Phys. Rev. Lett.* 96:087802. doi: 10.1103/PhysRevLett.96.087802

Dzubiella, J., Swanson, J. M., and McCammon, J. A. (2006b). Coupling nonpolar and polar solvation free energies in implicit solvent models. *J. Chem. Phys.* 124:084905. doi: 10.1063/1.2171192

Eisenberg, D., and McLachlan, A. D. (1986). Solvation energy in protein folding and binding. *Nature* 319, 199–203. doi: 10.1038/319199a0

Feig, M., and Brooks, C. L. III. (2004). Recent advances in the development and applications of implicit solvent models in biomolecule simulations. *Curr. Opin. Struct. Biol.* 14, 217–224. doi: 10.1016/j.sbi.2004.03.009

Guo, Z., Li, B., Dzubiella, J., Cheng, L. T., McCammon, J. A., and Che, J. (2013). Evaluation of hydration free energy by level-set variational implicit solvent model with coulomb-field approximation. *J. Chem. Theor. Comput.* 9, 1778–1787. doi: 10.1021/ct301087w

Guo, Z., Li, B., Dzubiella, J., Cheng, L.-T., McCammon, J. A., and Che, J. (2014). Heterogeneous hydration of p53-MDM2 complex. *J. Chem. Theor. Comput.* 9, 1302–1313. doi: 10.1021/ct400967m

Hua, L., Zangi, R., and Berne, B. J. (2009). Hydrophobic interactions and dewetting between plates with hydrophobic and hydrophilic domains. *J. Chem. J. Phys. Chem. C* 113, 5244–5253. doi: 10.1021/jp8088758

Huang, Q., Ding, S., Hua, C. Y., Yang, H. C., and Chen, C. L. (2004). A computer simulation study of water drying at the interface of protein chains. *J. Chem. Phys.* 121, 1969–1977. doi: 10.1063/1.1766017

Huang, X., Margulis, C. J., and Berne, B. J. (2003). Dewetting-induced collapse of hydrophobic particles. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11953–11958. doi: 10.1073/pnas.1934837100

Huang, X., Zhou, R., and Berne, B. J. (2005). Drying and hydrophobic collapse of paraffin plates. *J. Phys. Chem. B* 109, 3546–3552. doi: 10.1021/jp045520l

Hummer, G. (2010). Molecular binding: under water's influence. *Nat. Chem.* 2, 906–907. doi: 10.1038/nchem.885

Jeanmairet, G., Levesque, M., and Borgis, D. (2013). Molecular density functional theory of water describing hydrophobicity at short and long length scales. *J. Chem. Phys.* 139:154101. doi: 10.1063/1.4824737

Jensen, T. R., Østergaard Jensen, M., Reitzel, N., Balashev, K., Peters, G. H., Kjaer, K., et al. (2003). Water in contact with extended hydrophobic surfaces: direct evidence of weak dewetting. *Phys. Rev. Lett.* 90:086101. doi: 10.1103/PhysRevLett.90.086101

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935. doi: 10.1063/1.445869

Krimmer, S. G., Cramer, J., Schiebel, J., Heine, A., and Klebe, G. (2017). How nothing boosts affinity: hydrophobic ligand binding to the virtually vacated S1' pocket of thermolysin. *J. Am. Chem. Soc.* 139, 10419–10431. doi: 10.1021/jacs.7b05028

Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., et al. (1996). Structure of MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* 274, 948–953. doi: 10.1126/science.274.5289.948

Lee, L. P., and Tidor, B. (2001). Barstar is electrostatically optimized for tight binding to barnase. *Nat. Struct. Biol.* 8, 73–75. doi: 10.1038/nsb0901-789

Levy, Y., and Onuchic, J. N. (2006). Water mediation in protein folding and molecular recognition. *Annu. Rev. Biophys. Biomol. Struct.* 35, 389–415. doi: 10.1146/annurev.biophys.35.040405.102134

Liu, P., Huang, X., Zhou, R., and Berne, B. J. (2005). Observation of a dewetting transition in the collapse of the melittin tetramer. *Nature* 437, 159–162. doi: 10.1038/nature03926

Lum, K., Chandler, D., and Weeks, J. D. (1999). Hydrophobicity at small and large scales. *J. Phys. Chem. B* 103, 4570–4577. doi: 10.1021/jp984327m

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). Ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theor. Comput.* 11, 3696–3713. doi: 10.1021/acs.jctc.5b00255

Marrink, S. J., de Vries, A. H., and Mark, A. E. (2004). Coarse grained model for semiquantitative lipid simulations. *J. Phys. Chem. B* 108, 750–760. doi: 10.1021/jp036508g

Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P., and de Vries, A. H. (2007). The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* 111, 7812–7824. doi: 10.1021/jp071097f

Matthews, B. W., and Liu, L. (2009). A review about nothing: are apolar cavities in proteins really empty? *Protein Sci.* 18, 494–502. doi: 10.1002/pro.61

Monticelli, L., Kandasamy, S. K., Periole, X., Larson, R. G., Tieleman, D. P., and Marrink, S. J. (2008). The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theor. Comp.* 4, 819–834. doi: 10.1021/ct700324x

Parker, J. L., Claesson, P. M., and Attard, P. J. (1994). Bubbles, cavities, and the long-ranged attraction between hydrophobic surfaces. *Phys. Chem.* 98, 8468–8480. doi: 10.1021/j100085a029

Poynor, A., Hong, L., Robinson, I. K., Granick, S., Zhang, Z., and Fenter, P. A. (2006). How water meets a hydrophobic surface. *Phys. Rev. Lett.* 97:266101. doi: 10.1103/PhysRevLett.97.266101

Qvist, J., Davidovic, M., Hamelberg, D., and Halle, B. (2008). A dry ligand-binding cavity in a solvated protein. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6296–6301. doi: 10.1073/pnas.0709844105

Ramirez, R., and Borgis, D. (2005). Density functional theory and its relation to implicit solvent models. *J. Phys. Chem. B.* 109, 6754–6763. doi: 10.1021/jp045453v

Ricci, C. G., Li, B., Cheng, L. T., Dzubiella, J., and McCammon, J. A. (2017). "Martinizing" the variational implicit solvent method (VISM): solvation

free energy for coarse-grained proteins. *J. Phys. Chem. B* 121, 6538–6548. doi: 10.1021/acs.jpcb.7b04113

Roux, B., and Simonson, T. (1999). Implicit solvent models. *Biophys. Chem.* 78, 1–20. doi: 10.1016/S0301-4622(98)00226-9

Schmalhorst, P. S., Deluweit, F., Scherrers, R., Heisenberg, C. P., and Sikora, M. (2017). Overcoming the limitations of the MARTINI force field in simulations of Polysaccharides. *J. Chem. Theory Comput.* 13, 5039–5053. doi: 10.1021/acs.jctc.7b00374

Schwendel, D., Hayashi, T., Dahint, R., Pertsin, A., Grunze, M., Steitz, R., et al. (2003). Interaction of water with self-assembled monoloayers: neutron reflectivity measurement of the water density in the interface region. *Langmuir* 19, 2284–2293. doi: 10.1021/la026716k

Setny, P. (2007). Water properties and potential of mean force for hydrophobic interactions of methane and nanoscopic pockets studied by computer simulations. *J. Chem. Phys.* 127:054505. doi: 10.1063/1.2749250

Setny, P. (2008). Hydrophobic interactions between methane and a nanoscopic pocket: three dimensional distribution of potential of mean force revealed by computer simulations. *J. Chem. Phys.* 128:125105. doi: 10.1063/1.2839885

Setny, P., Baron, R., and McCammon, J. A. (2010). How can hydrophobic association be enthalpy driven? *J. Chem. Theor. Comput.* 6, 2866–2871. doi: 10.1021/ct1003077

Setny, P., and Geller, M. (2006). Water properties inside nanoscopic hydrophobic pocket studied by computer simulations. *J. Chem. Phys.* 125:144717. doi: 10.1063/1.2355487

Setny, P., Wang, Z., Cheng, L. T., Li, B., McCammon, J. A., and Dzubiella, J. (2009). Dewetting-controlled binding of ligands to hydrophobic pockets. *Phys. Rev. Lett.* 103:187801. doi: 10.1103/PhysRevLett.103.187801

Stark, A. C., Andrews, C. T., and Elcock, A. H. (2013). Toward optimized potential functions for protein-protein interactions in aqueous solutions: osmotic second virial coefficient calculations using the MARTINI coarse-grained force field. *J. Chem. Theor. Comp.* 9, 4176–4185. doi: 10.1021/ct4000 08p

Steitz, R., Gutberlet, T., Gauss, H., Klösgen, B., Krastev, R., Schemmel, S., et al. (2003). Nanobubbles and their precursor layer at the interface of water against a hydrophobic substrate. *Langmuir* 19, 2409–2418. doi: 10.1021/la026731p

Still, W. C., Tempczyk, A., Hawley, R. C., and Hendrickson, T. (1990). Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* 112, 6127–6129. doi: 10.1021/ja00172a038

ten Wolde, P. R., and Chandler, D. (2002). Drying-induced hydrophobic polymer collapse. *Proc. Natl. Acad. Sci. U.S.A.* 99, 6539–6543. doi: 10.1073/pnas.052153299

Tyrrell, J. W., and Attard, P. (2001). Images of nanobubblers on hydrophobic surfaces and their interactions. *Phys. Rev. Lett.* 87:176104. doi: 10.1103/PhysRevLett.87.176104

Wallqvist, A., Gallicchio, E., and Levy, R. M. (2001). A model for studying drying at hydrophobic interfaces: structural and thermodynamic properties. *J. Phys. Chem. B* 105, 6745–6753. doi: 10.1021/jp010945i

Wang, L., Berne, B. J., and Friesner, R. A. (2011). Ligand binding to protein-binding pockets with wet and dry regions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 1326–1330. doi: 10.1073/pnas.1016793108

Wang, Z., Che, J., Cheng, L. T., Dzubiella, J., Li, B., and McCammon, J. A. (2012). Level-set variational implicit-solvent modeling of biomolecules with the coulomb-field approximation. *J. Chem. Theor. Comp.* 8, 386–397. doi: 10.1021/ct200647j

Young, T., Abel, R., Kim, B., Berne, B. J., and Friesner, R. A. (2007). Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc. Natl. Acad. Sci. U.S.A.* 104, 808–813. doi: 10.1073/pnas.0610202104

Zhao, S., Ramirez, R., Vuilleumier, R., and Borgis, D. (2011). Molecular density theory of solvation: from polar solvents to water. *J. Chem. Phys.* 134:194102. doi: 10.1063/1.3589142

Zhao, Y., Bernard, D., and Wang, S. (2013). Small molecule inhibitors of MDM2-p53 and MDMX-p53 interaction as new cancer therapeutics. *BioDiscovery* 8, 1–15. doi: 10.7750/BioDiscovery.2013.8.4

Zhou, R., Huang, X., Margulis, C. J., and Berne, B. J. (2004). Hydrophobic collapse in multidomain protein folding. *Science* 305, 1605–1609. doi: 10.1126/science.1101176

Zhou, S., Cheng, L. T., Dzubiella, J., Li, B., and McCammon, J. A. (2014). Variational implicit solvation with Poisson-Boltzmann theory. *J. Chem. Theory Comput.* 10, 1454–1467. doi: 10.1021/ct401058w

Zhou, S., Cheng, L. T., Sun, H., Che, J., Dzubiella, J., Li, B., et al. (2015). LS-VISM: a software package for analysis of biomolecular solvation. *J. Comput. Chem.* 36, 1047–1059. doi: 10.1002/jcc.23890