THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

"Martinizing" the Variational Implicit Solvent Method (VISM): Solvation Free Energy for Coarse-Grained Proteins

Clarisse G. Ricci,*[,†] Bo Li,[‡] Li-Tien Cheng,[‡] Joachim Dzubiella,[§] and J. Andrew McCammon[†]

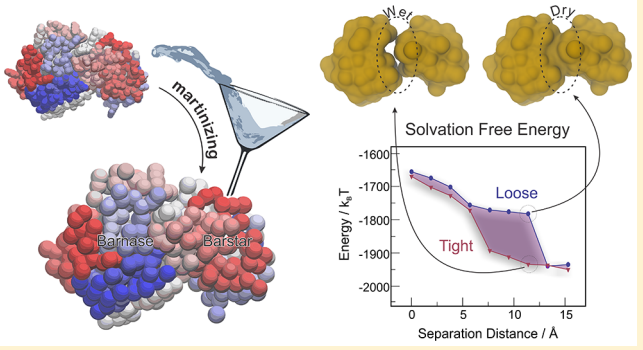
[†]Department of Pharmacology and Department of Chemistry & Biochemistry, Howard Hughes Medical Institute, National Biomedical Computation Resource, University of California at San Diego, La Jolla, California 92093, United States

[‡]Department of Mathematics and Quantitative Biology Graduate Program, University of California at San Diego, La Jolla, California 92093, United States

[§]Institut für Physik, Humboldt-Universität zu Berlin, D-12849, Berlin, Germany, and Soft Matter and Functional Materials, Helmholtz-Zentrum Berlin, D-14109, Berlin, Germany

Supporting Information

ABSTRACT: Solvation is a fundamental driving force in many biological processes including biomolecular recognition and self-assembly, not to mention protein folding, dynamics, and function. The variational implicit solvent method (VISM) is a theoretical tool currently developed and optimized to estimate solvation free energies for systems of very complex topology, such as biomolecules. VISM's theoretical framework makes it unique because it couples hydrophobic, van der Waals, and electrostatic interactions as a functional of the solvation interface. By minimizing this functional, VISM produces the solvation interface as an output of the theory. In this work, we push VISM to larger scale applications by combining it with coarse-grained solute Hamiltonians adapted from the MARTINI framework, a well-established mesoscale force field for modeling large-scale biomolecule assemblies. We show how MARTINI-VISM ([M]VISM) compares with atomistic VISM ([A]VISM) for a small set of proteins differing in size, shape, and charge distribution. We also demonstrate [M]VISM's suitability to study the solvation properties of an interesting encounter complex, barnase−barstar. The promising results suggest that coarse-graining the protein with the MARTINI force field is indeed a valuable step to broaden VISM's and MARTINI's applications in the near future.



INTRODUCTION

Solvation, or more specifically hydration, is a fundamental driving force in essentially all biological processes. Water and ions, and their interactions, play a huge role in protein folding, dynamics, and function,[1−3] self-assembly of membranes and proteins,[4] and molecular recognition and binding,[2,5] to name only a few. Many computational methods attempt to provide an accurate description of the solvation effects underlying such processes. The most efficient among these are the so-called dielectric boundary implicit solvation methods. By representing the solvent environment as a continuum medium separated from the solute by a well-defined solvation boundary, they avoid the recurrent issue of statistical sampling of solvent configurations. Instead, the average forces and thermal fluctuations giving rise to the polarity of the solvent are implicitly represented by a uniform dielectric constant.[6]

The majority of implicit solvation methods are based on electrostatics—described by the Poisson−Boltzmann[7−9] or generalized Born theory[10−12]—with hydrophobic, van der Waals, and first-shell solvation effects taken into account by means of empirical relations between solvation free energy and accessible surface area, also known as SASA (solvent accessible surface area) methods.[6,13] While successful in many cases, their performance is ultimately limited by the fact that these methods rely on fixed and pre-established solute−solvent interfaces—normally guessed as van der Waals-based surfaces. Therefore, they fail to capture subtle though important behavior deriving from heterogeneous solvation patterns, such as polymodal hydration and dewetting,[14] which are observed in explicit solvent simulations.[4,15,16] Moreover, Poisson−Boltzmann calculations are extremely sensitive to the chosen dielectric boundary,[17] so that a poorly guessed interface can lead to very significant errors.

The variational implicit solvent method (VISM) is a solvation free energy method that avoids guessing an *a priori* dielectric solvation boundary. Instead, VISM expresses the solvation contributions from hydrophobic effects, van der Waals interactions, and electrostatics as a functional of all possible solute−solvent interfaces (see eq 1).[18,19] VISM's variational formulation makes it possible to identify the

Received: May 1, 2017
Revised: June 13, 2017
Published: June 14, 2017

DOI: 10.1021/acs.jpcb.7b04113
J. Phys. Chem. B 2017, 121, 6538−6548

dielectric interface—or solvation state—that minimizes the solvation free energy, producing the optimal interface as an output of the theory. Moreover, the surface energy, van der Waals interactions, and electrostatics are not independent but coupled together through the solvation interface. In this way, sophisticated hydrophilic−hydrophobic compensation effects can be captured.[20,21]

VISM is currently developed to predict solvation free energies for systems of complex topology by means of a robust level-set method,[21−23] combined with a user-defined choice of the electrostatic formulation that varies from the Coulomb field approximation (CFA)[24,25] to the nonlinear Poisson−Boltzmann (PB) theory including ionic effects.[20] Besides predicting accurate solvation free energies of small solutes and small globular proteins, several studies have highlighted VISM's unique ability in capturing nontrivial solvation effects such as capillary evaporation,[20,21] dry−wet transitions in ligand binding events,[25−29] dewetting of druggable binding cavities,[30] and, since it is based on a true Hamiltonian approach, being extended to solvent equilibrium fluctuations.[31]

Given the proven potential and performance of VISM in describing the complex solvation phenomena in relatively small systems, it is urgently needed now to broaden its application range toward larger scale processes, such as multiprotein association and self-assembly. For modeling consistency and a not too large scale-separation, VISM can thus be connected with efficient coarse-grained (CG) models, desirably those that are efficiently coarse but keep chemical specificity and have been developed already for the use in a coarse-grained or implicit solvent environment. A popular state-of-the art CG model of that kind is provided by the MARTINI family of models.

The MARTINI force field (FF) is a well-established coarse-grained model originally developed to simulate lipids and surfactants in lipid bilayers[32,33] and later extended to proteins and other biomolecules.[34−38] Its parametrization philosophy strongly relies on reproducing water/oil or water/membrane partition coefficients. Although the majority of MARTINI applications to proteins involve membrane-embedded or membrane-anchored proteins,[39] there are a few studies focused on water-soluble proteins.[40,41] Among these, Stark et al. found that MARTINI tends to overestimate protein−protein interactions in aqueous environment, but the discrepancy to experiments could be counteracted by a simple downscaling of the Lennard-Jones energy parameter, ε.[40] Originally, MARTINI was developed for the use with a CG water model. The issues and performance problems of the latter have been discussed and a "dry" Martini version with an efficient, fully implicit water has been introduced[42] by adapting the nonbonded parameter matrix. Still, this implicit-solvent model in dry-Martini has the same limitations as the traditional (pre-VISM) implicit solvents discussed above.

In this work, we combined the current best of two worlds and adapted VISM to work with a for our purposes amended MARTINI FF to construct an efficient and quantitative model to attack new problems with VISM, such as biomolecular assembly on large scales. As a first demonstration, we predict the solvation free energies of some selected proteins. We tested MARTINI-VISM (^{M}VISM) against atomistic VISM (^{A}VISM) for a set of six proteins differing in shape, size, and charge distribution. With a few adjustments and simple corrections in the Lennard-Jones parameters, ^{M}VISM is capable of reproducing atomistic solvation free energies to within 10% for 5 out of 6 proteins, besides producing reasonable electrostatic potentials at the interacting surfaces of the proteins. Finally, both ^{M}VISM and ^{A}VISM seem to correctly predict the nature of dry−wet transitions in the encounter of barnase and barstar. These results pave the way for using ^{M}VISM to address more challenging problems such as solvation dynamics and many-protein interactions, in the near future.

THEORY AND METHODS

The Free Energy Functional. To estimate the solvation free energy of a protein (or other biomolecule) in the aqueous environment (water + ions), we start by dividing the system into three regions, as displayed in Figure 1. The solute region,

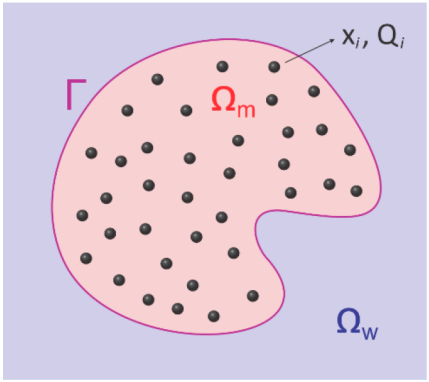


Figure 1. Schematic view of a solvation system with an implicit solvent. A solute−solvent interface, Γ, separates the solvent region, Ω_W, from the solute region, Ω_m. The solute atoms are located at x_1, ..., x_N and carry point charges Q_1, ..., Q_N, respectively.

Ω_m, contains all of the N atoms belonging to the solute molecule, which are located at x_1, ..., x_N inside Ω_m and carry point charges Q_1, ..., Q_N, respectively. The solvent region, Ω_w, is treated implicitly as a continuum. The third part consists of the solute−solvent interface, Γ, which geometrically separates the regions Ω_m and Ω_w.

In VISM, the solvation free energy is expressed as a functional of the solvation interface, Γ (eq 1). By minimizing this functional against all possible solvation interfaces, one can find the local minima that correspond to the stable hydration states of the system.

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

The first term in eq 1 is purely geometrical and accounts for the hydrophobic effect by integrating the surface tension along the solvation interface. Because for systems of nanometer scale the surface tension strongly depends on the curvature, we define the local surface tension as

$$\gamma = \gamma_0(1 - 2\tau H) \tag{2}$$

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, which essentially accounts for the relative size of the solvent molecules with respect to the solute local curvature.[18]

DOI: 10.1021/acs.jpcb.7b04113
J. Phys. Chem. B 2017, 121, 6538−6548

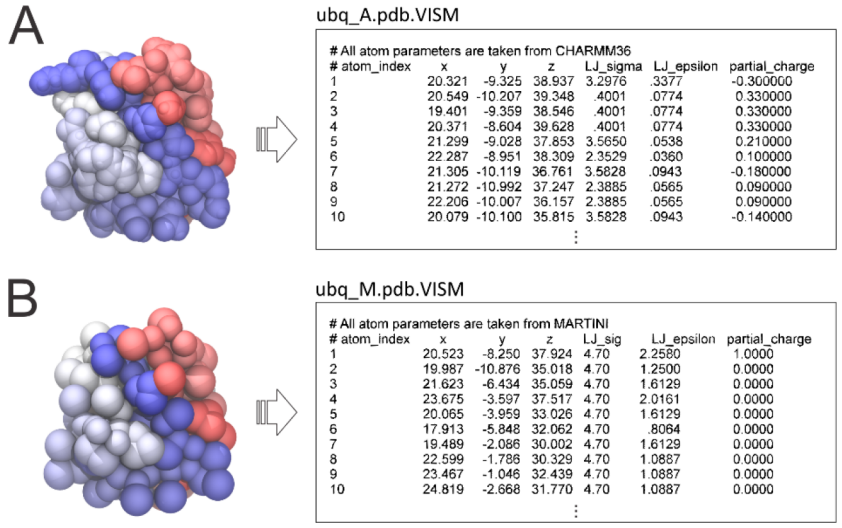


Figure 2. Atomistic (A) and MARTINI (B) representation of ubiquitin, with corresponding VISM input files. For the sake of clarity, only parts of the input files are shown, with σ and ε in Å and k_BT, respectively.

Table 1. Test Set of Proteins

system	surface area[a] (Å^2)	heavy atoms/beads	largest cross-section (Å)	net charge	charged + polar residues[b] (%)	PDB ID	reference
E3 ligase	318	350/93	~25	−2	33.3 + 0 = 33.3	2OOB	Peschard et al.[45]
ubiquitin	413	574/156	~25	−1	16.6 + 25.0 = 41.6	2OOB	
barstar	487	699/195	~30	−6	33.3 + 16.6 = 50.0	1BRS	Buckle et al.[46]
barnase	592	839/248	~35	+2	50.0 + 33.3 = 83.3	1BRS	
BLIP	795	1235/351	~40	−2	23.3 + 34.6 = 60.0	1JTG	Lim et al.[47]
β-lactamase	1092	2022/550	~50	−6	29.0 + 29.0 = 60.0	1JTG	

[a]Obtained from [A]VISM calculations starting from tight initials. [b]At the binding interfaces.

The second term in the functional accounts for dispersion attraction and Pauli repulsion, or van der Waals interactions between solute and solvent. For each atom type of the solute, we define a pairwise interaction energy with the solvent, U_i. As traditionally used in molecular dynamics (MD) simulations, U_i is modeled by a Lennard-Jones (LJ) function (eq 3):

$$U_i = 4\epsilon_i\left[\left(\frac{\sigma_i}{r}\right)^{12} - \left(\frac{\sigma_i}{r}\right)^{6}\right] \tag{3}$$

For the ith atom of the solute, located at x_i, the interaction energy is calculated as a volume integral over the solvent region, which is coarse-grained represented by grid points located at x, with a density prefactor, ρ_w.

Finally, the third term of the functional, $G_{elec}(\Gamma)$, is the electrostatic part of the solvation free energy. In the current VISM implementation, this term can be calculated with CFA or PB theory (in its linearized or nonlinear forms). Detailed descriptions of this term are given in ref 24 (for CFA) or ref 21 (for PB).

VISM Calculations. Atomistic protein structures were obtained from the Protein Data Bank,[43] and hydrogens were added. VISM input files were generated by combining the crystallographic atomic coordinates with LJ parameters and point charges borrowed from the CHARMM36 FF[44] (Figure 2A). LJ parameters for water were obtained from the TIP3P model and combined with the solute parameters by using the standard Lorentz−Berthelot rules:

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \tag{4}$$

$$\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \tag{5}$$

Throughout our [A]VISM calculations, we fixed T = 298 K, γ_0 = 0.1315 k_BT/Å^2, τ = 1.0 Å, ρ_w = 0.0333 Å^{-3}, ϵ_w = 78.0, and ϵ_m = 2.0.[21]

The electrostatic component of the free energy functional was calculated with CFA theory during the largest part of the minimization procedure and then refined in the last steps with linearized PB theory. We included +1 and −1 ionic species at a concentration of 0.1 M. We also pulled the final VISM interface closer to the solute molecules by δ = 1 Å (or δ = 2.5 Å, in [M]VISM calculations) before calculating the final electrostatic energies. This approach was proved to significantly improve the accuracy of electrostatic VISM energies in previous studies.[20]

VISM Calculations with MARTINI. For [M]VISM calculations, we submitted the atomistic PDB files to the martinize.py script with MARTINI 2.1,[34] which produced the corresponding coarse-grained PDB files. We then generated VISM input files by combining the bead coordinates from the coarse-grained PDB files with LJ parameters and point charges from MARTINI 2.1 (Figure 2B). Because MARTINI provides the Lennard-Jones parameters in the precombined form (U_{ij}), we had to adapt the VISM algorithm to skip the LJ combination rules (see eqs 4 and 5). We also used MARTINI's LJ potential in its shifted form from r_{shift} = 9 Å to r_{cut} = 12 Å, smoothly vanishing to zero beyond r_{cut}. Since one MARTINI water bead occupies the volume of approximately four TIP3P water molecules (data obtained from test simulations with MARTINI water and TIP3P water, at 298 K), we fixed ρ_w = 0.0333/4 = 0.008325 Å^{-3} in our [M]VISM calculations. Unless

DOI: 10.1021/acs.jpcb.7b04113
J. Phys. Chem. B 2017, 121, 6538−6548

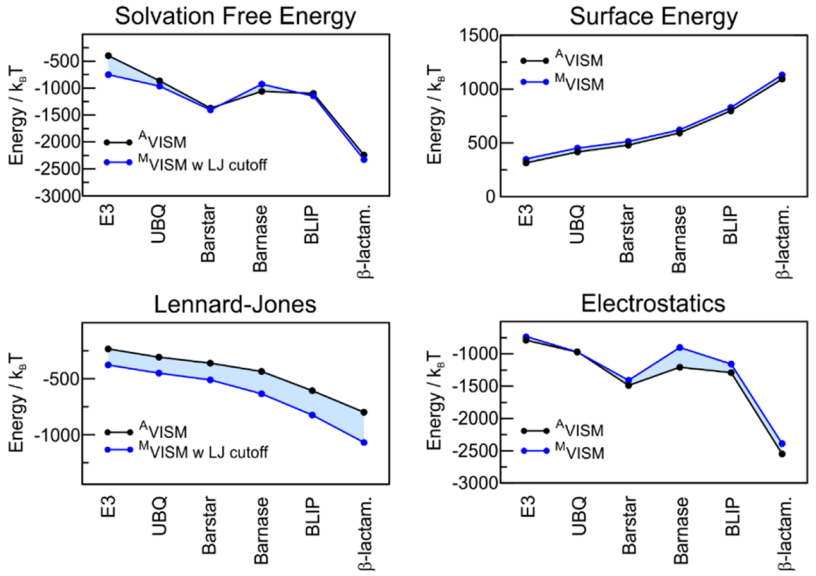


Figure 3. Solvation free energy and correspondent surface, Lennard-Jones, and electrostatic components computed with ^{A}VISM (in black) or ^{M}VISM (in blue) for six different proteins, distributed according to protein size (ascendant order).

specified otherwise, all remaining parameters in ^{M}VISM calculations were the same used in the ^{A}VISM calculations.

In both ^{A}VISM and ^{M}VISM calculations, the grid resolution ranged from approximately 0.3 to 0.5 Å, depending on the size of the protein (~0.3 for ubiquitin, E3, and barstar; ~0.4 Å for barnase and BLIP; and ~0.5 Å for β-lactamase). Each calculation was performed in a single CPU, and calculation times ranged from less than 1 h to ~32 h depending on the system size/geometry and the choice of the initial surfaces. Calculations starting from tight initial surfaces tend to converge significantly faster than calculations that start from loose initial surfaces (check ref 21 for more details).[21]

RESULTS AND DISCUSSION

To test and adjust ^{M}VISM performance, we chose a test set of six proteins differing in size, shape, and charge distribution (Table 1). These proteins also form pairwise complexes and do not undergo large conformational changes upon binding. We started by performing ^{M}VISM calculations and comparing the resulting solvation energies with those from ^{A}VISM calculations.

Figure 3 reports the solvation free energy and its decomposition in surface (hydrophobic), Lennard-Jones (van der Waals), and electrostatic components, for the six individual proteins, distributed according to their size (in ascendant order). As expected, both the surface and the LJ components correlate well with protein size, as they are both dependent on the protein surface area. We also found that ^{M}VISM calculations correctly capture the overall trend in the solvation free energy and in its energy components but tend to systematically overestimate the magnitude of the LJ energies by ~40% with respect to the atomistic calculations.

Overestimation of LJ interactions is a known trait of the of MARTINI FF[40] and likely a deliberate way of compensating (enthalpically) for the loss of degrees of freedom—and consequent loss of entropically driven hydrophobic effects—in MD simulations with the coarse-grained model.[39] In VISM, since hydrophobic effects are taken into account implicitly in the geometric term of the free energy functional, we feel justified to downscale the MARTINI-based LJ interactions in ^{M}VISM.

Downscaling LJ Interactions by Tuning the ε Parameter. We chose to tune the LJ parameter, ε, by interpolating the well-depth of the LJ potential between (i) the value set in the original MARTINI FF and (ii) the value employed by MARTINI to describe nominally "repulsive" interactions, as originally proposed by Stark et al.[40] The degree of scaling is controlled by a scaling parameter, α, with $\alpha = 1$ corresponding to the original MARTINI FF and $\alpha = 0$ corresponding to the use of the weakest possible bead–bead interaction type for all bead–bead interactions (for which MARTINI uses $\varepsilon = 2.0$ kJ mol^{-1}). For a given value of α, the new ε_α is determined using

$$\varepsilon_\alpha = \varepsilon_{\text{repulsive}} + \alpha(\varepsilon_{\text{original}} - \varepsilon_{\text{repulsive}}) \quad (6)$$

Calculations performed with different α values reveal that a good agreement between atomistic and coarse-grained LJ energies is obtained with $\alpha = 0.5$ (Figure 4A and Table S1). Interestingly, the significant improvement in the Lennard-Jones energies does not systematically improve the solvation free energy for all proteins, as shown in Figure 4B. This is a direct effect of VISM's theoretical formulation, which allows for different energy components to couple through the solvation interface. Indeed, a more detailed analysis of each energy term reveals a compensation between the LJ and electrostatic components, such that downscaling the coarse-grained LJ interactions concomitantly allows for the coarse-grained electrostatic energies to become more negative and closer to atomistic values (Figure 4C). Overall, downscaling the LJ interactions in ^{M}VISM not only improves the description of the LJ energies, but it also produces a more accurate distribution of the solvation free energy among each of its energy components. We thus decided to fix $\alpha = 0.5$ in all subsequent ^{M}VISM calculations.

The corresponding solvation surfaces produced by VISM are illustrated in Figure 5, along with the original molecular complexes. The coarse-grained interfaces strongly resemble the atomistic ones, although they are slightly larger and less curved

DOI: 10.1021/acs.jpcb.7b04113
J. Phys. Chem. B 2017, 121, 6538–6548

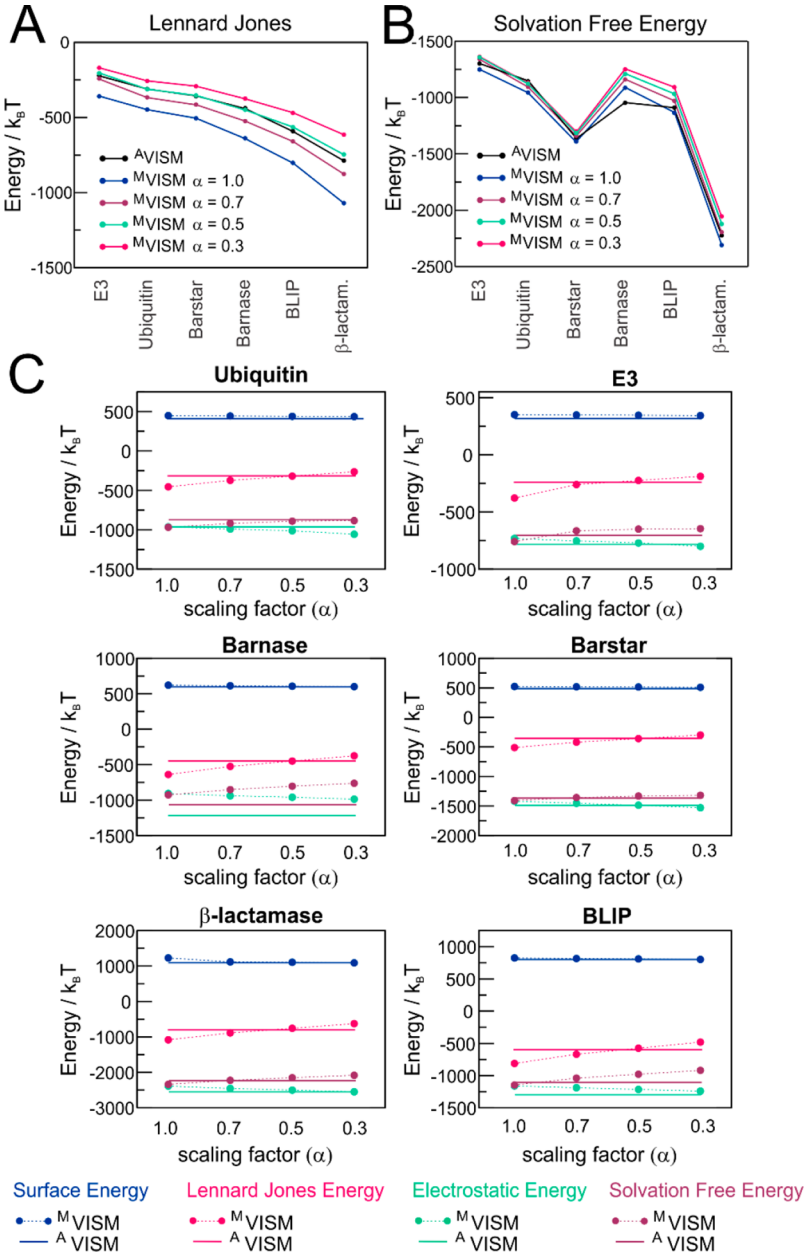


Figure 4. Lennard-Jones (A) and solvation free energies (B) for ^{M}VISM calculations (colored lines) with different levels of ε scaling in comparison with ^{A}VISM (in black). (C) Detailed analysis of the effects of downscaling ε on each energy term. The energies depicted as dashed lines correspond to the coarse-grained energies with different levels of ε scaling. For comparison, the corresponding atomistic energies are depicted as solid lines.

due to the size and shape of MARTINI beads. Because in VISM one minimizes the free energy functional with respect to the solvation interface, the final results depend on the initial guess of the solvation surface. To account for different solvation states, it is common to start VISM calculation from tight and loose surface initials.[20,21,24,26] A loose initial corresponds to a large surface loosely encompassing all of the solute atoms, while a tight initial corresponds to a van der Waals surface of the solute atoms. In this way, VISM calculations starting from loose initials are more likely to produce "dry" states—with water expulsion near hydrophobic patches—whereas tight initials tend to produce "wet", fully solvated states. For the six proteins tested, the solvation energies obtained from tight or loose initials are extremely similar, in both ^{A}VISM and ^{M}VISM calculations (Figure S1 and Table S2), as are the corresponding solvation interfaces (Figures S2, S3, and S4). Therefore, these proteins do not appear to display multiple hydration states in the apo-state.

The Curvature Correction Coefficient, τ. The microscopic curvature correction coefficient (τ) is a parametric coefficient used in the geometric, or hydrophobic, part of the VISM functional (see eqs 1 and 2). It implicitly accounts for how sensitive the organization of the solvent molecules is with respect to the local curvature of the solvation surface, which is related to the solvent size. In atomistic VISM calculations, τ is typically set to 1 Å.[21] Since combining groups of atoms into beads affects the protein local curvatures, we next investigate how adjustments in τ can affect the ^{M}VISM performance with respect to atomistic calculations.

Several ^{A}VISM and ^{M}VISM calculations were performed with τ values ranging from 0.5 to 2.0 Å, using the ubiquitin and barnase proteins as test cases. As expected, we found that the τ

DOI: 10.1021/acs.jpcb.7b04113
J. Phys. Chem. B 2017, 121, 6538−6548

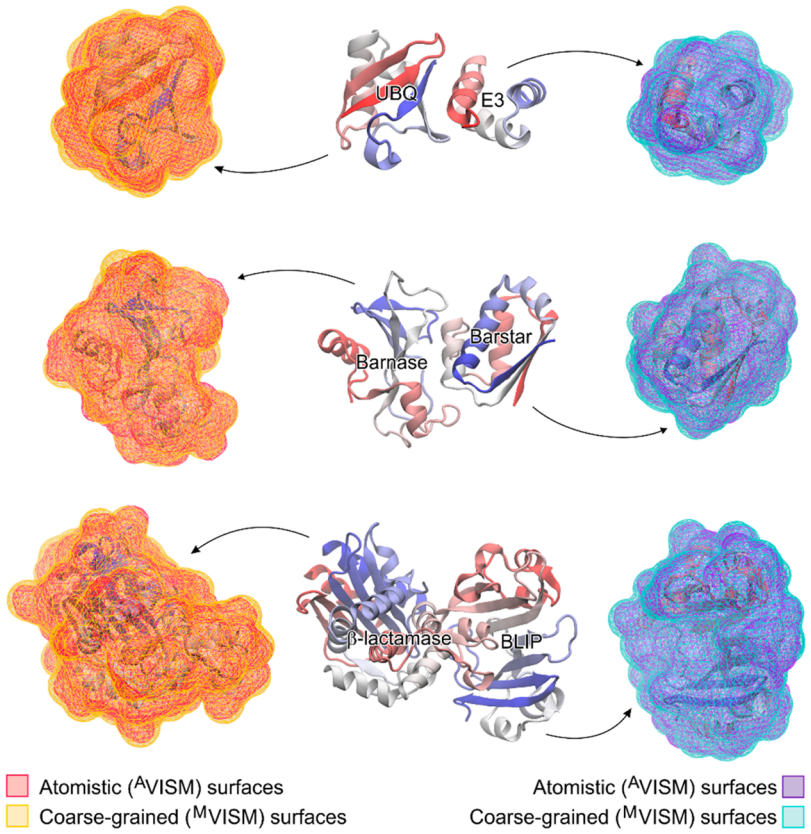


Figure 5. Atomistic (magenta and violet) or coarse-grained (yellow and cyan) solvation surfaces obtained with ^{A}VISM and ^{M}VISM, respectively. These calculations started from loose initial surfaces.

parameter mainly affects the surface term, with small concomitant changes in the Lennard-Jones or electrostatic terms (<20 k_BT, not shown). For both proteins, the surface energies decrease with increasing τ (Figure 6A). This is consistent with the fact that globular proteins display predominantly convex surfaces, where a larger τ decreases the local surface tension and, therefore, also the surface energy. In concave regions—such as binding pockets—a larger τ has the opposite effect of increasing the local surface tension and therefore enhancing the hydrophobic effect.[27] As a consequence, larger τ values tend to produce solvation surfaces with less pronounced pockets. This effect is illustrated in Figure 6, which shows the barstar binding cavity located in the barnase protein, in atomistic (B) and coarse-grained (C) representations. Comparison of the surfaces obtained with small (0.5 Å) or large (2.0 Å) τ shows that the more pronouncedly concave regions of the binding cavity become shallower with larger τ values (see, for instance, the cyan patch in Figure 6C). A physical interpretation for this is that a larger τ makes the binding cavities more hydrophobic, with a higher tendency to expel water (dewetting effect).

Quantitative analysis of the surface energies shows that, in order to match the ^{M}VISM surface energies to the atomistic (^{A}VISM) energies obtained with τ = 1 Å, one should use slightly larger τ values. In the case of ubiquitin, a good agreement is obtained with τ = 1.5 Å, whereas, for barnase, a good agreement is obtained with τ = 1.2 Å (see dotted arrows in Figure 6A). Since barnase is larger than ubiquitin, our results suggest that the need to adjust τ might decrease with protein size. For the remaining ^{M}VISM calculations in this work, we used τ = 1 Å, for simplicity.

Electrostatic Energies and the Shifting Parameter, δ. Besides affecting the Lennard-Jones and surface energies, coarse-graining the protein with MARTINI 2.1 also affects the electrostatic energies, since it rearranges atomic partial charges into unit integer negative or positive bead charges (+1 or −1). We next investigated to which extent coarse-graining the charges impacts the electrostatic potential at binding interfaces. Figure 7 shows the electrostatic potential at the binding interfaces of barnase and barstar obtained with ^{A}VISM (A) or ^{M}VISM (B). Clearly, the electrostatic potential produced by ^{M}VISM is a coarse-grained one as compared to the atomistic, as would be expected, since the nonpolarizable version of MARTINI does not use partial charges. However, the overall electrostatic complementarity is maintained, with a large negatively charged patch in barstar matching a largely positive binding cavity in barnase. Similar results were obtained with ubiquitin−E3 and β-lactamase−BLIP complexes (Figures S5 and S6, respectively), which is reassuring for future applications of ^{M}VISM in protein binding and assembly.

Martinizing the protein also has the effect of burying the point charges deeper in the solute region than they would be in atomistic representations of the proteins, due to the larger size of the beads and the location of the charge sites in MARTINI 2.1 beads. This causes the electrostatic attraction between protein charges and the (polar) solvent medium to decrease, leading to overestimated (less negative) electrostatic energies. One way to counteract this artifact is by simply adjusting the final shift (δ) that is applied to the solvation boundary, prior to the calculation of the final PB electrostatic energies. In atomistic VISM calculations, a 1 Å shift of the final interface toward the protein has been proved to favorably account for the

DOI: 10.1021/acs.jpcb.7b04113
J. Phys. Chem. B 2017, 121, 6538−6548

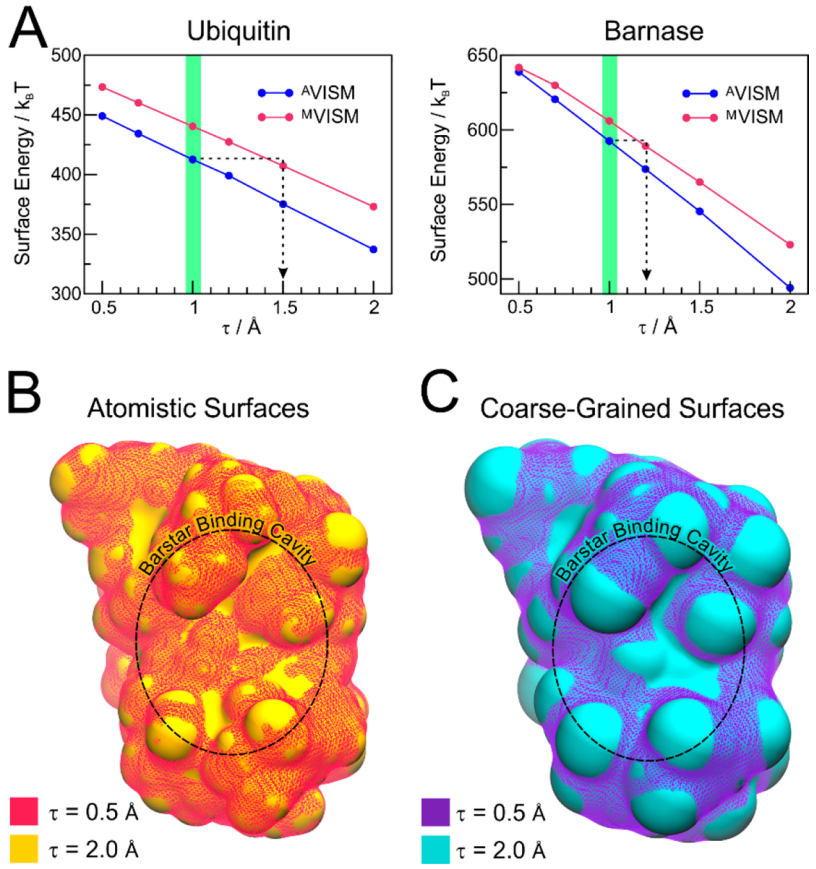


Figure 6. Effect of the curvature correction coefficient on the surface energies and solvation interfaces. (A) Surface energies with different τ values for ubiquitin (left) and barnase (right). (B) Atomistic solvation interfaces obtained with ^{A}VISM, using τ = 0.5 Å (magenta) or τ = 2.0 Å (yellow) for barnase. (C) Coarse-grained solvation interfaces obtained with ^{M}VISM, using τ = 0.5 Å (purple) or τ = 2.0 Å (cyan) for barnase.

fact that the SASA-like VISM interface is slightly different from the dielectric boundary interface.[20,21] In ^{M}VISM, this parameter can also be tuned to counteract the deeper burial of point charges in MARTINI 2.1. We thus report the electrostatic energies obtained with ^{M}VISM for the six proteins, using different shifting values (Figure 8). As we increase the shifting of the VISM interface toward the solute atoms, the electrostatic energies become more negative, reflecting the favorable interaction between charged solute groups and the aqueous environment. As shown in Figure 8, a good agreement between coarse-grained and atomistic electrostatic energies is systematically obtained by performing ^{M}VISM calculations with a slightly larger shift of 2.5 Å.

It is worth noting that there is still room to further improve ^{M}VISM electrostatic energies (and corresponding electrostatic potentials) by employing more sophisticated versions of the MARTINI force field. The polarizable MARTINI 2.2P, for instance, can improve the electrostatic description of polar proteins by including auxiliary partially charged sites for polar residues.[48] Whether ^{M}VISM calculations of highly polar proteins such as barnase could benefit from the polarizable MARTINI 2.2P remains to be tested, but it is certainly an interesting direction should one need more accurate electrostatics in ^{M}VISM applications.

The Parametrized ^{M}VISM Method. After parametrization, we suggest applying the ^{M}VISM method with (i) σ and (2-fold downscaled) ε Lennard-Jones parameters from the MARTINI 2.1 FF for the van der Waals part of the free energy functional; (ii) τ = 1 Å for the surface term in the free energy functional (though one could use a larger τ to get slightly more accurate agreement of the surface energies); and (iii) applying a slightly larger shift of the solvation boundary (δ = 2.5 Å) prior to the PB electrostatic calculations.

An Application Example. One of the most interesting aspects of the VISM method is its ability to capture different solvation states at regions where an accurate solvation description is critical, as in binding pockets or interfaces. As an application example, we decided to test whether ^{M}VISM can capture dry−wet transitions during a molecular encounter of the barnase and barstar proteins. Barstar and barnase form a tight complex which has achieved its extremely fast kinetics of binding by means of optimized electrostatic interactions.[49] As such, both barnase and barstar binding interfaces are rich in charged residues and polar residues, as shown in Table 1.

To create a reasonable encounter pathway for this complex, we separated the two proteins along the axis formed by their geometrical centers. Configurations were saved every 2 Å, with separation distances ranging from 0 (native bound complex) to 15 Å (unbound proteins). For each configuration, we performed VISM calculations starting from both tight and loose initials. Besides the solvation energies, we kept track of the resulting surfaces, from which we could see whether the calculations produced "wet" or "dry" encounter complexes.

Figure 9A displays the energetic profiles obtained for the association between barnase and barstar, starting from tight or loose initials. We found that ^{A}VISM calculations converge at separation distances <3 Å and >10 Å regardless of the initial surfaces. At less than 3 Å, ^{A}VISM calculations necessarily

DOI: 10.1021/acs.jpcb.7b04113
J. Phys. Chem. B 2017, 121, 6538−6548

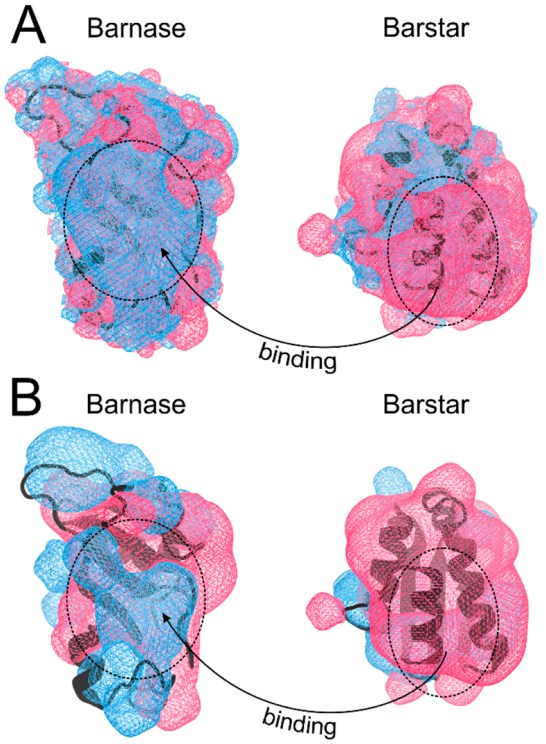


Figure 7. Electrostatic potential at the binding interface of the barnase and barstar proteins, calculated with ^{A}VISM (A) or ^{M}VISM (B). Positive and negative regions are displayed in blue and red, respectively, with the interacting interfaces indicated by dotted circles.

produce "dry" states because water cannot fit in the space between the interaction interfaces, being squeezed out. At separation distances larger than 10 Å, ^{A}VISM always produces "wet" states because the hydrophobic cost of producing large concave surfaces in the space between the proteins becomes too high. In ^{M}VISM calculations, convergence between tight and loose calculations is achieved at slightly larger separation distances. This is likely due to the larger size of water beads—which are squeezed out at separations smaller than 5 Å—and due to less favorable electrostatic interactions with water—which allow for dewetting to occur at separations as large as 11.4 Å.

At intermediate separation distances, however, both ^{A}VISM and ^{M}VISM capture distinctive "dry" or "wet" solvation states depending on the initial surface shape. This is illustrated by the solvation surfaces displayed in Figure 9B (atomistic) and C (coarse-grained), and also evident from the hysteresis in the energy profiles in Figure 9A. Interestingly, partial desolvation has been suggested to occur during the binding encounter of barstar and barnase, at separation distances up to 7 Å.[50] It has been argued that, without the solvent to screen the electrostatic interactions between the two proteins, strong long-range electrostatic interactions pull the two domains together. This effect would contribute favorably to the fast association of the proteins, even if destabilizing the final thermodynamics of binding.[51] Still, both ^{A}VISM and ^{M}VISM calculations reveal that the "wet" encounter pathway is significantly favored by approximately 150−180 k_BT with respect to the "dry" pathway, which makes sense as this is a highly hydrophilic complex. Not surprisingly, the difference in free energy is dominated by the electrostatic component, since the interacting surfaces are highly charged and thus display very favorable electrostatic interactions with water, favoring the "wet" states.

Overall, ^{M}VISM seems to correctly predict the nature of dry−wet transitions in the encounter between barnase and barstar. Combining the VISM model with coarse-grained molecular dynamics simulations is one of our ultimate goals to study protein−protein interactions. This application example thus serves to highlight the potential of using ^{M}VISM to account for solvation effects in future coarse-grained simulations with implicit solvent.

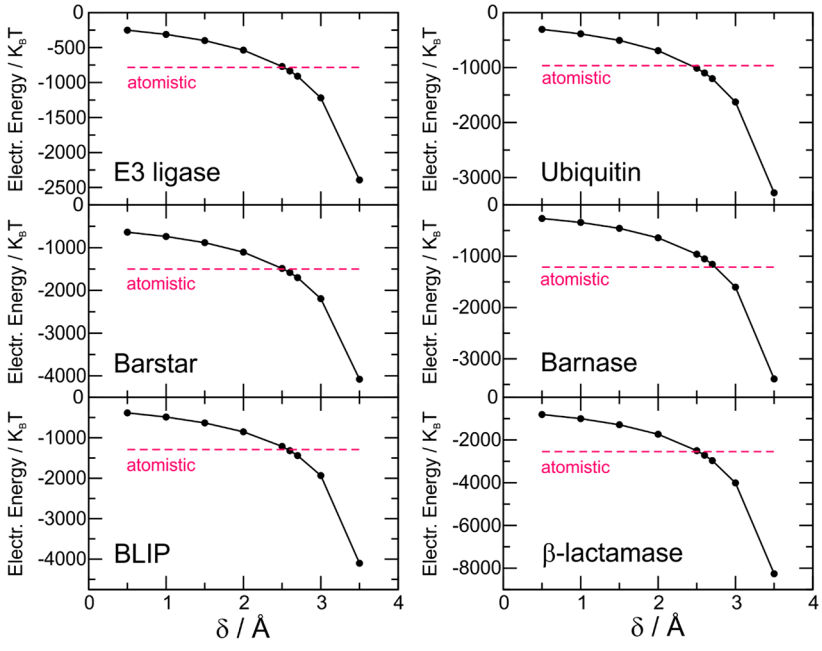


Figure 8. Electrostatic energies calculated with ^{M}VISM using different shifting (δ) values. As a reference, atomistic electrostatic energies obtained with δ = 1 Å are indicated by pink dotted lines.

DOI: 10.1021/acs.jpcb.7b04113
J. Phys. Chem. B 2017, 121, 6538−6548

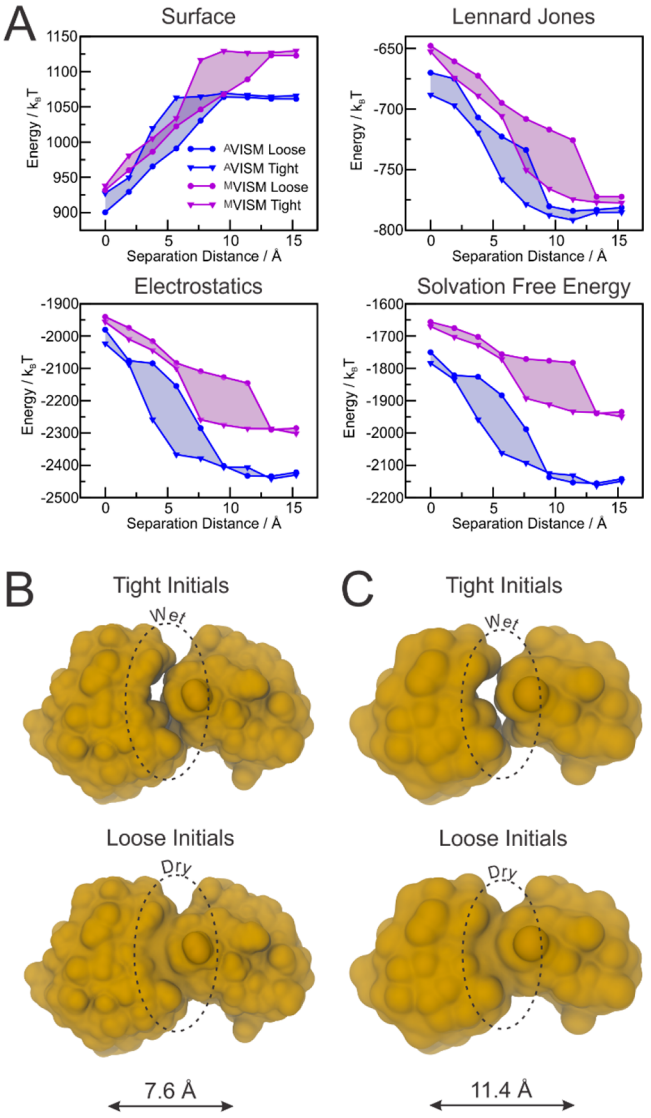


Figure 9. Dry–wet transitions in the encounter complex of barstar and barnase. (A) ^{A}VISM (blue) and ^{M}VISM (purple) solvation energies obtained from tight (triangles) or loose (circles) initial surfaces. (B) Different solvation states captured by ^{A}VISM at 7.6 Å. (C) Different solvation states captured by ^{M}VISM at 11.4 Å.

CONCLUSIONS

In this work, we adapted the VISM method to work with the MARTINI 2.1 coarse-grained force field. Using atomistic VISM energies as a reference, the main adjustment of the MARTINI Hamiltonian to produce VISM solvation energies consisted of tuning the LJ parameter ε to correct for overestimated van der Waals interactions. MARTINI's overestimation of LJ interactions arises in part as an enthalpic compensation to the loss of hydrophobic interactions due to reduced degrees of freedom of coarse-grained water. In VISM, such compensation is not required because hydrophobic interactions are accounted for implicitly in the surface term of the free energy functional.

After downscaling of LJ interactions, ^{M}VISM displayed good agreement with atomistic (^{A}VISM) calculations, not only capturing the trends in solvation free energy of six different proteins but also correctly partitioning the free energy into hydrophobic, van der Waals, and electrostatic components. Despite the use of simplified coarse-grained charges, ^{M}VISM is capable of producing reasonable electrostatic complementarity between binding interfaces, which is a necessary feature for subsequent protein–protein binding studies. In terms of electrostatics, ^{M}VISM could be further improved with the use of more sophisticated versions of MARTINI, such as the polarizable MARTINI 2.2P.[48]

In the present formulation, ^{M}VISM can qualitatively capture and reproduce the dry–wet transitions in the binding of barnase and barstar, as observed in atomistic VISM calculations and supported by previous theoretical studies.[50] As such, the results reported herein are the first step toward the development of a hybrid approach combining the coarse-grained simulations of solute atoms and our VISM description of completely implicit solvent. We expect this work to broaden VISM applicability to study complex solvation properties of important protein assemblies in the near future.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.7b04113.

Additional tables showing downscaling of the ε values and ^{A}VISM and ^{M}VISM energies and figures showing ^{A}VISM and ^{M}VISM calculations, ubiquitin and E3 protein solvation surfaces, barnase and barstar solvation surfaces, β-lactamase and BLIP solvation surfaces, electrostatic potential of the ubiquitin and E3 ligase, and electrostatic potential of the β-lactamase and BLIP (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: cgravinaricci@ucsd.edu. Phone: +1 858 405 0579.

ORCID

Clarisse G. Ricci: 0000-0002-3289-2248

Joachim Dzubiella: 0000-0001-6751-1487

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the U.S. National Science Foundation (NSF), the National Institutes of Health (NIH), the Howard Hughes Medical Institute (HHMI), the National Biomedical Computation Resources (NBCR), and the San Diego Supercomputer Center (SDSC). L.-T.C. and B.L. were partially supported by the NSF grants DMS-1319731 and DMS-1620487. The authors acknowledge Dr. Paulo C. T. Souza for valuable discussions on the MARTINI FF.

ABBREVIATIONS

BLIP, β-lactamase inhibitor protein; CG, coarse-grained; FF, force field; LJ, Lennard-Jones; MD, molecular dynamics; PB, Poisson–Boltzmann; UBQ, ubiquitin; VISM, variational implicit solvent method

REFERENCES

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

(2) Ansari, A.; Jones, C. M.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. The Role of Solvent Viscosity in the Dynamics of Protein Conformational Changes. *Science* **1992**, *256*, 1796–1798.

DOI: 10.1021/acs.jpcb.7b04113
J. Phys. Chem. B 2017, 121, 6538–6548

(3) Fenimore, P. W.; Frauenfelder, H.; McMahon, B. H.; Parak, F. G. Slaving: Solvent Fluctuations Dominate Protein Dynamics and Functions. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 16047−16051.
(4) Liu, P.; Huang, X.; Zhou, R.; Berne, B. J. Observation of a Dewetting Transition in the Collapse of the Mellitin Tetramer. *Nature* **2005**, *437*, 159−162.
(5) Baron, R.; McCammon, J. A. Molecular Recognition and Ligand Association. *Annu. Rev. Phys. Chem.* **2013**, *64*, 151−175.
(6) Cramer, C. J.; Truhlar, D. G. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem. Rev.* **1999**, *99*, 2161−2200.
(7) Davis, M. E.; McCammon, J. A. Electrostatics in Biomolecular Structure and Dynamics. *Chem. Rev.* **1990**, *90*, 509−521.
(8) Sharp, K. A.; Honig, B. Electrostatic Interactions in Macromolecules. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301−332.
(9) Baker, N. A. Improving Implicit Solvent Simulations: A Poisson-Centric View. *Curr. Opin. Struct. Biol.* **2005**, *15*, 137−143.
(10) 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.
(11) Bashford, D.; Case, D. A. Generalized Born Models of Macromolecular Solvation Effects. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129−152.
(12) Feig, M.; Brooks, C. L. Recent Advances in the Development and Application of Implicit Solvent Models in Biomolecule Simulations. *Curr. Opin. Struct. Biol.* **2004**, *14*, 217−224.
(13) Eisenberg, D.; McLachlan, A. D. Solvation Energy in Protein Folding and Binding. *Nature* **1986**, *319*, 199−203.
(14) Chandler, D. Interfaces and the Driving Force of Hydrophobic Assembly. *Nature* **2005**, *437*, 640−647.
(15) Huang, Q.; Ding, S.; Hua, C.-Y.; Yang, H.-C. A Computer Simulation Study of Water Drying at the Interface of Protein Chains. *J. Chem. Phys.* **2004**, *121*, 1969−1977.
(16) Zhou, R.; Huang, X.; Margulis, C. J.; Berne, B. J. Hydrophobic Collapse in Multidomain Protein Folding. *Science* **2004**, *305*, 1605−1609.
(17) Dong, F.; Vijaykumar, M.; Zhou, H. X. Comparison of Calculation and Experiment Implicates Significant Electrostatic Contributions to the Binding Stability of Barnase and Barstar. *Biophys. J.* **2003**, *85*, 49−60.
(18) Dzubiella, J.; Swanson, J. M. J.; McCammon, J. A. Coupling Hydrophobicity, Dispersion, and Electrostatics in Continuum Solvent Models. *Phys. Rev. Lett.* **2006**, *96*, 087802.
(19) Dzubiella, J.; Swanson, J. M. J.; McCammon, J. A. Coupling Nonpolar and Polar Solvation Free Energies in Implicit Solvent Models. *J. Chem. Phys.* **2006**, *124*, 084905.
(20) Zhou, S.; Cheng, L.-T.; Dzubiella, J.; Li, B.; McCammon, J. A. Variational Implicit Solvation with Poisson-Boltzmann Theory. *J. Chem. Theory Comput.* **2014**, *10*, 1454−1467.
(21) Zhou, S.; Cheng, L.-T.; Sun, H.; Che, J.; Dzubiella, J.; Li, B.; McCammon, J. A. LS-VISM: A Software Package for Analysis of Biomolecular Solvation. *J. Comput. Chem.* **2015**, *36*, 1047−1059.
(22) Cheng, L.-T.; Dzubiella, J.; McCammon, J. A.; Li, B. Application of the Level-Set Method to the Implicit Solvation of Nonpolar Molecules. *J. Chem. Phys.* **2007**, *127*, 084503.
(23) Cheng, L. T.; Xie, Y.; Dzubiella, J.; McCammon, J. A.; Che, J.; Li, B. Coupling the Level-Set Method with Molecular Mechanics for Variational Implicit Solvation of Nonpolar Molecules. *J. Chem. Theory Comput.* **2009**, *5*, 257−266.
(24) Wang, Z.; Che, J.; Cheng, L.-T.; Dzubiella, J.; Li, B.; McCammon, J. A. Level-Set Variational Implicit-Solvent Modeling of Biomolecules with the Coulomb-Field Approximation. *J. Chem. Theory Comput.* **2012**, *8*, 386−397.
(25) Guo, Z.; Li, B.; Dzubiella, J.; Cheng, L.-T.; McCammon, J. A.; Che, J. Evaluation of Hydration Free Energy by Level-Set Variational Implicit Solvent Model with Coulomb-Field Approximation. *J. Chem. Theory Comput.* **2013**, *9*, 1778−1787.
(26) Cheng, L.-T.; Wang, Z.; Setny, P.; Dzubiella, J.; Li, B.; McCammon, J. A. Interfaces and Hydrophobic Interactions in Receptor-Ligand Systems: A Level-Set Variational Implicit Solvent Approach. *J. Chem. Phys.* **2009**, *131*, 144102.
(27) Setny, P.; Wang, Z.; Cheng, L.-T.; Li, B.; McCammon, J. A.; Dzubiella, J. Dewetting-Controlled Binding of Ligands to Hydrophobic Pockets. *Phys. Rev. Lett.* **2009**, *103*, 187801.
(28) Zhou, S.; Rogers, K. E.; de Oliveira, C. A. F.; Baron, R.; Cheng, L.-T.; Dzubiella, J.; Li, B.; McCammon, J. A. Variational Implicit-Solvent Modeling of Host-Guets Binding: A Case Study on Cucurbit[7]uril. *J. Chem. Theory Comput.* **2013**, *9*, 4195−4204.
(29) Guo, Z.; Li, B.; Dzubiella, J.; Cheng, L.-T.; McCammon, J. A.; Che, J. Heterogeneous Hydration of p53-MDM2 Complex. *J. Chem. Theory Comput.* **2014**, *10*, 1302−1313.
(30) Guo, Z.; Li, B.; Cheng, L.-T.; Zhou, S.; McCammon, J. A.; Che, J. Identification of Protein-Ligand Binding Sites by the Level-Set Variational Implicit-Solvent Approach. *J. Chem. Theory Comput.* **2015**, *11*, 753−765.
(31) Zhou, S.; Sun, H.; Cheng, L.-T.; Dzubiella, J.; Li, B.; McCammon, J. A. Stochastic Level-Set Variational Implicit-Solvent Approach to Solute-Solvent Interfacial Fluctuations. *J. Chem. Phys.* **2016**, *145*, 054114.
(32) Marrink, S. J.; de Vries, A. H.; Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **2004**, *108*, 750−760.
(33) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.
(34) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819−834.
(35) Lopez, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hunenberger, P. H.; Marrink, S. J. Martini Coarse-Grained Force Field: Extension to Carbohydrates. *J. Chem. Theory Comput.* **2009**, *5*, 3195−3210.
(36) Lopez, C. A.; Sonova, Z.; van Eerden, F. J.; de Vries, A. H.; Marrink, S. J. Martini Force Field Parameters for Glycolipids. *J. Chem. Theory Comput.* **2013**, *9*, 1694−1708.
(37) De Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schafer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J. Chem. Theory Comput.* **2013**, *9*, 687−697.
(38) Uusitalo, J. J.; Ingolfsson, H. I.; Akhshi, P.; Tieleman, D. P.; Marrink, S. J. Martini Coarse-Grained Force Field: Extension to DNA. *J. Chem. Theory Comput.* **2015**, *11*, 3932−3945.
(39) Marrink, S. J.; Tieleman, D. P. Perspective on the Martini Model. *Chem. Soc. Rev.* **2013**, *42*, 6801−6822.
(40) Stark, A. C.; Andrews, C. T.; Elcock, A. H. 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. Theory Comput.* **2013**, *9*, 4176−4185.
(41) Negami, T.; Shimizu, K.; Terada, T. Coarse-grained Molecular Dynamics of Protein-ligand Binding. *J. Comput. Chem.* **2014**, *35*, 1835−1845.
(42) Arnarez, C.; Uusitalo, J. J.; Masman, M. F.; Ingólfsson, H. I.; de Jong, D. H.; Melo, M. N.; Periole, X.; de Vries, A. H.; Marrink, S. J. Dry Martini, a Coarse-Grained Force Field for Lipid Membrane Simulations with Implicit Solvent. *J. Chem. Theory Comput.* **2015**, *11*, 260−175.
(43) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(44) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D., Jr. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain X_1 and X_2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257−3273.
(45) Peschard, P.; Kozlov, G.; Lin, T.; Mirza, I. A.; Berghuis, A. M.; Lipkowitz, S.; Park, M.; Gehring, K. Structural Basis for Ubiquitin-

DOI: 10.1021/acs.jpcb.7b04113
J. Phys. Chem. B 2017, 121, 6538−6548

Mediated Dimerization and Activation of the Ubiquitin Protein Ligase Cbl-b. *Mol. Cell* **2007**, *27*, 474−485.
(46) Buckle, A. M.; Schreiber, G.; Fersht, A. R. Protein-Protein Recognition: Crystal Structural Analysis of a Barnase-Barstar Complex at 2.0 Å Resolution. *Biochemistry* **1994**, *33*, 8878−8889.
(47) Lim, D.; Park, H. U.; De Castro, L.; Kang, S. G.; Lee, H. S.; Jensen, S.; Lee, K. J.; Strynadka, N. C. Crystal Structure and Kinetic Analysis of Beta-Lactamase Inhibitor Protein-II in Complex with TEM-1 Beta-Lactamase. *Nat. Struct. Biol.* **2001**, *8*, 848−852.
(48) de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Parameters for the Coarse-Grained Protein Force Field. *J. Chem. Theory Comput.* **2013**, *9*, 687−697.
(49) Lee, L.-P.; Tidor, B. Barstar is Electrostatically Optimized for Tight Binding to Barnase. *Nat. Struct. Biol.* **2001**, *8*, 73−76.
(50) Gabdoulline, R. R.; Wade, R. C. On the Protein-Protein Diffusional Encounter Complex. *J. Mol. Recognit.* **1999**, *12*, 226−234.
(51) Elcock, A. H.; Gabdoulline, R. R.; Wade, R. C.; McCammon, J. A. Computer Simulation of Protein-Protein Association Kinetics: Acetylcholinesterase-Fasciculin. *J. Mol. Biol.* **1999**, *291*, 149−162.

DOI: 10.1021/acs.jpcb.7b04113
J. Phys. Chem. B 2017, 121, 6538−6548