Title

Coupling Monte Carlo, Variational Implicit Solvation, and Binary Level-Set for Simulations of Biomolecular Binding

Journal

Journal of Chemical Theory and Computation, 17(4)

ISSN

1549-9618

Authors

Zhang, Zirui
Ricci, Clarisse G
Fan, Chao
et al.

Publication Date

2021-04-13

DOI

10.1021/acs.jctc.0c01109

Peer reviewed

Published in final edited form as:

J Chem Theory Comput. 2021 April 13; 17(4): 2465–2478. doi:10.1021/acs.jctc.0c01109.

Coupling Monte Carlo, Variational Implicit Solvation, and Binary Level-Set for Simulations of Biomolecular Binding

Zirui Zhang[1,*], **Clarisse G. Ricci**[2,*], **Chao Fan**[1], **Li-Tien Cheng**[1,†], **Bo Li**[3,‡], **J. Andrew McCammon**[2]

[1]Department of Mathematics, University of California, San Diego, La Jolla, CA 92093-0112, USA

[2]Department of Chemistry and Biochemistry, Department of Pharmacology, University of California, San Diego, La Jolla, CA 92093-0365, USA

[3]Department of Mathematics and Quantitative Biology Ph.D. Program, University of California, San Diego, La Jolla, CA 92093-0112, USA

Abstract

We develop a hybrid approach that combines the Monte Carlo (MC) method, a variational implicit-solvent model (VISM), and a binary level-set method for the simulation of biomolecular binding in an aqueous solvent. The solvation free energy for the biomolecular complex is estimated by minimizing the VISM free-energy functional of all possible solute-solvent interfaces that are used as dielectric boundaries. This functional consists of the solute volumetric, solute-solvent interfacial, solute-solvent van der Waals interaction, and electrostatic free energy. A technique of shifting the dielectric boundary is used to accurately predict the electrostatic part of the solvation free energy. Minimizing such a functional in each MC move is made possible by our new and fast binary level-set method. This method is based on the approximation of surface area by the convolution of an indicator function with a compactly supported kernel, and is implemented by simple flips of numerical grid cells locally around the solute-solvent interface. We apply our approach to the p53-MDM2 system for which the two molecules are approximated by rigid bodies. Our efficient approach captures some of the poses before the final bound state. All-atom molecular dynamics simulations with most of such poses quickly reach to the final bound state. Our work is a new step toward realistic simulations of biomolecular interactions. With further improvement of coarse graining and MC sampling, and combined with other models, our hybrid approach can be used to study the free-energy landscape and kinetic pathways of ligand binding to proteins.

Keywords

Molecular binding; dewetting/desolvation; variational implicit-solvent model; binary level-set method; Monte Carlo simulations; p53-MDM2

[†]To whom correspondence should be addressed. lcheng@math.ucsd.edu. [‡]To whom correspondence should be addressed. bli@math.ucsd.edu.
[*]Contributed equally.

I. INTRODUCTION

Biomolecular binding in aqueous solvent is fundamental to biological functions yet extremely complex due to the many-body interactions spanning across multiple temporal and spatial scales. Recent years have seen a growing interest in understanding the mechanisms of such biomolecular processes, due to particularly the rapid development in rational drug design.[1–6] However, there are several bottleneck issues in the current computational study of biomolecular binding and unbinding. One of them is the efficient description of the effect of water in the hydration of biomolecules. Water is recognized as an important player in many biomolecular activities, including protein conformational changes, protein-ligand binding and unbinding, and protein-protein interactions.[7–12] This role often results from the collective behaviors of the network of many water molecules. It can therefore be costly to describe the effect of water by including many individual water molecules in computer simulations. Another one is the general issue of crossing free-energy barriers in the binding process in which solute and solvent fluctuations are critical. In general, the process of biomolecular binding can be dominated by either “conformational change”, or “induced fit”, or the mix of these.[13–19] Regardless, before reaching a “binding ready” pose, a biomolecular complex is in a diffusional mode, often taking much time of the entire binding process.[20]

In this work, we develop a hybrid computational approach to the simulation of biomolecular binding and unbinding processes. This approach combines the Monte Carlo (MC) method, a variational implicit-solvent model (VISM),[21,22] and a new and fast binary level-set method. The MC method is used here to simulate the diffusion of individual proteins and formation of the biomolecular complex, while the VISM with implementation by the binary level-set method is for the efficient estimation of the solvation free energy with an implicit solvent through the solute-solvent interfaces. If we consider two biomolecules immersed in an aqueous solvent, then the free energy for these two molecules to bind or unbind consists of contributions arising from the solute-solute interactions and the solvation of these molecules. Our MC-VISM simulation consists of a sequence of MC moves each of which can be accepted or rejected by the Metropolis criterion with respect to the total interaction free energy.

Central in VISM is an effective VISM solvation free-energy functional of all possible solute-solvent interfaces that are used as dielectric boundaries. This functional consists of the solute volumetric, solute-solvent interfacial, solute-solvent van der Waals (vdW) interaction, and electrostatic free energy. For a fixed conformation of the biomolecules, which means particularly that all the solute atoms and their partial charges are fixed, we obtain an estimate of the solvation free energy by numerically minimizing the VISM free-energy functional. In recent years, we have developed the level-set numerical method for such minimization in three-dimensional space with complex protein geometries. Our series of works have demonstrated that the level-set VISM can capture well the solvation free energies, particularly for nonpolar systems, the effect of electrostatics such as the net electrostatic force acting on biomolecules and the subtle step-by-step hydration for charged systems, and the dry and wet hydration states.[23–33] Recently, the efficiency of the VISM is further improved by the coarse-graining through the Martini force-field.[34] Moreover,

combined with the string method and Brownian dynamics with a multi-state potential, the level-set VISM is also applied to predicting the pathways of dry-wet transitions as well as the kinetics of the molecular binding and unbinding for a model system.[35]

Electrostatics is one of the dominant components of the solvation of charged molecules in an aqueous solvent. In VISM, the electrostatic part of the solvation free energy can be incorporated through the dielectric-boundary Poisson–Boltzmann (PB) theory.[36–41] Here, however, we use the Coulomb-field approximation (CFA), as it requires no solution to partial differential equations, and hence is rather efficient.[27,28,30,42] We also use a technique of shifting the dielectric boundary to predict more accurately the electrostatic part of the solvation free energy.

The key to our MC-VISM simulations is our new and fast, binary level-set method for minimizing the VISM free energy done in each of the Metropolis MC moves. This method is based on the approximation of surface area by the convolution of an indicator function with a compactly supported kernel. Instead of a continuous level-set function, here a binary level-set function defined on numerical grid cells is used to define the solute-solvent interface, or equivalently the solute and solvent regions. With such a binary level-set function, the VISM free-energy functional can be approximated by summing over the contributions from all the grid cells. We can find a minimum conformation (i.e., the optimal dielectric boundary) of the free-energy functional by iteratively flipping the signature of the grid cell in a steepest-descent fashion. This formulation does not require solving a partial differential equation. Moreover, the flipping is only done locally around the boundary. Therefore, the method is very fast compared with the classical, continuous level-set method. It is fast enough to be coupled with the MC method for biomolecular simulations. We shall test the convergence, accuracy, and speed of our numerical algorithm by considering single ions in solvent for which analytical and experimental results are available.

We apply our approach to the p53-MDM2 system. The p53-MDM2 interaction is a relevant pharmacological target for anti-cancer therapeutics[43–47] and an important model for the study of protein-protein binding due to the abundance of structural information.[48–53] MDM2 has a highly concave and hydrophobic binding pocket that undergoes dewetting fluctuations prior to the binding of p53, as seen in explicit solvent molecular dynamics (MD) simulations and level-set VISM calculations.[28,30,54] Here, we first calculate the solvation free energy of this protein complex and obtain the potential of mean force with respect to some separation distance of the two molecules. We then approximate each of them as a rigid body, and carry out the MC-VISM simulations of their binding process. We show our efficient approach can capture some of the poses before the final bound state. All-atom MD simulations starting with such poses quickly reach the final bound state.

The rest of the paper is organized as follows: In section II, we describe our MC-VISM theory. In particular, we review the VISM free-energy functional with the CFA of the electrostatic solvation free energy. In section III, we describe the binary level-set method for minimizing the VISM free-energy functional, and report the results of testing the binary level-set VISM for the solvation of ions. In section IV, we apply our binary level-set VISM

to the solvation of p53-MDM2 complex, and show the result of rigid-body MC-VISM simulations of the binding of p53-MDM2. Finally, in section V, we draw our conclusions.

II. THEORY

A. The total interaction free energy

We consider two molecules A (with M atoms) and B (with N atoms) in an aqueous solvent. We denote by $\mathbf{r}_i^A$, Q_i^A ($i = 1, \ldots, M$) and $\mathbf{r}_i^B$, Q_i^B ($j = 1, \ldots, N$) all the solute atoms and their partial charges of A and B, respectively. The total interaction free energy of this molecular complex in the solvent is

$$G_{\text{total}} = G_{\text{vdW,sol–sol}} + G_{\text{elec,sol–sol}} + G_{\text{solvation}} \tag{1}$$

The first two terms are the solute-solute van der Waals (vdW) and electrostatic interaction energies, given by the Lennard-Jones (LJ) and Coulombic interaction potentials, respectively,

$$G_{\text{vdW,sol–sol}} = \sum_{i=1}^{M}\sum_{j=1}^{N} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{\left|\mathbf{r}_i^A - \mathbf{r}_j^B\right|}\right)^{12} - \left(\frac{\sigma_{ij}}{\left|\mathbf{r}_i^A - \mathbf{r}_j^B\right|}\right)^{6}\right],$$

$$G_{\text{elec,sol–sol}} = \frac{1}{4\pi\varepsilon_0\varepsilon_{\text{w}}}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{Q_i^A Q_j^B}{\left|\mathbf{r}_i^A - \mathbf{r}_j^B\right|}.$$

Here, ε_{ij} and σ_{ij} are the energy and length parameters of the LJ potential for the interaction between $\mathbf{r}_i^A$ and $\mathbf{r}_j^B$, ε_0 is the vacuum permittivity, and ε_{w} is the relative permittivity of the solvent. The solvation free energy $G_{\text{solvation}}$ is given by

$$G_{\text{solvation}} = \min_{\Gamma} G_{\text{VISM}}[\Gamma].$$

Here, $G_{\text{VISM}}[\Gamma]$ is the VISM solvation free energy of a solute-solvent interface Γ that encloses all the solute atoms, and the minimum is taken over all such interfaces Γ. This functional $G_{\text{VISM}}[\Gamma]$ is described in more details for a general set up in the next subsection.

B. The VISM free-energy functional

Consider one or more molecules of solute atoms located at $\mathbf{r}_i$ and carrying partial charges Q_i ($i = 1, \ldots, L$) (In the case of two molecules as described above, the total number of solute atoms is $L = M + N$.) For any closed surface Γ that encloses all the solute atoms $\mathbf{r}_i$, we denote by Ω_{m} and Ω_{w} the interior and exterior of the surface Γ, and call them the solute and solvent regions, respectively; cf. Fig. 1 (Left). We also call Γ a solute-solvent interface. We shall denote by ε_{m} the dielectric permittivity of the solute region Ω_{m}.

The VISM solvation free-energy functional of all possible solute-solvent interfaces Γ is defined by[21,22,31,33]

$$G_{\mathrm{VISM}}[\Gamma] = \Delta P\ \mathrm{vol}(\Omega_{\mathrm{m}}) + \int_{\Gamma} \gamma dS + \rho_{\mathrm{w}} \sum_{i=1}^{L} \int_{\Omega_{\mathrm{w}}} U_i(|\mathbf{r}-\mathbf{r}_i|) dV_{\mathbf{r}} + G_{\mathrm{elec}}[\Gamma]\,. \tag{2}$$

Here, the first term $\Delta P\,\mathrm{vol}\,(\Omega_{\mathrm{m}})$ describes the energy of creating the solute region Ω_{m} in the solvent, where ΔP is the difference of the pressure of the solvent liquid and solute vapor, respectively. In this work, we shall neglect this term as it is rather small compared with other terms.

The second term is the surface energy, where γ is the solute-solvent interfacial surface tension. In general, we can take the form $\gamma = \gamma_0(1 - 2\tau H)$, where γ_0 is the constant surface tension for a planar liquid-vapor interface, τ is the curvature correction coefficient known as the Tolman length,[55,56] and H is the local mean curvature (the average of the two principal curvatures) that is positive for a spherical solute. While the Tolman correction is often found to be useful in many cases,[23,26,27,31,33] it can also be more complicated and costly in computations. Therefore, in this work, we shall neglect this correction.

The third term is the solute-solvent vdW type interaction energy. The constant ρ_{w} is the solvent number density. For each i, the term $U_i(|\mathbf{r} - \mathbf{r}_i)$ is the vdW type interaction potential between the solute atom at $\mathbf{r}_i$ and a solvent molecule or ion at $\mathbf{r}$. We employ the LJ potential

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

where the parameters ε_i of energy and σ_i of length can vary with solute atoms as in a conventional force-field.

The last term $G_{\mathrm{elec}}[\Gamma]$ in (2) is the electrostatic part of the solvation free energy. Here we shall use the Coulomb-field approximation (CFA):[27,28,30,42]

$$G_{\mathrm{elec}}[\Gamma] = \frac{1}{32\pi^2\varepsilon_0}\left(\frac{1}{\varepsilon_{\mathrm{w}}} - \frac{1}{\varepsilon_{\mathrm{m}}}\right)\int_{\Omega_{\mathrm{w}}} \left|\sum_{i=1}^{L} \frac{Q_i(\mathbf{r}-\mathbf{r}_i)}{|\mathbf{r}-\mathbf{r}_i|^3}\right|^2 dV_{\mathbf{r}}\,. \tag{4}$$

For a given set of solute atomic positions $\mathbf{r}_i$ and partial charges Q_i ($i = 1, \ldots, L$), we can now minimize the VISM solvation free-energy functional $G_{\mathrm{VISM}}[\Gamma]$ to obtain an optimal solute-solvent interface. It has been found that such a VISM surface, i.e., a VISM free-energy minimizing surface, often represents the surface with the first peak of water density determined using the position of oxygen atoms in water molecules,[27,28,30,31] and may not be necessary the best choice of dielectric boundary. In fact, if we use a VISM surface as the dielectric boundary to calculate the electrostatic solvation energy, then the error can be sometimes significant.[27,28,30,31] Here we use a previously developed technique to shift the VISM surface by a constant distance ξ (usually $\xi = 1$ Å) toward the solute region; cf. Fig.

2 (Left). The shifted boundary is an effective dielectric boundary, and is used to the final calculation of the electrostatic free energy.

C. The MC-VISM algorithm and parameters

We now consider the binding of two molecules A and B as described in Section II A. The solute atomic positions and the corresponding partial charges are $\mathbf{r}_i^A$, Q_i^A ($i = 1, \ldots, M$) and $\mathbf{r}_j^B$, Q_j^B ($j = 1, \ldots, N$), respectively. The total binding free energy of such a molecular complex is given in Eq. (1). To explore the complex binding process, we approximate the two molecules as rigid bodies. This means that the position of the molecules A and B are determined by their centers of mass $\mathbf{R}^{\mathrm{A}}$ and $\mathbf{R}^{\mathrm{B}}$, and the orientation of the smaller molecule relative to the larger one. We also fix the larger molecule throughout the simulation.

The MC-VISM algorithm for the binding of two molecules.—Step 1. Initialize the system: set up the initial atomic positions of the moelcules, and input all the parameters.

Step 2. Randomly perturb the smaller molecule. Perturbations include both rigid-body rotations and translations.

Step 3. Calculate the solute-solute vdW and Coulombic interaction energies.

Step 4. Calculate the solvation free energy by minimizing the VISM free-energy functional.

Step 5. Calculate the total binding free energy $G_{\mathrm{total}}^{\mathrm{new}}$ and the free energy difference $\Delta G = G_{\mathrm{total}}^{\mathrm{new}} - G_{\mathrm{total}}^{\mathrm{old}}$.

Step 6. If $\Delta G < 0$, then accept the MC move. Otherwise, generate a random number $\alpha \in [0, 1]$ and accept the move if and only if $e^{-\Delta G/k_{\mathrm{B}}T} \leq \alpha$. If the move is rejected, go to Step 2.

We use the LJ parameters from the force-field in CHARMM36.[57,58] In Table I, we list the VISM parameters.

III. NUMERICAL METHOD

A. A binary level-set method

We describe in detail our new, binary level-set method for numerically minimizing the VISM free-energy functional Eq. (2) (with $\Delta P = 0$ and $\tau = 0$). In our numerical computations, we choose the solvation region Ω to be a box: $\Omega = (-A, A)^3$ for some length parameter A (in υ) of the side of the box, and cover it with a uniform finite difference grid with grid size (i.e., the side of each grid cell) h. We define a binary level-set function φ on Ω with a such a grid by $\varphi(\mathbf{x}) = 1$ or -1 on each grid cell.[59] With such a binary level-set function, we obtain the (approximate) solute region Ω_{m} and solvent region Ω_{w} to be the union of all the grid cells with φ-value -1 and that with 1, respectively; cf. Fig. 1 (Right). We shall consider only those binary level-set functions with the corresponding solute region Ω_{m} containing all the solute atoms $\mathbf{r}_1, \ldots, \mathbf{r}_L$. The solute-solvent interface Γ is still defined

to be the interface that separates the solute and solvent regions. But it is now composed of a collection of faces of grid cells.

With a binary level-set formulation, we can discretize the VISM free-energy functional Eq. (2) (with $\Delta P = 0$ and $\tau = \gamma_0$). The vdW interaction term (the third term), and the electrostatic interaction term $G_{\text{elec}}[\Gamma]$ (cf. Eq. (4)) are both integrals over the solvent region Ω_{w}. They can be simply approximated by the center-point integration rule as sums over all the grid cells composing Ω_{w},

$$G_{\text{vdW}}[\Gamma] = \sum_{\mathbf{x}_j \in \Omega_{\text{w}}} (G_{\text{vdW}})_j + \mathcal{O}(h), \tag{5}$$

$$G_{\text{elec}}[\Gamma] = \sum_{\mathbf{x}_j \in \Omega_{\text{w}}} (G_{\text{elec}})_j + \mathcal{O}(h), \tag{6}$$

where $\mathbf{x}_j$ is the center of the jth grid cell in the solvent region Ω_{w} and the notation $\mathcal{O}(h)$ indicates that the discretization error is of the order of h. For each j,

$$(G_{\text{vdW}})_j = \rho_{\text{w}} \sum_{i=1}^{L} U_i(|\mathbf{x}_j - \mathbf{r}_i|) h^3, \tag{7}$$

$$(G_{\text{elec}})_j = \frac{1}{32\pi^2 \varepsilon_0} \left(\frac{1}{\varepsilon_{\text{w}}} - \frac{1}{\varepsilon_{\text{m}}} \right) \left| \sum_{i=1}^{L} \frac{Q_i(\mathbf{x}_j - \mathbf{r}_i)}{|\mathbf{x}_j - \mathbf{r}_i|^3} \right|^2 h^3 . \tag{8}$$

These are the contributions from the jth grid cell to the total vdW energy and electrostatic energy, respectively. We remark that, if a large number of solute atoms are fixed in simulation, then we could compute all the corresponding contributions $(G_{\text{vdw}})_j$ and $(G_{\text{elec}})_j$ before any interation loop, and store these values for use throughout the entire simulation.

With $\tau = 0$ in Eq. (2), the surface energy is $\gamma_0 \text{Area}(\Gamma)$. To approximate the surface area using a binary level-set function φ, we introduce a kernel function $K = K(\mathbf{x})$ ($\mathbf{x} \in \mathbb{R}^3$). We assume that the kernel is positive and radially symmetric, i.e., $K(\mathbf{x})$ is a one-variable function of $|\mathbf{x}|$, and that the kernel vanishes outside the unit ball $B_1(\mathbf{0})$ (the ball centered at the origin with radius 1) of the three-dimensional space $\mathbb{R}^3$; cf. Fig. 3. We approximate the surface area of Γ by

$$\text{Area}(\Gamma) = C(\delta) \int_{\mathbf{x} \in \Omega_{\text{m}}} \int_{\mathbf{y} \in \Omega_{\text{w}}} K\left(\frac{\mathbf{x}-\mathbf{y}}{\delta} \right) d\mathbf{y} d\mathbf{x} + \mathcal{O}(\delta^2) \quad \text{for } 0 < \delta \ll 1, \tag{9}$$

where

$$C(\delta) = \left(\delta^4 \int_0^1 a_0(s) ds \right)^{-1} \quad \text{and} \quad a_0(s) = \int_{B_1(0) \cap \{y_3 > s\}} K(\mathbf{y}) d\mathbf{y}.$$

In the integral region for the definition of $>a_0(s)$, y_3 is the third component of the position vector **y**. Similar formula can be found in literature on diffusion generated motion by mean curvature,[60–65] where the area is approximated by convolution with the Gaussian kernel. The idea here is that the surface area of the interface between two regions is related to the amount of substance that diffuses from one region to the other. Optimizing the area formula with respect to δ leads to the choice of $\delta \sim \sqrt{h}$. The discretization of the double-integral in Eq. (9) by the center-point numerical integration rule then leads to

$$\gamma_0 \mathrm{Area}(\Gamma) = \sum_{\mathrm{x}_j \in \Omega_\mathrm{m}} (G_\mathrm{surf})_j + \mathcal{O}(h), \tag{10}$$

$$(G_\mathrm{surf})_j = \gamma_0 C(\delta) h^6 \sum_{\substack{\mathbf{x}_k \in \Omega_\mathrm{w} \\ |\mathbf{x}_k - \mathbf{x}_j| \le \delta}} K(\mathbf{x}_j - \mathbf{x}_k). \tag{11}$$

In our implementation of the binary level-set method, we use the kernel function

$$K(\mathbf{x}) = \begin{cases} \sin^2(\pi |\mathbf{x}|) & \text{if } |\mathbf{x}| \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

We also choose $\delta = 3\sqrt{h}$. From numerical experimentation, we find this choice of kernel function and rescaled kernel radius δ produce robust results.

It now follows from Eqs. (5), (6), and (10) that the final, discretized VISM free energy is given by

$$G_\mathrm{VISM}^\mathrm{disc} = \sum_{\mathbf{x}_j \in \Omega_\mathrm{m}} (G_\mathrm{surf})_j + \sum_{\mathbf{x}_j \in \Omega_\mathrm{w}} \left[(G_\mathrm{vdW})_j + (G_\mathrm{elec})_j \right].$$

Note that we can use this formula to calculate the free-energy change if we flip the signature (i.e., the sign) of the binary level-set function at the center $\mathbf{x}_j$ of an arbitrary grid cell. Suppose we change a grid cell centered at $\mathbf{x}_j$ from Ω_m to Ω_w, which corresponds to filling the void by water, then the change in $G_\mathrm{VISM}^\mathrm{disc}$ due to the flipping is

$$\begin{aligned} \Delta G_j = {} & \gamma_0 C(\delta) h^6 \sum_{\mathbf{x}_k \in \Omega_m} K(\mathbf{x}_j - \mathbf{x}_k) \\ & - \gamma_0 C(\delta) h^6 \sum_{\substack{\mathbf{x}_k \in \Omega_w \\ |\mathbf{x}_k - \mathbf{x}_j| < \delta}} K(\mathbf{x}_j - \mathbf{x}_k) + (G_\mathrm{vdW})_j + (G_\mathrm{elec})_j . \end{aligned} \tag{12}$$

Notice that the first two terms are the difference between the kernel in water region and the kernel in solute region. Similarly, if we change $\mathbf{x}_j$ from Ω_w to Ω_m, which corresponds to removing water from this cell, the resulting energy change is $-\Delta G_j$.

Our binary level-set method for minimizing the VISM free-energy functional is an optimization method of the steepest descent type. Therefore, due to the non-convexity of the VISM free-energy functional, different initial surfaces may relax to different local minimizers of the free-energy functional that are metastable equilibria. Such local minimizers correspond to polymodal hydration states. In order to capture different local minimizers, we usually use two types of initial surfaces: a tight wrap that is a union of the surfaces of vdW spheres centered at solute atoms with reduced radii, and a loose wrap that is a large surface loosely enclosing all the solute atoms.

After a surface is initialized, we can calculate the difference ΔG_j (cf. Eq. (12)) at every center of grid cell near the interface Γ. The minimization of the total (discrete) free energy $G_{\mathrm{VISM}}^{\mathrm{disc}}$ is done by repeatedly flipping the grid cell with the most negative ΔG_j, thus leading to a steepest descent in total energy. After each flipping, the value of ΔG_j at neighboring grid cells within the rescaled kernel radius δ needs to be updated. The algorithm stops when $\Delta G_j > 0$ for every grid cell, which means there's no single flipping that could decrease the total energy, and we reach a local minimum. We note that there may be simultaneous flipings at multiple grid cells that can lead to a global minimum. Since we need to repeatedly look up the grid cell with the minimum value of ΔG_j, we can use the min-heap data structure, which takes logarithmic time to remove the smallest element and insert an element.[66]

B. Algorithm

Algorithm of the binary level-set method.—Step 1. Input all the parameters γ_0, ρ_w, ε_0, ε_m, ε_w, and atomic parameters $\mathbf{r}_i$, ε_i, σ_i, and Q_i for all $i = 1, \ldots, L$. Choose a computational box according to the atomic coordinates and discretize the box uniformly with the prescribed computational grid size h. Initialize the kernel function and the binary level-set function (tight or loose).

Step 2. Compute and store $(G_{\mathrm{vdW}})_j$ (cf. Eq. (7)) and $(G_{\mathrm{elec}})_j$ (cf. Eq. (8)) at centers $\mathbf{x}_j$ of all grid cells.

Step 3. Compute $(G_{\mathrm{geom}})_j$ (cf. Eq. (11)) and ΔG_j (cf. Eq. (12)) at each center $\mathbf{x}_j$ of grid cell. Insert the pair $(\mathbf{x}_j, \Delta G_j)$ to the heap data structure.

Step 4. Find the grid cell with minimum ΔG_j, flip its signature, and update ΔG_k at the neighboring center point $\mathbf{x}_k$ with $|\mathbf{x}_j - \mathbf{x}_k| \leq \delta$.

Step 5. Repeat Step 4 until $\Delta G_j 0$ for all grid cells. At this point, we reach a local minimum where there is no single flipping that can decrease the energy.

C. Test on single ions

To test the VISM and the binary level-set method, we consider an ion, or more generally a spherical molecule with a single charged atom at its center (assumed to be the origin $\mathbf{0}$) carrying a partial charge Q. The VISM free-energy functional Eq. (2) (with $\Delta P = 0$ and $\tau = 0$) is then a function of the radius R of that spherical solute region. It is given by

$$G(R) = 4\pi R^2\gamma + 16\pi\rho_w\varepsilon\left(\frac{\sigma^{12}}{9R^9} - \frac{\sigma^6}{3R^3}\right) + \frac{Q^2}{8\pi\varepsilon_0 R}\left(\frac{1}{\varepsilon_w} - \frac{1}{\varepsilon_m}\right), \quad (13)$$

where $\sigma = \sigma_1$ and $\varepsilon = \varepsilon_1$ are the LJ parameters in the LJ potential for the interaction between the charged molecule and a water molecule; cf. Eq. (2) and Eq. (3). The function $G(R)$ can be minimized very accurately.

We use the parameters in Table I, and set $\sigma = 3.5$ Å and $\varepsilon = 0.3\ k_B T$. We consider different partial charge values $Q = 0$ e, 0.5 e, and 1 e. For each of these values, we minimize the function $G(R)$ (cf. Eq. (13)) to get the minimum value of the solvation free energy. We also use a continuous level-set method and our new, binary level-set method to minimize the VISM free-energy functional Eq. (2) with the current parameters and with the computational box $\Omega = (-8, 8)^3$ Å^3. In Table II, we show the results of our computations. The result of the minimization of the function $G(R)$ (cf. Eq. (13)) is labeled "analytic". The result obtained by the continuous level-set method is labeled "continuous", while that by the binary level-set method is labeled "binary". It is clear that our binary level-set VISM is very accurate compared with the continuous level-set VISM for a single charged particle.

We also apply our binary level-set VISM to the solvation of single ions K^+, Na^+, Cl^-, and F^-. We take the LJ parameters for these ions from the publication.[67] In our calculations, the dielectric boundary of the anion Cl^- or F^- is obtained by shifting the VISM equilibrium surface by $\xi = 1$ Å, which is the length of the water OH bond.[27,28,68–71] In Table III, we display the solvation free energy obtained by our continuous and binary level-set VISM calculations, and the experimental values of solvation free energy[72] for these ions. We see that our VISM result agrees well with experiment.

In Table IV, we show a comparison of the calculation speed among the continuous level-set VISM with CPU, the binary level-set VISM with CPU, and the binary level-set VISM with GPU for the one-particle system with $Q = 0$ as in Table II. All the calculations are performed on a 2017 iMac, with *3.5 GHz Intel Core i5* CPU and *Radeon Pro 575 4096 MB* GPU. The continuous VISM is accurate but slow, and thus it becomes impractical if we need to compute the energy many times in a simulation. The binary level-set VISM is efficient and it also approximates well the continuous level-set VISM.

IV. P53-MDM2: SIMULATION RESULTS AND ANALYSIS

A. Solvation free energy of p53-MDM2

To test whether the binary level-set VISM can capture the dewetting effects in protein interfaces, we used it to study the solvation behavior in the binding cavity of MDM2 in response to the approach of p53 transactivation domain peptide. To investigate the heterogeneous hydration induced by p53 in the MDM2 binding pocket with our approach, we generated an artificial dissociation pathway along the axis formed by the geometrical centers of the two proteins in the bound complex (PDB ID 1YCR). The inter-protein distance along this reaction coordinate varied from $d = 0$ (bound crystallographic complex) to $d = 24$ Å (unbound), with configurations saved every 1 Å. For each configuration,

we calculated binary level-set VISM solvation free energies and solute-solvent interfaces, starting from both loose and tight initial surfaces.

Figure 4 shows the differences of the solvation free energy calculated from tight and loose initial surfaces, and the individual components of the solvation free energy along the dissociation coordinate, d. For small (d < 10 Å) or large (d > 14 Å) inter-domain distances, calculations starting from tight or loose initials converge to the same solvation free energy, indicating they capture similar solvation states ("dry" for short distances and "wet" for large distances). For intermediate distances (10 < d < 14 Å), "branching" of the solvation free energies along the reaction coordinate reveals the existence of heterogeneous solvation states. While tight initial conditions produce fully solvated states, loose initial conditions produce states where water is completely excluded from the inter-domain region, as illustrated in Figure 5 (A) for d = 13 Å. These results show that binary level-set VISM preserves a significant feature of the original continuous level-set VISM that is the ability to capture different stable minima in the solvation landscape.

In the case of MDM2 and p53, the main difference between the dry and wet "branches" consists of the surface component—which favors the "dry branch" by ~ 16 kcal/mol—and of the LJ component—which favors the "wet branch" by ~ 9 kcal/mol. In terms of the total solvation free energy, however, both "dry" and "wet" branches are similarly stable, indicating that the binding cavity of MDM2 is likely to become desolvated at the approach of p53. The location of the "branching" along the reaction coordinate indicates the critical distance at which wet-dry transitions occur during binding or unbinding. In the case of p53 and MDM2, binary level-set VISM calculations indicate that dewetting transitions occur at 10–14 Å, in relatively good agreement with previous continuous level-set VISM calculations, in which branching was detected at 7.5 – 15 Å.[30] Differences between binary and continuous level-set VISM results can be attributed to the "pixelation" of solvation boundaries produced by binary VISM and to the fact that in this work we are not using the Tolman coefficient to adjust the surface tension to the local curvature. Figure 5 (A) displays the solvation boundaries obtained by binary level-set VISM at d = 13 Å starting from loose or tight initial surfaces. Although pixelated and not including the Tolman correction for local curvature, these solvation boundaries overall capture the same solvation states depicted by continuous level-set VISM at the same separation distance; cf. Figure 5.

B. Rigid-body MC-VISM simulations of the binding of p53-MDM2

We then set out to test if the MC-VISM simulations could capture binding events between p53 peptide and MDM2, starting from random unbound configurations. We start with 10 different initial unbound conformations, marked in Table V by s0a and s0b (same conformation), and s1, …, s9. These conformations are generated by pulling p53 by 15 Å away from MDM2 along the axis connecting their geometric centers in the bound complex, and then, except for the first one (marked s0a and s0b in Table V), randomly rotating p53 by less than 90°. Initial positions with steric clashes are rejected. Similar setup is used by Zhang *et al.*[73] to investigate different MC methods. For the first initial conformation, we perform 10 trials; see rows s0a and s0b in Table V. For each of the other initial conformations s1, …, s9, we perform 5 trials. Each trial consists of 100,000 MC moves. The direction

of translation and the axis of rotation is uniformly distributed on the unit sphere. The magnitude of translation is uniformly distributed between 0 and 1 Å. The magnitude of rotation is uniformly distributed between 0°and 3.72°. As a metric for binding, we used the average of pairwise distances as proposed by Zhou *et al.* (2017)[53] (cf. also Figure 8 (A)) and herein called the binding distance. Due to the way pairwise distances are combined, the binding distance reflects not only the proximity between the two proteins, but also the orientation between them—a large binding distance could correspond to an unbound state or to an incorrectly bound state. Table V is a summary of the minimum binding distance and minimum total binding free energy of all the trials.

Figure 6 (A) and Figure 7 show the distribution of the many MC-VISM trajectories along the conformational binding space, with initial poses marked by asterisks, final poses marked by circles, and intermediate poses colored from blue to yellow. Many simulations resulted in large binding distances with some decrease in the binding energy, suggesting that p53 engaged in some kind of non-specihc interactions with MDM2, as consistent with the typical rugged topology expected for energy landscapes of binding. Some simulations, however, produced binding distances < 12 Å that were accompanied by a sharp and favorable decrease in the binding energy, indicating the formation of specific interactions between p53 and MDM2 (highlighted area in Figure 6 (A)). These were considered productive simulations, as they resulted in productive (specific) interactions between p53 and MDM2.

A visual inspection of productive MC simulations reveals that they produced essentially the same binding mode, with the N-terminal portion of the p53 peptide well positioned for binding while the central Y23 and the C-terminal portion are not yet buried within the MDM2 binding cleft (Figure 6 (B)). More specifically, Glu17 (p53) is well positioned to engage in electrostatic interactions with Lys94 (MDM2), Thr18 (p53) interacts with Gln72 (MDM2) and Phe19 (p53) is anchored by hydrophobic interactions with Val93 in the binding cleft of MDM2 (not shown). These poses thus correspond to a pre-bound state whereby the p53 peptide is initially anchored to MDM2 by its N-terminal end.

It is not surprising that MC-VISM cannot sample the fully bound state given that these simulations are not (yet) including conformational flexibility, which is important for the interaction between p53 and MDM2.[48–50] Recent experimental and MD simulation studies suggest that p53 binding to MDM2 follows an “induced fly-casting” mechanism, whereby MDM2 initially binds to a partially disordered p53 that only then folds into its final (and more ordered) binding structure.[53] Interestingly, these studies agree that most of the folding occurs in the C-terminal portion and that initial binding occurs with the N-terminal portion of the p53 peptide—in agreement with the pre-bound state captured by our rigid MC-VISM simulations. The binding mode captured by our MC-VISM is thus similar to the first half of the “coupled binding-folding” mechanism as proposed by Zhou *et al.*.[53]

It seems reasonable to assume that the main obstacle preventing MC-VISM from reaching the final binding pose is the lack of conformational flexibility. To further investigate this aspect, we used the pre-bound states produced by MC-VISM as starting points for explicit solvent MD simulations. In six out of the seven MD simulations, p53 quickly tucked the key W23 and C-terminal tail within the MDM2 binding pocket, reaching fully bound states

in less than 6 ns of simulations (Figure 8 (B) and (C)). The final conformations refined by MD simulations were very similar to the crystallographic complex, as shown by the RMSD calculations and visual inspection (Figure 8 (D) and (E)). In only one simulation, p53 reached an alternative binding mode in which the side-chain of Y19 occupied the central pocket of the MDM2 binding cleft. We thus conclude that the binding mode predicted by rigid MC-VISM consist of a pre-bound state which easily leads to correct binding once the proteins are allowed some degree of conformational flexibility.

V. CONCLUSIONS

We have developed a hybrid approach combining the MC method, the VISM for solvation of biomolecules with an implicit solvent, and a fast binary level-set method for the simulation of biomolecular binding process. We have tested the convergence of our new model and method, and applied our approach to the study of protein complex p53-MDM2. We have demonstrated that our binary level-set VISM can efficiently capture heterogeneous hydration states of protein complex p53-MDM2, and that our binary level-set method is fast enough to be coupled with the rigid-body MC simulations of protein-protein interactions. Our extensive rigid-body MC-VISM simulations of binding of p53-MDM2 have captured some initial binding poses of the complex, and MD simulations starting with such poses quickly reach the final bound state. This indicates that the rigid-body approximations of proteins are rational and efficient in the description of early stages of protein binding process. The protein flexibility is more crucial as two proteins are relatively closer to each other.

Our future studies shall address several issues. One of them is the efficient sampling of different solvation states that can be captured by setting different initial conformations in the VISM relaxation. Another issue is the relaxation of the restrictions arising from the the rigid-body approximations of proteins. In this regard, it is possible to develop some coarse-grained models and implementations. Finally, to speed up the MC simulations, we need to accelerate the sampling of conformations. An immediate next step can be to implement existing speed-up sampling techniques and combine them into our MC-VISM simulations.

Acknowledgments.

This work was supported in part by the US National Science Foundation through the grant DMS-1913144 (L.-T.C. & B.L.), by the US National Institutes of Health (NIH) through the grant R01GM132106 (B.L.), and by a 2019–2020 Lattimer Faculty Research Fellowship, Division of Physical Sciences, University of California, San Diego (B.L.). Work in the McCammon group is supported in part by NIH (GM31749), National Biomedical Computation Resource, and San Diego Supercomputer Center.

References

1. Babine RE; Bender SL Molecular recognition of protein-ligand complexes: Applications to drug design. Chem. Rev 1997, 97, 1359–1472. [PubMed: 11851455]
2. Morra G; Genoni A; Neves MA; Merz KMJ; Colombo G Molecular recognition and drug-lead identification: What can molecular simulations tell us? Curr. Med. Chem 2010, 17, 25–41. [PubMed: 19941480]
3. Baron R; McCammon JA Molecular recognition and ligand association. Annu. Rev. Phys. Chem 2013, 64, 151–175. [PubMed: 23473376]

4. Brooijmans N; Kuntz ID Molecular recognition and docking algorithms. Annu. Rev. Biophys. Biomol. Struct 2013, 32, 335–373.
5. Michel J Current and emerging opportunities for molecular simulations in structure-based drug design. Phys. Chem. Chem. Phys 2014, 16, 4465–4477. [PubMed: 24469595]
6. Sulimov VB; Kutov DC; Sulimov AV Advances in Docking. Curr. Med. Chem 2019, 26, 7555–7580. [PubMed: 30182836]
7. Levy Y; Onuchic J Water mediation in protein folding and molecular recognition. Ann. Rev. Biophys. Biomol. Struct 2006, 35, 389–415. [PubMed: 16689642]
8. Prabhu N; Sharp K Protein-solvent interactions. Chem. Rev 2006, 106, 1616–1623. [PubMed: 16683747]
9. Rasaiah JC; Garde S; Hummer G Water in nonpolar confinment: From nanotubes to proteins and beyond. Annu. Rev. Phys. Chem 2008, 59, 713–740. [PubMed: 18092942]
10. Baron R; Setny P; McCammon JA Water in cavity-ligand recognition. J. Am. Chem. Soc 2010, 132, 12091–12097. [PubMed: 20695475]
11. Hummer G Molecular binding: Under water's influence. Nat. Chem 2010, 2, 906–907. [PubMed: 20966940]
12. Bellissent-Funel M-C; Hassanali A; Havenith M; Henchman R; Pohl P; Sterpone F; van der Spoel D; Xu Y; Garcia AE Water determines the structure and dynamics of proteins. Chem. Rev 2016, 116, 7673–7697. [PubMed: 27186992]
13. Boehr DD; Wright PE How do proteins interact? Science 2008, 320, 1429–1430. [PubMed: 18556537]
14. Okazaki K; Takada S Dynamic energy landscape view of coupled binding and protein conformational change: Induced-fit versus population-shift mechanisms. Proc. Natl. Acad. Sci. USA 2008, 105, 11182–11187. [PubMed: 18678900]
15. Boehr DD; Nussinov R; Wright PE The role of dynamic conformational ensembles in biomolecular recognition. Nat. Chem. Biol 2009, 5, 789–796. [PubMed: 19841628]
16. Hammes GG; Chang Y; Oas TG Conformational selection or induced fit: A flux description of reaction mechanism. Proc. Natl. Acad. Sci. USA 2009, 106, 13737–13741. [PubMed: 19666553]
17. Wlodarski T; Zagrovic B Conformational selection and induced fit mechanism underlie specificity in noncovalent interactions with ubiquitin. Proc. Natl. Acad. Sci. USA 2009, 106, 19346–19351. [PubMed: 19887638]
18. Csermely P; Palotai R; Nussinov R Induced fit, conformational selection an independent dynamic segments: An extended view of binding events. Trends Biochem. Sci, 2010, 35, 539–546. [PubMed: 20541943]
19. Zhou H-X From induced fit to conformational selection: A continuum of binding mechanism controlled by the timescale of conformational transitions. Biophys. J 2010, 98, L15–L17. [PubMed: 20303846]
20. McCammon JA Darwinian biophysics: Electrostatics and evolution in the kinetics of molecular binding. Proc. Natl Acad. Sci. USA 2009, 106, 7683–7684. [PubMed: 19416830]
21. Dzubiella J; Swanson JMJ; McCammon JA Coupling hydrophobicity, dispersion, and electrostatics in continuum solvent models. Phys. Rev. Lett 2006, 96, 087802. [PubMed: 16606226]
22. Dzubiella J; Swanson JMJ; McCammon JA Coupling nonpolar and polar solvation free energies in implicit solvent models. J. Chem. Phys 2006, 124, 084905. [PubMed: 16512740]
23. Cheng L-T; Dzubiella J; McCammon JA; Li B Application of the level-set method to the implicit solvation of nonpolar molecules. J. Chem. Phys 2007, 127, 084503. [PubMed: 17764265]
24. Cheng L-T; Wang Z; Setny P; Dzubiella J; Li B; McCammon JA Interfaces and hydrophobic interactions in receptor-ligand systems: A level-set variational implicit solvent approach. J. Chem. Phys 2009, 131, 144102. [PubMed: 19831428]
25. Cheng L-T; Xie Y; Dzubiella J; McCammon JA; 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. [PubMed: 20150952]
26. Setny P; Wang Z; Cheng L-T; Li B; McCammon JA; Dzubiella J Dewetting-controlled binding of ligands to hydrophobic pockets. Phys. Rev. Lett 2009, 103, 187801. [PubMed: 19905832]

27. Wang Z; Che J; Cheng L-T; Dzubiella J; Li B; McCammon JA Level-set variational implicit solvation with the Coulomb-field approximation. J. Chem. Theory Comput 2012, 8, 386–397. [PubMed: 22346739]
28. Guo Z; Li B; Dzubiella J; Cheng L-T; McCammon JA; Che J Evaluation of hydration free energy by the level-set variational implicit-solvent model with the Coulomb-Field approximation. J. Chem. Theory Comput 2013, 9, 1778–1787. [PubMed: 23505345]
29. Zhou S; Rogers KE; de Oliveira CF; Baron R; Cheng L-T; Dzubiella J; Li B; McCammon JA Variational implicit-solvent modeling of host-guest binding: A case study on Cucurbit[7]uril. J. Chem. Theory Comput 2013, 9, 4195–4204. [PubMed: 24039554]
30. Guo Z; Li B; Dzubiella J; Cheng L-T; McCammon JA; Che J Heterogeneous hydration of p53/MDM2 complex. J. Chem. Theory Comput 2014, 10, 1302–1313. [PubMed: 24803860]
31. Zhou S; Cheng L-T; Dzubiella J; Li B; McCammon JA Variational implicit solvation with Poisson–Boltzmann theory. J. Chem. Theory Comput 2014, 10, 1454–1467. [PubMed: 24803864]
32. 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. [PubMed: 27497546]
33. Ricci CG; Li B; Cheng L-T; Dzubiella J; McCammon JA Tailoring the variational implicit solvent method for new challenges: Biomolecular recognition and assembly. Front. Mol. Biosci 2018, 5.
34. Ricci CG; Li B; Cheng L-T; Dzubiella J; McCammon JA 'Martinizing' the variational implicit solvent method (VISM): Solvation free energy for coarse-grained proteins. J. Phys. Chem. B 2017, 121, 6538–6548. [PubMed: 28613904]
35. Zhou S; Weiß RG; Cheng L-T; Dzubiella J; McCammon JA; Li B Variational implicit-solvent predictions of the dry-wet transition pathways for ligand-receptor binding and unbinding kinetics. Proc. Natl Acad. Sci. USA 2019, 116, 14989–14994. [PubMed: 31270236]
36. Davis ME; McCammon JA Electrostatics in biomolecular structure and dynamics. Chem. Rev 1990, 90, 509–521.
37. Sharp KA; Honig B Electrostatic interactions in macromolecules: Theory and applications. Annu. Rev. Biophys. Chem 1990, 19, 301–332.
38. Baker NA Poisson–Boltzmann methods for biomolecular electrostatics. Method Enzy. Mol 2004, 83, 94–118.
39. Grochowski P; Trylska J Continuum Molecular Electrostatics, Salt Effects and Counterion Binding. A Review of the Poisson-Boltzmann Model and its Modifications. Biopolymers 2008, 89, 93–113. [PubMed: 17969016]
40. Che J; Dzubiella J; Li B; McCammon JA Electrostatic free energy and its variations in implicit solvent models. J. Phys. Chem. B 2008, 112, 3058–3069. [PubMed: 18275182]
41. Cai Q; Wang J; Hsieh MJ; Ye X; Luo R Poisson–Boltzmann implicit solvation models. Annu. Rep. Comput. Chem 2012, 8, 149–162.
42. Bashford D; Case DA Generalized Born models of macromolecular solvation effects. Ann. Rev. Phys. Chem 2000, 51, 129–152. [PubMed: 11031278]
43. Chène P Inhibiting the p53-MDM2 interaction: an important target for cancer therapy. Nat. Rev. Cancer 2003, 3, 102–109. [PubMed: 12563309]
44. Khoo K; Verma C; Lane D Drugging the p53 pathway: understanding the route to clinical efficacy. Nat. Rev. Drug Discov 2014, 13, 217–236. [PubMed: 24577402]
45. Anifowose A; Agbowuro AA; Yang X; Wang B Anticancer strategies by upregulating p53 through inhibition of its ubiquitination by MDM2. Med. Chem. Res 2020, 29, 1105–1121.
46. Duffy MJ; Synnott NC; O'Grady S; Crown J Targeting p53 for the treatment of cancer. Seminars in Cancer Biology 2020,
47. Jiang L; Zawacka-Pankau J The p53/MDM2/MDMX-targeted therapies—a clinical synopsis. Cell Death Dis. 2020, 11, 237. [PubMed: 32303678]
48. Uhrinova S; Uhrin D; Powers H; Watt K; Zheleva D; Fischer P; McInnes C; Barlow PN Structure of free MDM2 N-terminal domains reveals conformational adjustments that accompany p53-binding. J. Mol. Biol 2005, 350, 587–598. [PubMed: 15953616]

49. Zondlo SC; Lee AE; Zondlo NJ Determinants of specificity of MDM2 for the activation domains of p53 and p65: proline27 disrupts the MDM2-binding motif of p53. Biochemistry 2006, 45, 11945–11957. [PubMed: 17002294]
50. Chen HF; Luo R Binding induced folding in p53-MDM2 complex. J. Amer. Chem. Soc 2007, 129, 2930–2937. [PubMed: 17302414]
51. Dastidar SG; Lane DP; Verma CS Multiple peptide conformations give rise to similar binding affinities: molecular simulations of p53-MDM2. J. Amer. Chem. Soc 2008, 130, 13514–13515. [PubMed: 18800837]
52. Borcherds W; Theillet FX; Katzer A; Finzel A; Mishall KM; Powell AT; Wu H; Manieri W; Dieterich C; Selenko P; Loewer A; Daughdrill GW Disorder and residual helicity alter p53-MDM2 binding affinity and signaling in cells. Nat. Chem. Biol 2014, 10, 1000–1002. [PubMed: 25362358]
53. Zhou G; Pantelopulos GA; Mukherjee S; Voelz VA Bridging microscopic and macroscopic mechanisms of p53-MDM2 binding with kinetic network models. Biophys. J 2017, 113, 785–793. [PubMed: 28834715]
54. Ricci CG; McCammon JA Heterogeneous solvation in distinctive protein-protein interfaces revealed by molecular dynamics simulations. J. Phys. Chem. B 2018, 122, 11695–11701. [PubMed: 30252476]
55. Tolman RC The effect of droplet size on surface tension. J. Chem. Phys 1949, 17, 333–337.
56. Chen J; Brooks III CL Critical importance of length-scale dependence in implicit modeling of hydrophobic interactions. J. Am. Chem. Soc 2007, 129, 2444. [PubMed: 17288425]
57. Best RB; Zhu X; Shim J; Lopes PEM; Mittal J; Feig M; MacKerell AD Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ_1 and χ_2 dihedral angles. J. Chem. Theory Comput 2012, 8, 3257–3273. [PubMed: 23341755]
58. Huang J; Rauscher S; Nawrocki G, e. a. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat. Methods 2017, 14, 71–73. [PubMed: 27819658]
59. Lie J; Lysaker M; Tai X-C A binary level set model and some applications to Mumford–Shah image segmentation. IEEE Trans. Image Proc 2006, 15, 1171–1181.
60. Merriman B; Bence J; Osher S Diffsion generated motion by mean curvature. Computational Crystal Growers Workshop. 1992; pp 73–83.
61. Ruuth SJ Efficient algorithms for diffusion-generated motion by mean curvature. J. Comput. Phys 1998, 144, 603–625.
62. Ruuth SJ; Merriman B; Osher S Convolution generated motion as a link between cellular automata and continuum pattern dynamics. J. Comput. Phys 1999, 151, 836–861.
63. Ruuth SJ; Merriman B Convolution-thresholding methods for interface motion. J. Comput. Phys 2001, 169, 678–707.
64. Esdoğlu S; Jacobs M; Zhang P Kernels with prescribed surface tension and mobility for threshold dynamics schemes. J. Comput. Phys 2017, 337, 62–83.
65. Wang D; Li H; Wei X; Wang X-P An efficient iterative thresholding method for image segmentation. J. Comput. Phys 2017, 350, 657–667.
66. Kleinberg J; Tardos E Algorithm Design; Addison-Wesley Longman Publishing Co., Inc.: USA, 2005.
67. Horinek D; Mamatkulov SI; Netz RR Rational design of ion force fields based on thermodynamic solvation properties. J. Chem. Phys 2009, 130, 124507. [PubMed: 19334851]
68. Roux B; Yu HA; Karplus M Molecular basis for the Born mdel of ion solvation. J. Phys. Chem 1990, 94, 4683–4688.
69. Banavali NK; Roux B Atomic radii for continuum electrostatics calculation on nucleic acids. J. Phys. Chem. B 2002, 106, 11026–11035.
70. Mobley DL; Barber II AE; Fennell CJ; Dill KA Charge asymmetries in hydration of polar solutes. J. Phys. Chem. B 2008, 112, 2405–2414. [PubMed: 18251538]

71. Mukhopadhyay A; Fenley AT; Tolokh IS; Onufriev AV Charge hydration asymmetry: The basic principle and how to use it to test and improve water models. J. Phys. Chem. B 2012, 116, 9776–9783. [PubMed: 22762271]

72. Marcus Y Thermodynamics of solvation of ions. Part 5. Gibbs free energy of hydration at 298.15 K. J. Chem. Soc. Faraday Trans 1991, 87, 2995–2999.

73. Zhang Z; Schindler CEM; Lange OF; Zacharias M Application of enhanced sampling Monte Carlo methods for high-resolution protein-protein docking in Rosetta. PLOS One 2015, 10, e0125941. [PubMed: 26053419]

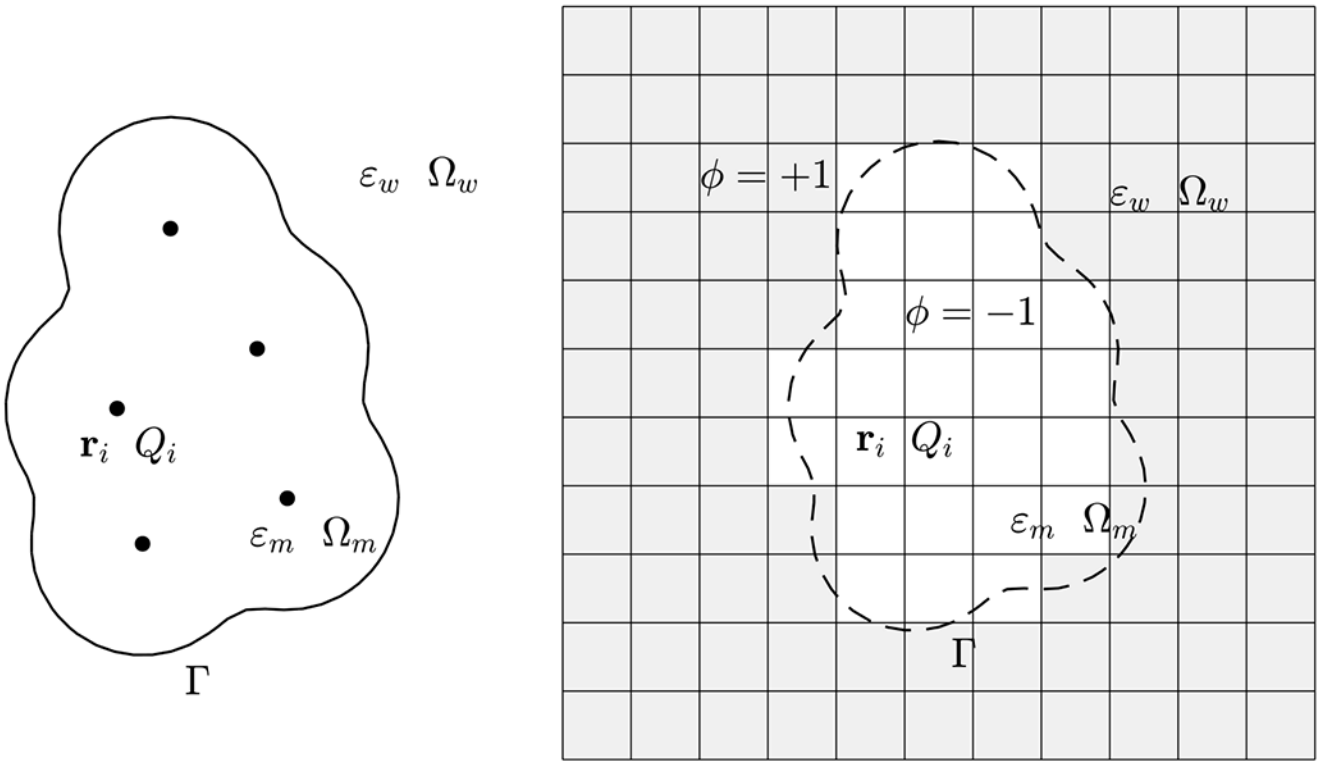

FIG. 1.

Left: 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 $\mathbf{r}_1, \ldots, \mathbf{r}_L$ and carry partial charges $Q_1, \ldots, Q_L$, respectively. The dielectric permittivities of the solute and solvent regions are denoted by ε_m and ε_w, respectively. Right: In the binary level-set formulation, the computational domain is discretized into grid cells. A binary level-set function is used to approximate the dielectric boundary Γ. It takes the value −1 on any grid cell inside the solute region Ω_m and +1 on any cell inside the solvent region Ω_w; cf. Section III A.

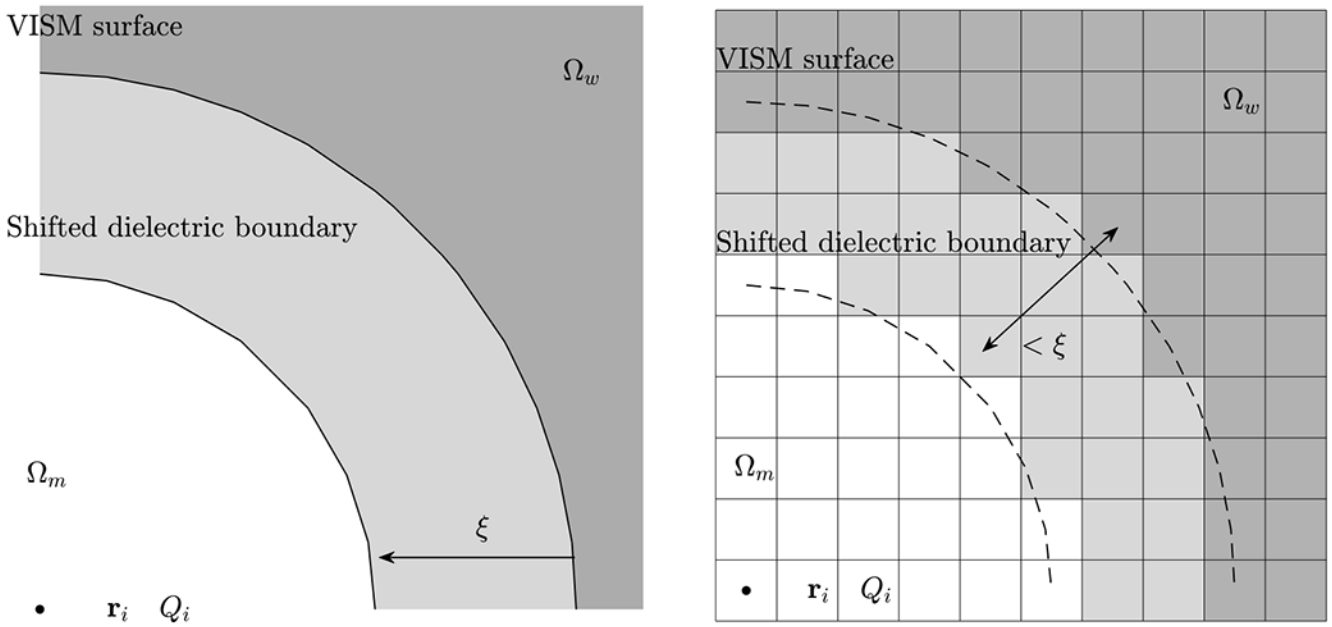

FIG. 2.

Left: An effective dielectric boundary is obtained by shifting the VISM surface inward to the solute region by ξ (Å). Right: In the binary level-set implementation, a grid cell in the solute region Ω_m contributes to the the electrostatic energy, if it has a center-to-center distance less than ξ to some grid cell in the solvent region Ω_w.

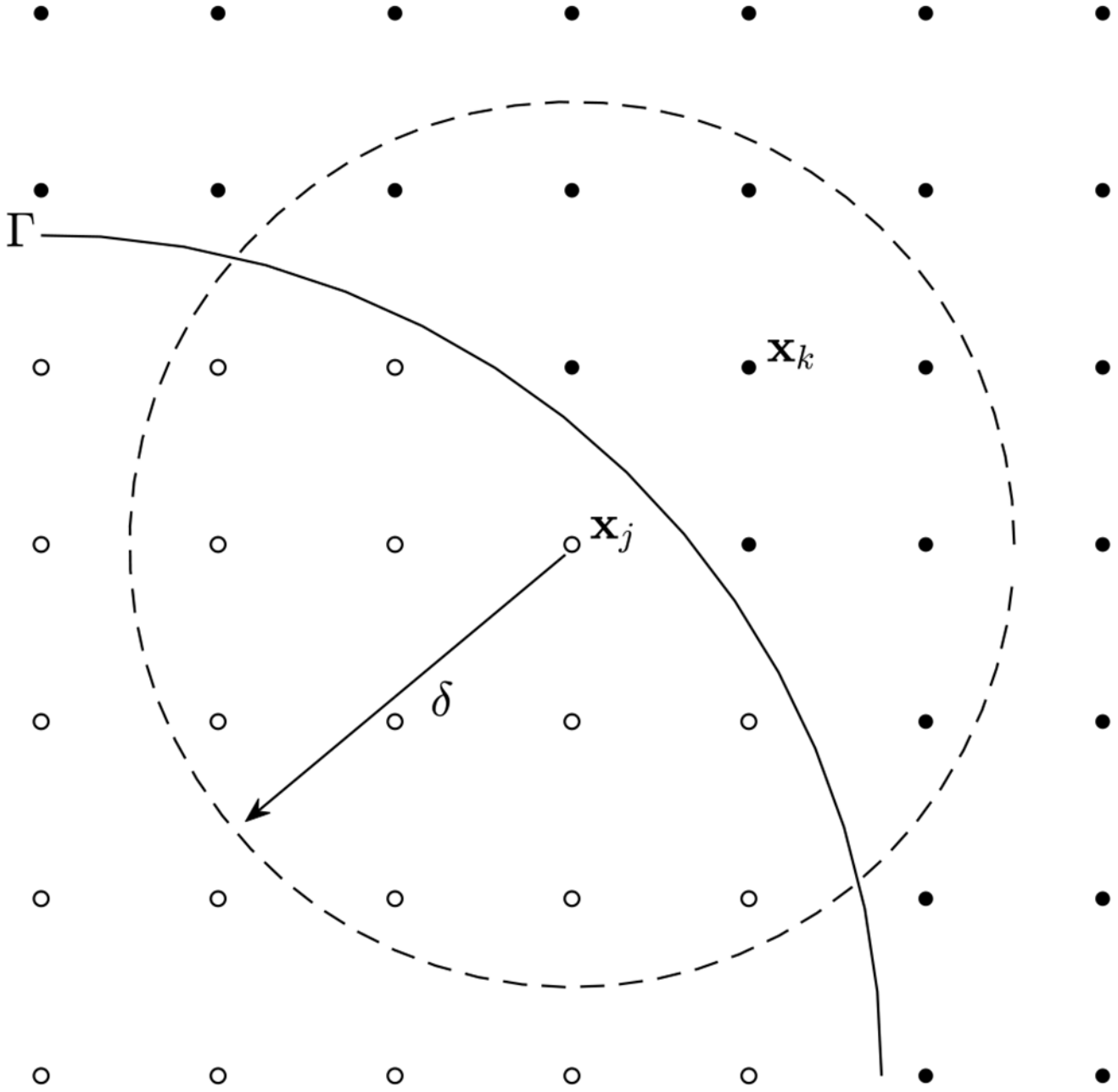

FIG. 3.

Illustration of a scaled kernel centered at the center $\mathbf{x}_i$ of a grid cell and vanishing outside a sphere (indicated by the broken lines). Black dots represent centers of grid cells in the solvent region Ω_w and circles represent the centers of grid cells in the solute region Ω_m.

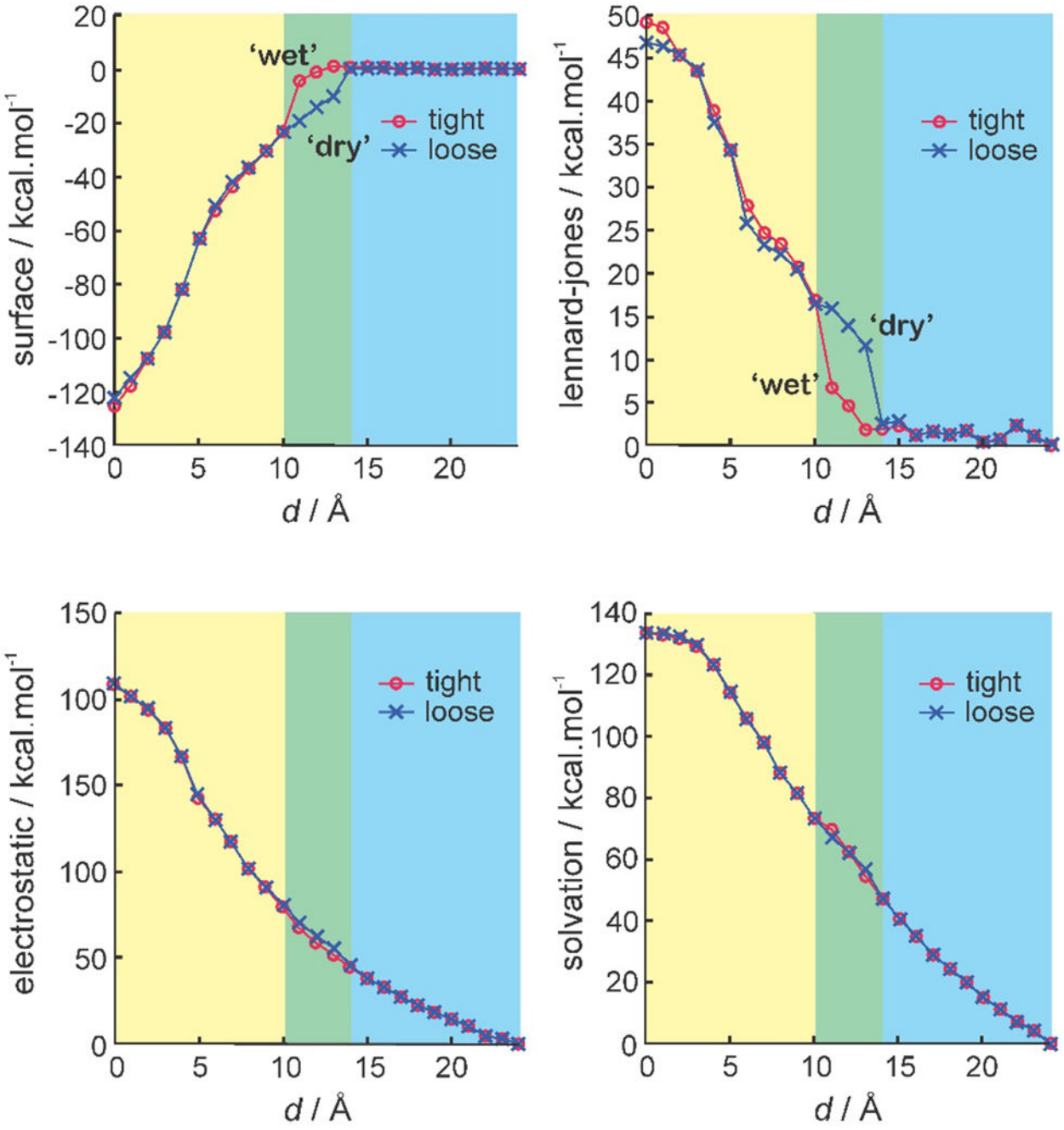

FIG. 4.

Solvation free energy (and relative components) of MDM2 and p53 along the reaction coordinate, *d*, obtained from tight and loose initial conditions. Highlighted in yellow and blue are the regions for which loose and tight calculations converge producing either desolvated or solvated states, respectively, and highlighted in green is the region where tight and loose calculations diverge producing different solvation boundaries depending on the initial conditions ("branching").

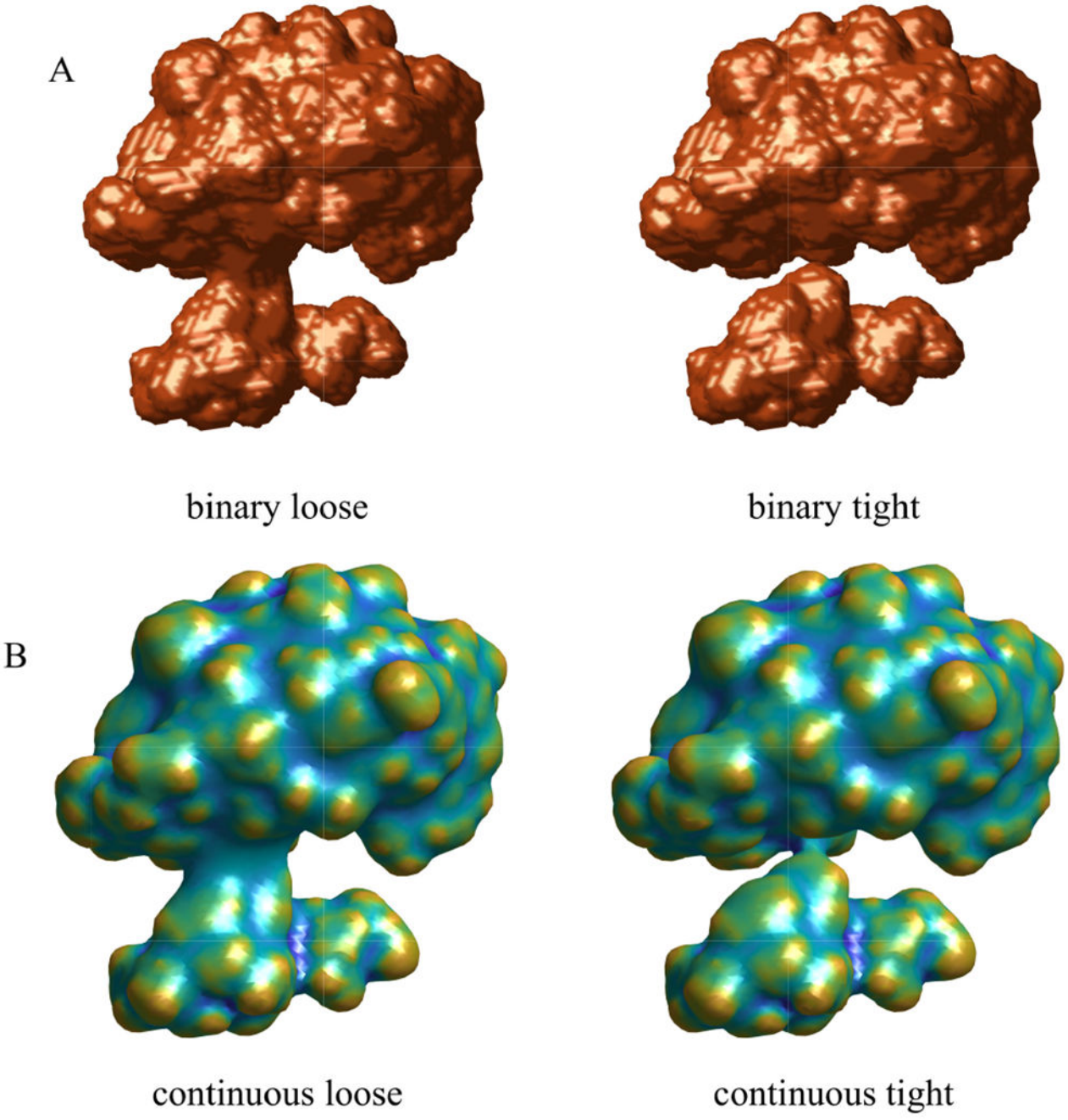

FIG. 5.

Stable equilibrium solute-solvent interfaces of p53-MDM2 obtained at $d = 13$ Å by binary (A) or continuous level-set VISM (B), starting from loose (left) or tight (right) initials. In the surfaces produced by the continuous level-set VISM (B), the color of the surface represents the mean local curvature.

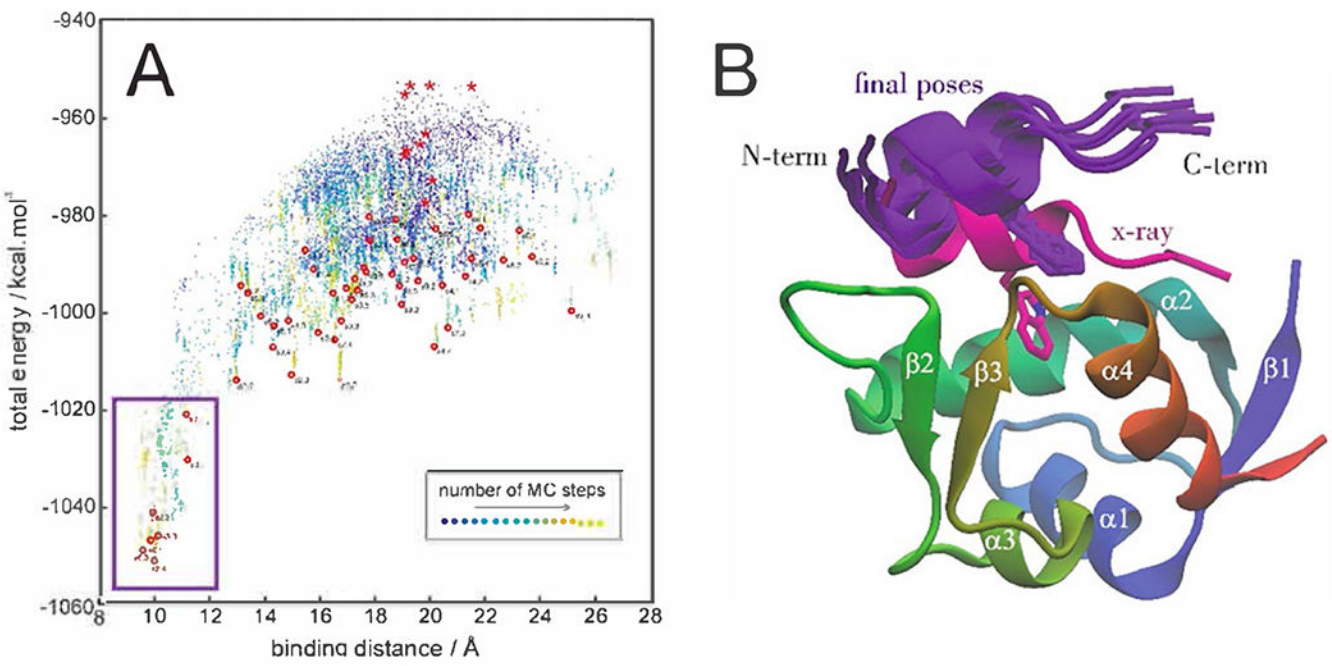

FIG. 6.

(A) Scatter plot of the total energy as a function of the binding distance for all MC-VISM simulations. Initial (randomly generated) configurations are marked by red asterisks; final configurations are marked by red circles; and configurations sampled throughout the simulations are colored from blue to yellow. MC-VISM simulations resulting in productive binding encounters between p53 and MDM2 are highlighted. (B) Superimposition of the final binding poses from productive MC-VISM simulations (purple) and the x-ray complex (PDB ID lycr, in magenta). For reference, the central W23 residue is displayed. MDM2 secondary structure is colored from N- to C-terminal.

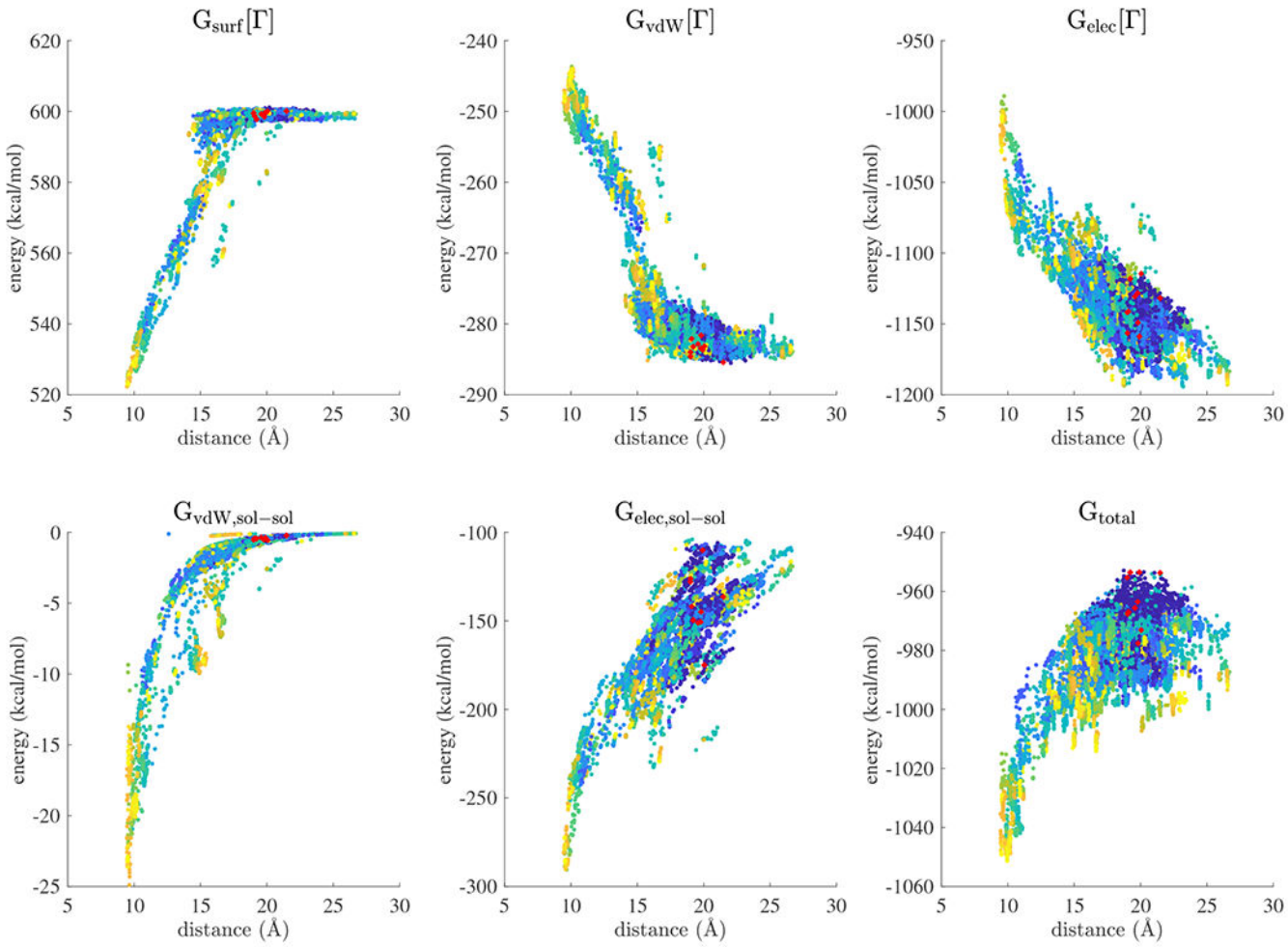

FIG. 7.

Scatter plot of individual energy components of the total binding energy versus the binding distance for all MC-VISM simulations. Initial (randomly generated) configurations are marked by red circles and configurations sampled throughout the simulations are colored from blue to yellow.

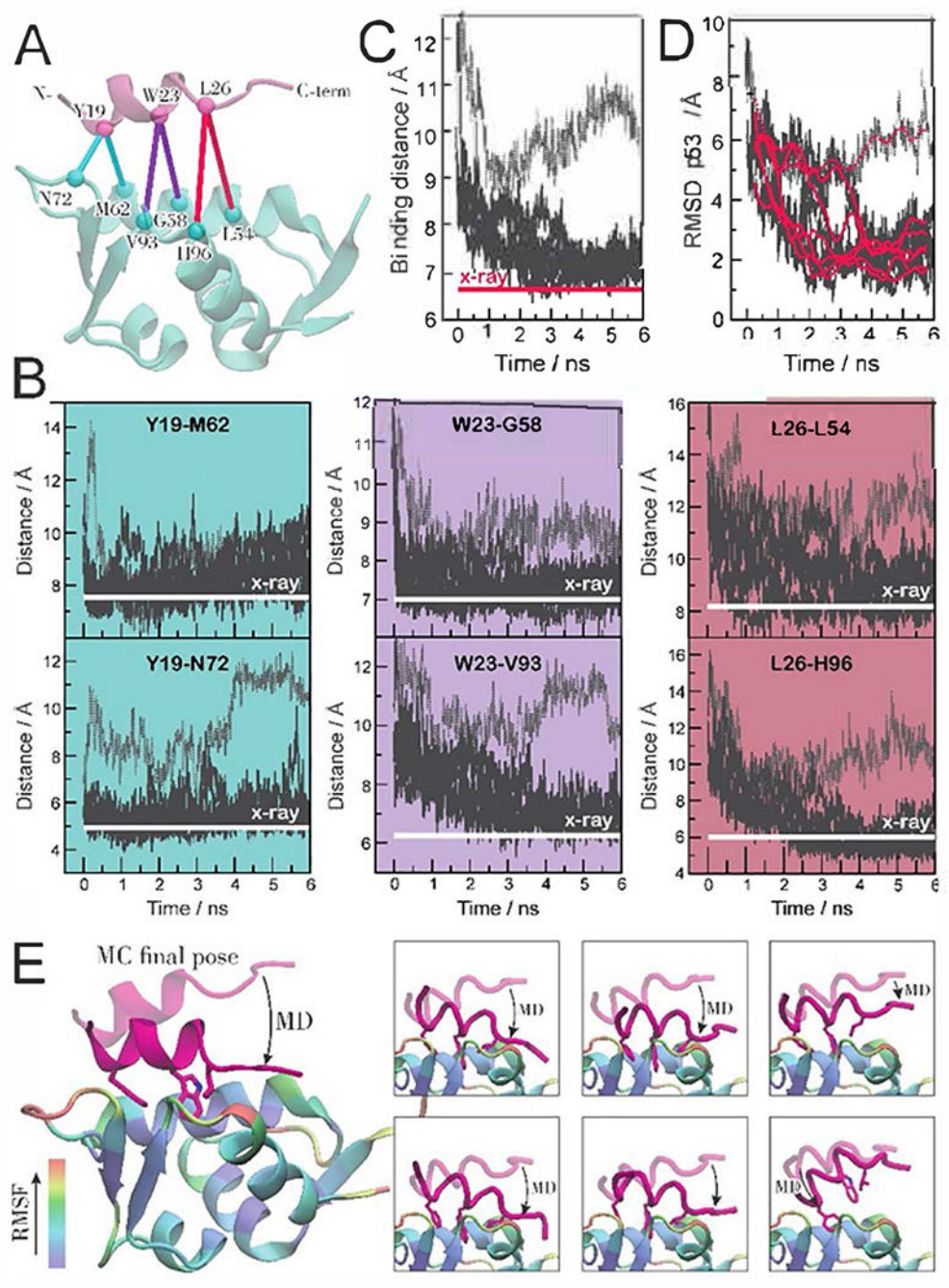

FIG. 8.

(A) Binding distance as defined by the average of six pairwise distances between p53 and MDM2 α-carbons. (B) Evolution of the pairwise distances involving the N-terminal (cyan panels), the central segment (purple panels) and the C-terminal (red panels) of p53 during the MD simulations. As a reference, the distances as measured in the x-ray complex (PDB ID 1ycr) are shown in white. (C) Evolution of the averaged binding distance during the MD simulations. (D) RMSD of p53 α-carbons with respect to the x-ray structure (PDB ID 1ycr). Alignment was performed based on the α-carbons of MDM2. (E) Refined final posed

obtained by MD simulations. MDM2 is colored by its backbone RMSF values obtained during the MD simulations. Six out of seven MD simulations rapidly produce the correct binding mode by means of insertion of the C-terminal portion of p53 into MDM2 binding cleft.

TABLE I.

The VISM parameters.

Parameter	Symbol	Value	Unit
temperature	T	298	K
solvent number density	ρ_w	0.0333	$Å^{-3}$
surface tension	γ_0	0.174	$k_BT/Å^2$
solute dielectric constant	ε_m	1	
solvent dielectric constant	ε_w	80	
boundary shift	ξ	1	Å

TABLE II.

Solvation free energy (in $k_B T$) and its components for a particle with different charge values Q (in e).

Q	surface energy			vdW energy			electrostatic energy		
	analytical	continuous	binary	analytical	continuous	binary	analytical	continuous	binary
0	20.51	20.68	20.28	−2.64	−2.76	−2.66	0.00	0.00	0.00
0.5	19.27	19.43	19.16	−1.05	−1.24	−1.26	−23.17	−23.08	−23.10
1	16.89	17.01	16.72	5.11	4.78	4.89	−99.01	−98.65	−98.87

TABLE III.

The solvation free energy (in k_BT) for each of the single ions K^+, Na^+, Cl^-, and F^-: the level-set VISM calculations vs. experiment.[72]

ions	ϵ (k_BT)	σ (Å)	experiment	continuous level-set	binary level-set
K^+	0.008	3.85	−117.5	−112.3	−103.1
Na^+	0.008	3.49	−145.4	−131.1	−123.1
Cl^-	0.21	3.78	−135.4	−126.7	−113.4
F^-	0.219	3.3	−185.2	−171.9	−158.7

TABLE IV.

Solvation free energy (k_BT) and computation time (s) for different grid numbers with $Q = 0$. Here, cont. stands for the continuous level-set method and binary for the binary level-set method.

Grid number	Surface energy		vdW energy		Time		
	cont.	binary	cont.	binary	cont. (CPU)	binary (CPU)	binary (GPU)
25^3	21.46	20.64	−2.86	−3.31	1.10	0.14	0.01
50^3	20.87	20.45	−2.78	−3.02	11.97	3.03	0.10
100^3	20.68	20.28	−2.76	−2.66	186.44	52.30	1.41
200^3	20.80	20.37	−2.91	−2.68	5032.03	1198.97	26.11

TABLE V.

Summary of all simulations: the minimum binding distance (Å) and minimum total binding free energy (kcal/mol) for all the trials. For the initial conformation, 10 trials are performed; cf. the rows s0a and s0b. For each of the other initial configuration, 5 trials are performed.

initial	min binding dist					min total binding free energy				
s0a	9.8	18.8	15.9	15.8	18.6	−1047.18	−988.76	−1002.05	−980.55	−981.13
s0b	15.1	18.1	18.8	17.7	19.1	−978.75	−974.63	−974.46	−972.26	−976.81
s1	19.8	17.2	15.9	15.5	15.2	−982.90	−993.36	−1014.14	−991.34	−996.30
s2	11.1	9.5	14.9	9.8	12.8	−1021.23	−1041.45	−1013.12	−1051.32	−1014.25
s3	10.9	15.0	10.1	15.1	14.9	−1030.65	−987.51	−1046.41	−1004.48	−997.68
s4	19.4	19.1	14.1	19.7	16.8	−994.73	−992.93	−1001.94	−1007.26	−995.31
s5	14.2	15.9	17.2	19.3	13.1	−1003.03	−991.89	−995.71	−989.15	−996.28
s6	13.6	12.7	9.4	15.8	18.4	−1001.05	−994.82	−1049.21	−985.41	−989.17
s7	17.3	18.9	17.2	16.5	15.8	−989.93	−1003.54	−992.42	−1005.80	−991.07
s8	19.2	19.2	17.9	19.9	18.7	−980.08	−989.50	−985.27	−983.36	−983.07
s9	19.4	18.7	16.1	13.9	15.0	−999.90	−993.79	−998.63	−1007.38	−994.79