Variational implicit-solvent predictions of the dry-wet transition pathways for ligand-receptor binding and unbinding kinetics

Shenggao Zhou[a], **R. Gregor Weiß**[b], **Li-Tien Cheng**[c], **Joachim Dzubiella**[d], **J. Andrew McCammon**[e,1], **and Bo Li**[c,1]

[a]Department of Mathematics and Mathematical Center for Interdiscipline Research, Soochow University, 1 Shizi Street, Suzhou 215006, Jiangsu, China; [b]Laboratory of Physical Chemistry, ETH Zürich, Vladimir-Prelog-Weg 2, CH-8093 Zürich, Switzerland; and Institut für Physik, Humboldt-Universität zu Berlin, Newtonstrasse 15, D-12489 Berlin, Germany; [c]Department of Mathematics, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0112, USA; [d]Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Hermann-Herder-Straße 3, 79104 Freiburg, Germany; and Research Group Simulations of Energy Materials (EE-GSEM), Helmholtz-Zentrum Berlin, Hahn-Meitner-Platz 1, 14109, Berlin, Germany; [e]Department of Chemistry and Biochemistry, Department of Pharmacology, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0365, USA

This manuscript was compiled on June 27, 2019

Ligand-receptor binding and unbinding are fundamental biomolecular processes and particularly essential to drug efficacy. Environmental water fluctuations, however, impact the corresponding thermodynamics and kinetics and thereby challenge theoretical descriptions. Here, we devise a holistic, implicit-solvent, multi-method approach to predict the (un)binding kinetics for a generic ligand-pocket model. We use the variational implicit-solvent model (VISM) to calculate the solute-solvent interfacial structures and the corresponding free energies, and combine the VISM with the string method to obtain the minimum energy paths and transition states between the various metastable ("dry" and "wet") hydration states. The resulting dry-wet transition rates are then used in a spatially-dependent multi-state continuous-time Markov chain Brownian dynamics simulations, and the related Fokker–Planck equation calculations, of the ligand stochastic motion, providing the mean first-passage times for binding and unbinding. We find the hydration transitions to significantly slow down the binding process, in semi-quantitative agreement with existing explicit-water simulations, but significantly accelerate the unbinding process. Moreover, our methods allow the characterization of non-equilibrium hydration states of pocket and ligand during the ligand movement, for which we find substantial memory and hysteresis effects for binding versus unbinding. Our study thus provides a significant step forward towards efficient, physics-based interpretation and predictions of the complex kinetics in realistic ligand-receptor systems.

Ligand-receptor binding/unbinding kinetics | dry-wet transitions | variational implicit-solvent model | level-set method | string method

The complex process of ligand-receptor binding and unbinding in aqueous environment is fundamental to biological function. Understanding the thermodynamics and kinetics of such processes has far-reaching practical significance, particularly in rational drug design (1, 2). Water is a key player in ligand-receptor binding and unbinding, and in molecular recognition in general (3, 4). In particular, it has been well established that hydrophobic interactions can drive the association and dissociation of biological molecules (5–8).

Hydration contributes significantly to the ligand-receptor binding free energy, determining the thermodynamic stability of the bound unit (9, 10). Recent experimental and theoretical studies have indicated that the kinetics of ligand-receptor binding and unbinding is crucial for drug effectiveness and efficacy (2, 11, 12). Often, a ligand binds to a hydrophobic pocket on the surface of a receptor molecule (13–16). Water molecules fluctuate around such an apolar pocket, leading to metastable "dry" or "wet" hydration states of the binding site, separated by an energetic barrier which is on the order of k_BT (17). Such a moderate energetic hurdle facilitates repeated condensation and evaporation of water in the pocket region, leading to large collective hydration fluctuations (18). In general, the dewetting of local regions generates strong hydrophobic forces in molecular association and dissociation (6, 7, 19, 20). In particular, it has been demonstrated that the dry-wet transitions are a precursor of the ligand-receptor binding and unbinding (17, 21, 22). Besides being the origin for the thermodynamically driven forces, water fluctuations also modify the friction and kinetics of associating hydrophobic molecules (23–27), slowing down the binding kinetics and giving rise to local non-Markovian effects (18, 27).

While water plays a critical role in molecular recognition, efficient modeling of water is rather challenging due to an overwhelming number of solvent degrees of freedom, many-body effects, and the multi-scale nature of molecular interactions. Explicit-water molecular dynamics (MD) simulations have been the main tool in most of the existing studies of the kinetics of ligand-receptor binding and unbinding (18, 22, 25, 26, 28–

Significance Statement

The kinetics of ligand-receptor (un)binding—how fast a ligand binds into and resides in a receptor—cannot be inferred solely from the binding affinity which describes the thermodynamic stability of the bound complex. A bottleneck in understanding such kinetics, which is critical to drug efficacy, lies in the modeling of the collective water fluctuations in apolar confinement. We develop a new theoretical approach that couples a variational implicit-solvent model with the string method to describe the dry-wet transition pathways, which then serve as input for the ligand multi-state Brownian dynamics. Without explicit descriptions of individual water molecule, our theory predicts the key thermodynamic and kinetic properties of unbinding and binding, the latter in quantitative agreement with explicit-water molecular dynamics simulations.

JD, JAM, and BL designed research. SZ, RGW, and LTC performed research. SZ, LTC, and BL developed numerical methods and LTC wrote the initial level-set code. SZ, RGW, and JD analyzed computational results. SZ, RGW, JD, JAM, and BL wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: jmccammon@ucsd.edu or bli@math.ucsd.edu

33). While explicitly tracking water molecules, MD simulations are still limited to systems of relatively small sizes and events of relatively short time scales. In particular, slow and rare water fluctuations and large ligand residence times in the pocket still challenge the prediction of unbinding times.

In this work, we develop a holistic, multi-method, implicit-solvent approach to study the kinetics of ligand-receptor binding and unbinding in a generic pocket-ligand model exactly as studied previously by explicit-water MD simulations (18), focusing on the effect of solvent fluctuations and multiple hydration states on such processes.

Our approach is based on the variational implicit-solvent model (VISM) that we have developed in recent years (34–38). In VISM, one minimizes a solvation free-energy functional of solute-solvent interfaces to determine a stable, equilibrium conformation, and to provide an approximation of the solvation free energy. The functional couples the solute surface energy, solute-solvent van der Waals (vdW) dispersive interactions, and electrostatics. This theory resembles that of Lum–Chandler–Weeks (39) [cf. also (40, 41)], and is different from the existing SAS (solvent-accessible surface) type models. We have designed and implemented a robust level-set method to numerically minimize the VISM functional with arbitrary 3D geometry (36–38, 42).

Here, for our model ligand-pocket system, we use our level-set VISM to obtain different hydration states and their solvation free energies, and use the VISM-string method (43, 44) to find the minimum energy paths connecting such states and the corresponding transition rates. Such rates are then used in our continuous-time Markov chain Brownian dynamics simulations, and the related Fokker–Planck equation calculations, of the ligand stochastic motion to obtain the mean first-passage times for the ligand binding and unbinding. We compare our results with existing explicit-water MD simulations.

The model ligand-receptor system. The generic pocket-ligand model (45) consists of a hemispherical pocket and a methane-like molecule; cf. Fig. 1 (A). The pocket, with the radius $R = 8$ Å and centered at $(0, 0, 0)$, is embedded in a rectangular wall, composed of apolar atoms aligned in a hexagonal close-packed grid of lattice constant 1.25 Å. The wall surface is oriented in xy-plane. The ligand, a single neutral Lennard-Jones (LJ) sphere, is placed along the pocket symmetry axis, the z-axis, which is taken to be the reaction coordinate. Fig. 1 (B)–(D) depict the cross sections of all the possible VISM surfaces, i.e., the stable solute-solvent interfaces separating the solute region $\Omega_{\rm m}$ and solvent region $\Omega_{\rm w}$, representing different hydration states for a fixed position of ligand.

Results and Analysis

Multiple hydration states and the potential of mean force (PMF). We use our level-set method to minimize the VISM solvation free-energy functional (cf. Eq. [2] in Theory and Methods) and obtain a VISM surface. By choosing different initial solute-solvent interfaces, we obtain different VISM surfaces describing different hydration states; cf. Fig. 1.

Fig. 2 (A) shows the solvation free energies for different VISM surfaces against the reaction coordinate z. For $z < -0.5$ Å, there is only one VISM surface, 1s-dry; cf. Fig. 1 (B). In addition to 1s-dry, a second VISM surface, 2s-wet, appears for $-0.5 < z < 5$ Å; cf. Fig. 1 (D). For $5 < z < 8$ Å, there are

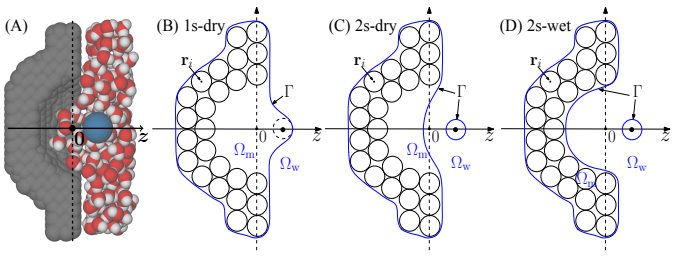

Fig. 1. (A) A schematic of the ligand (blue sphere), explicit water, and the pocket of a concave wall. (B) 1s-dry: The VISM surface (blue line) is a single surface enclosing all the wall atoms and also the ligand atom, hence a dry state of the pocket. (C) 2s-dry: The VISM surface has two disjoint components, one enclosing all the wall atoms with a dry pocket, and one enclosing the ligand. (D) 2s-wet: The VISM surface has two components, tightly wrapping up the wall and ligand, respectively, with no space for water, hence a wet pocket.

three VISM surfaces. In addition to 1s-dry and 2s-wet, the third one is 2s-dry; cf. Fig. 1 (C). Once the ligand is away from the pocket with $z > 8$ Å, there are only two VISM surfaces: 2s-dry and 2s-wet.

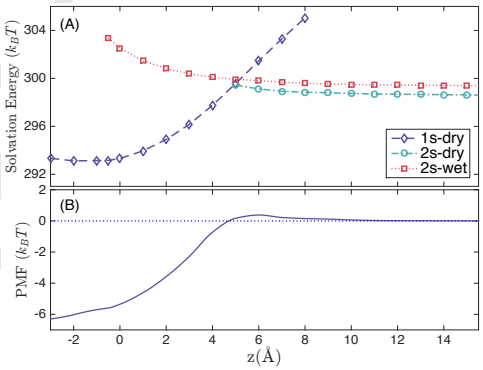

Fig. 2. (A) Solvation free energies of different VISM surfaces vs. the ligand location. (B) The equilibrium PMF.

Fig. 2 (B) shows the equilibrium PMF, defined as

$$V(z) = -k_{\rm B}T \ln\left(\sum_{\Gamma(z)} e^{-G[\Gamma(z)]/k_{\rm B}T}\right) + U_0(z) + V_\infty, \qquad [1]$$

where $\Gamma(z)$ runs over all the VISM surfaces with $G[\Gamma(z)]$ the VISM solvation free energy at $\Gamma(z)$, and $U_0(z) = \sum_{\mathbf{r}_i} U_{\rm LJ}(|\mathbf{r}_i - \mathbf{r}_z|)$ with $\mathbf{r}_z$ the ligand position vector, $\mathbf{r}_i$ running through all the wall atoms, and $U_{\rm LJ}(r)$ a 12–6 LJ potential. The constant V_∞ is chosen so that $V(\infty) = 0$. The PMF agrees well with the result from MD simulations (17, 46, 47).

Dry-wet transition paths and energy barriers. At a fixed reaction coordinate z with multiple hydration states, we use our level-set VISM coupled with the string method to calculate the minimum energy paths (MEPs) that connect these states, and the corresponding transition states, energy barriers, and ultimately the transition rates. A string or path here consists of a family of solute-solvent interfaces, and each point of a string, which is an interface in our case, is called an image.

In Fig. 3, we display the solvation free energies of images on MEPs that connect the three hydration states, 1s-dry, 2s-dry, and 2s-wet, at $z = 6$ Å. There are two MEPs connecting 1s-dry (marked (I)) and 2s-dry (marked (IV)). One of them passes

through the axisymmetric transition state marked (III), and the other passes through the axiasymmetric transition state marked (II). Here, symmetry or asymmetry refers to that of the 3D conformation of the VISM surface. Energy barriers in the transition from the state 1s-dry to 2s-dry along the two transition paths are estimated to be 1.09 k_BT and 0.52 k_BT, respectively. Only one MEP is found to connect 2s-dry (marked (IV)) and 2s-wet (marked (VI)), and the corresponding transition state (marked (V)) is also found. The MEP from 1s-dry to 2s-wet always passes through the state 2s-dry.

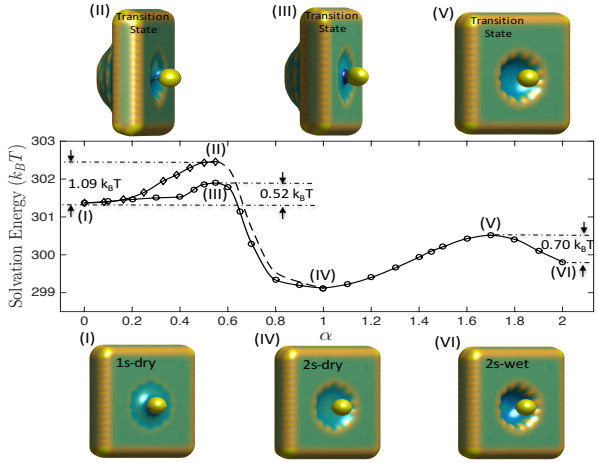

Fig. 3. Solvation free energies of images on MEPs that connect the hydration states 1s-dry (I), 2s-dry (IV), and 2s-wet (VI) (shown in the bottom) with transition states (II), (III), and (V) (shown on top) and the transition energy barriers for $z = 6$ Å. In the middle plots, the horizontal axis is the string parameter α.

Fig. 4 summarizes all the energy barriers in the transitions from one hydration state to another for each reaction coordinate z. For $0 \le z \le 4$ Å shown in the top of Fig. 4, there are only two hydration states: 1s-dry and 2s-wet. The 1s-dry has a lower free energy; cf. Fig. 2 (A), and hence the barrier in the wetting transition from 1s-dry to 2s-wet (shown in red) is higher than that in the dewetting transition from 2s-wet to 1s-dry (shown in blue). The dewetting barrier first increases as the ligand approaches the entrance of the pocket (from $z = 4$ to $z = 1$ Å), and then decreases after the ligand enters the pocket (from $z = 1$ to $z = -0.5$ Å). This is because that more attractive solute-solvent vdW interaction is lost in dewetting as the ligand-pocket distance reduces from $z = 4$ to $z = 1$ Å, and that the decrease in interfacial energy outweighs the vdW contribution to the solvation free energy as the distance further reduces from $z = 1$ to $z = -0.5$ Å. Our predictions agree well with those by the explicit-water MD simulations (17).

For $5 \le z \le 8$ Å, there are three hydration states 1s-dry, 2s-wet, and 2s-dry; cf. Fig. 2 (A). In the middle of Fig. 4, we plot for z in this range the energy barriers along the MEPs, both axisymmetric and axiasymmetric, connecting the two states 1s-dry and 2s-dry; cf. Fig. 3. Note that, as the ligand approaches the pocket, the solute-solvent interfacial energy changes rapidly, and hence the barrier in the transition from 1s-dry to 2s-dry increases quickly, while the barrier in the reverse transition decreases quickly.

In the bottom of Fig. 4, we plot energy barriers for transitions between the states 2s-dry and 2s-wet in the range $5 \le z \le 12$ Å; cf. Fig. 2 (A). As the ligand-pocket distance increases, the barrier for the wetting transition (marked red)

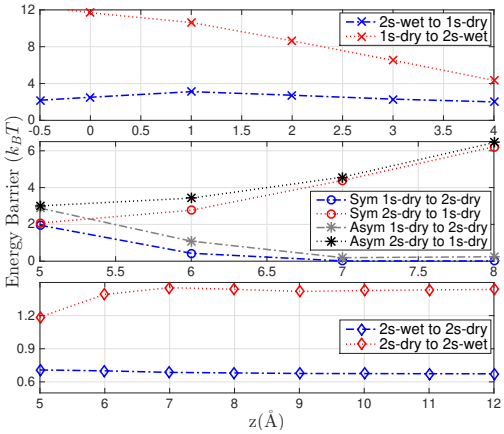

Fig. 4. Transition energy barriers vs. the reaction coordinate z with $-0.5 \le z \le 4$ Å (top) and $5 \le z \le 12$ Å (middle and bottom). Sym or Asym stands for a MEP with an axisymmetric or axiasymmetric transition state.

first increases, since the newly created solvent region with attractive solute-solvent vdW interaction decreases. It then reaches a plateau after the distance is greater than 7 Å. The pocket dewetting barrier (marked blue) is slightly larger when the ligand is close to the pocket, since contributions of solute-solvent vdW interaction are lost during the pocket dewetting.

Kinetics of binding and unbinding. We perform continuous-time Markov chain (CTMC) Brownian dynamics (BD) simulations and solve the related Fokker–Planck equation (FPE) calculations for the ligand stochastic motion with the pocket dry-wet fluctuations; see Theory and Methods. For comparison, we also perform the usual BD simulations and FPE calculations without including such fluctuations.

Fig. 5 (A) and (B) show the mean first-passage times (MFPTs) for the binding and unbinding, respectively. Note that the BD simulations and FPE calculations agree with each other perfectly for both binding and unbinding, without and with the pocket dry-wet fluctuations, respectively. This validates mutually the accuracy of our numerical schemes. Note also that the binding/unbinding MFPT increases/decreases monotonically as the ligand-pocket distance increases, due to elongated/shortened ligand travel.

In Fig. 5 (A), we see that the MFPT for binding is very small if $z < -0.5$ Å. This is because the ligand diffusion constant $D_{\rm in}$ inside the pocket is large and the PMF is highly attractive; cf. Fig. 2 (B). As the initial position z increases from 0 Å to 5 Å, the difference between the two MFPTs with and without the pocket dry-wet fluctuations increases from nearly 0 ps to 100 ps. Such an increasing difference results from the existence of the hydration state 2s-wet in this range, and the solvation free energy of this state increases as the ligand moves from $z = 5$ Å to $z = 0$ Å; cf. Fig. 2 (A). The pocket dry-wet fluctuations thus decelerate considerably the ligand-pocket association. Such deceleration has been explained by the reduced diffusivity of the ligand in the vicinity of pocket entrance due to the slow solvent fluctuations (18).

Our predictions of the MFPT for binding, with the dry-wet fluctuations included, agree very well with the explicit-water MD simulations (18), improving significantly over those without such fluctuations. Note that our model predicts somewhat shorter binding times than the MD simulations for $1 < z < 6$ Å.

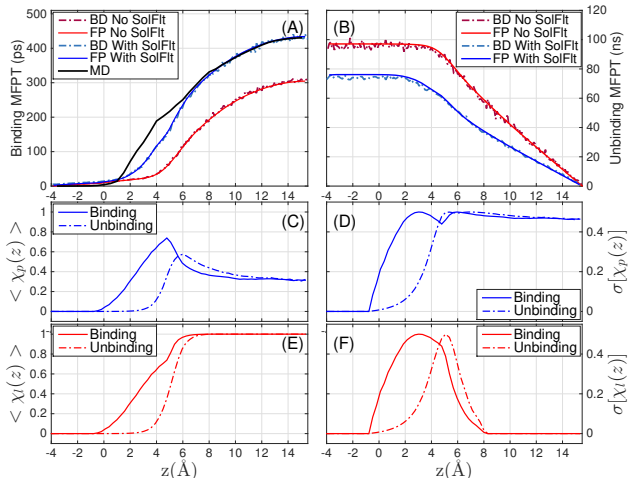

Fig. 5. The MFPT for: (A) the binding of ligand that starts from $z_{\rm init} = z$ and reaches the pocket at $z_{\rm L} = -4$ Å; and (B) the unbinding of ligand that starts from $z_{\rm init} = z$ and reaches $z_{\rm R} = 15.5$ Å, predicted by: BD simulations without (BD No SolFlt) and with (BD With SolFlt) the dry-wet fluctuations; and FPE calculations without (FP No SolFlt) and with (FP With SolFlt) the dry-wet fluctuations, respectively. Note that the time unit on the vertical axis in (B) is ns while that in (A) is ps. The MFPT obtained by explicit-water MD simulations (MD) (18) is also shown in (A). (C)–(F) The mean values and standard deviations of the pocket and ligand hydration states $\chi_p(z)$ and $\chi_l(z)$, respectively, against the ligand location z during the *nonequilibrium* binding process from the BD simulations starting at $z_{\rm init} = 6$ Å (cf. (C) and (E)) and the unbinding process starting at $z_{\rm init} = -2$ Å (cf. (D) and (F)).

In this region, the hydration fluctuations are maximal, and this visible but relatively small (when compared to the MFPT from the farthest distance) discrepancy reflects some of the approximations of our implicit-solvent theory and the model reduction on just a few states.

Fig. 5 (B) shows that the timescale for unbinding is significantly larger than that of the binding, by nearly three orders of magnitude. Without the pocket dry-wet fluctuations, the unbinding MFPT is constant for $z < 4$ Å and decreases linearly for $z > 4$ Å. Note that the MFPT for binding in this case also starts to increase significantly at $z = 4$ Å; cf. Fig. 5 (A). With the pocket dry-wet fluctuations, the unbinding MFPT is much smaller, since the solvation free energy of the 2s-wet state is higher when the ligand is closer to the pocket (cf. Fig. 2 (A)), favoring the ligand unbinding. In this case, the MFPT remains constant up to $z = 2$ Å and then decays almost linearly. This suggests that the wetting transitions occur if $z > 2$ Å. Note from Fig. 5 (A) that the binding MFPT starts increasing rapidly also around $z = 2$ Å.

We now study the interesting hydration of the pocket and ligand individually during the *non-equilibrium* binding/unbinding processes. For this, we define a pocket hydration parameter to be $\chi_{\rm p}(z) = 0$ or 1 if the pocket is dry or wet, respectively. Analogously, we set for the ligand $\chi_{\rm l}(z) = 0$ or 1 if the ligand is dry or wet, respectively. The values 0 and 1 of these ligand-position dependent random variables $\chi_{\rm p}(z)$ and $\chi_l(z)$ are defined by the three hydration states 1s-dry, 2s-dry, and 2s-wet (cf. Fig. 1 (B)–(D)) as follows:

$\chi_{\rm p}(z) = 0$ and $\chi_l(z) = 0$ for a 1s-dry VISM surface;
$\chi_{\rm p}(z) = 0$ and $\chi_l(z) = 1$ for a 2s-dry VISM surface;
$\chi_{\rm p}(z) = 1$ and $\chi_l(z) = 1$ for a 2s-wet VISM surface.

Fig. 5 (C)–(F) show the mean values, $\langle\chi_{\rm p}(z)\rangle$ and $\langle\chi_{\rm l}(z)\rangle$, and the standard deviations, $\sigma[\chi_{\rm p}(z)]$ and $\sigma[\chi_l(z)]$, during the binding and the unbinding processes, respectively.

When the ligand is far away, there are only two VISM surfaces, 2s-dry and 2s-wet, cf. Fig. 2 (A). For such a case, our BD simulations predict the probability 32% of a wet pocket (i.e., $\chi_p = 0.32$ for large z) in the binding and unbinding processes. This is perfectly consistent with the equilibrium probability $e^{-G[\Gamma_{\rm 2s-wet}]/k_{\rm B}T}/(e^{-G[\Gamma_{\rm 2s-dry}]/k_{\rm B}T} + e^{-G[\Gamma_{\rm 2s-wet}]/k_{\rm B}T})$ predicted by our VISM theory. We observe that the pocket hydration peaks at the entrance of the pocket in binding, agreeing well with MD simulations (17, 18), where it was argued that stronger pocket hydration is induced by the penetration of the ligand solvation shell. When the ligand enters the pocket the latter becomes dry as anticipated.

In comparison, the maximum pocket hydration for unbinding is shifted a bit away from the pocket. This kinetic asymmetry or "translational mismatch" can be explained as well by the asymmetric hydration states of the ligand, see Fig. 5 (E), which exits the pocket without a complete solvation shell. This behavior is reminiscent of a hysteresis, that is, the hydration states during the ligand passage depend on the history of the ligand, i.e., where it comes from.

The standard deviations of pocket hydration shown in Fig. 5 (D) depict that the dry-wet fluctuations have local maxima close to the pocket entrance ($z \simeq 3-5$ Å) and behave also significantly different for binding and unbinding. The corresponding standard deviations of ligand hydration shown in Fig. 5 (F) show massively unstable hydration (i.e., large peaks) close to the pocket entrance, while inside and far away from the pocket the fluctuations are zero, indicating a very stable (de)hydration state. Again the peaks are at different locations for binding versus unbinding, reflecting the hysteresis and memory of dry-wet transitions during ligand passage.

Conclusions

We have developed an implicit-solvent approach, coupling our VISM, the string method, and multi-state CTMC BD simulations, for studying the kinetics of ligand-receptor binding and unbinding, particularly the influence of collective solvent fluctuations on such processes. Without any explicit descriptions of individual water molecules, our predictions of the MFPT for the binding process, which is decelerated by the solvent fluctuations around the pocket, agree very well with the less efficient explicit-water MD simulations. Moreover, we find surprisingly that the solvent fluctuations accelerate the ligand unbinding from the pocket, which involves a much larger timescale and is thus more challenging for explicit-water MD simulations (26, 30). Importantly, our implicit-solvent approach indicates that the water effects are controlled by a few key physical parameters and mechanisms, such as polymodal nano-capillarity based on surface tension of the solute-solvent interface and the coupling of the random interface forces to the ligand's diffusive motion.

Our approach provides a promising new direction in efficiently probing the kinetics, and thermodynamics, of the association and dissociation of complex ligand-receptor systems, which have been studied mostly using enhanced sampling techniques (18, 25, 26, 28, 30, 32). Our next step is to extend our approach for more realistic systems with general reaction coordinates and different techniques for sampling transition paths (48, 49). Our VISM can treat efficiently the electrostatic interactions using the Poisson–Boltzmann theory (38). To account for the flexibility of the ligand and receptor in their

binding and unbinding, we shall expand our solvation model to include the solute molecular mechanical interactions (50).

Theory and Methods

Variational implicit-solvent model (VISM). We consider the solvation of solute molecules, with all the solute atomic positions $\mathbf{r}_1, \ldots, \mathbf{r}_N$, in an aqueous solvent that is treated implicitly as a continuum. (For our model ligand-pocket system, the solute atoms include those of the concave wall and the single atom of the ligand; cf. Fig. 1.) A solute-solvent interface Γ is a closed surface that encloses all the solute atoms but no solvent molecules. The interior and exterior of Γ are the solute and solvent regions, denoted $\Omega_{\rm m}$ and $\Omega_{\rm w}$, respectively. We introduce the VISM solvation free-energy functional (34, 35):

$$G[\Gamma] = \Delta P \,\mathrm{vol}\,(\Omega_{\rm m}) + \int_\Gamma \gamma \, dS + \rho_0 \int_{\Omega_{\rm w}} U(\mathbf{r}) \, dV + G_{\rm e}[\Gamma]. \quad [2]$$

Here, ΔP is the difference of pressures across the interface Γ, γ is the solute-solvent interface surface tension, ρ_0 is the bulk solvent (i.e., water) density, and $U(\mathbf{r}) = \sum_{i=1}^N U_i(|\mathbf{r}-\mathbf{r}_i|)$ with each U_i a standard 12–6 LJ potential. We take $\gamma = \gamma_0(1-2\tau H)$, where γ_0 is the surface tension for a planar interface, τ is the curvature correction coefficient often known as the Tolman length (51), and H is the local mean curvature. The last term $G_{\rm e}[\Gamma]$ is the electrostatic part of the solvation free energy, which we will not include in this study.

Minimizing the functional Eq. [2] among all the solute-solvent interfaces Γ determines a stable, equilibrium, solute-solvent interface, called a VISM surface, and the corresponding solvation free energy. A VISM surface is termed dry, representing a dry hydration state, if it loosely wraps up all the solute atoms with enough space for a few solvent molecules, or wet, representing a wet hydration state, if it tightly wraps up all the solute atoms without extra space for a solvent molecule.

Implementation by the level-set method. Beginning with an initially guessed solute-solvent interface, our level-set method evolves the interface step by step in the steepest descent direction until a VISM surface is reached. Different initial surfaces may lead to different final VISM surfaces. See Supporting Information (SI) for more details of implementation.

The level-set VISM-string method for minimum energy paths (MEPs). Let us fix all the solute atomic positions and assume that Γ_0 and Γ_1 are two VISM surfaces (e.g., dry and wet surfaces). We apply the string method (43, 44) to find a MEP that connects Γ_0 and Γ_1. A string or path here is a family of solute-solvent interfaces $\{\Gamma_\alpha\}_{\alpha\in[0,1]}$ that connects the two states Γ_0 and Γ_1. Such a string is a MEP, if it is orthogonal to the level surfaces of the VISM free-energy functional. To find a MEP connecting Γ_0 and Γ_1, we select some initial images (i.e., points of a string), and then update them iteratively to reach a MEP. Different initial images may lead to different MEPs. Once a MEP is found, we can then find a saddle point on the MEP. Alternatively, we can fix one of the VISM surfaces, select some initial images, and allow the last image to climb up to reach a saddle point, and then find the MEP connecting the two VISM surfaces passing the saddle point. We refer to SI for more details on our implementation of the method.

Consider now our ligand-pocket system; cf. Fig. 1. For any reaction coordinate z, we label all the three hydration states 1s-dry, 2s-dry, and 2s-wet (cf. Fig. 1) as the states 0, 1, and 2, respectively. We define for each $i \in \{0,1,2\}$ the potential

$$V_i(z) = G_i(z) + U_0(z), \quad [3]$$

where $G_i(z)$ is the solvation free energy of the ith state at z (cf. Fig 2 (A)) and $U_0(z)$ is the ligand-pocket vdW interaction potential defined below Eq. [1]. We set $V_i(z) = 0$ if the ith state does not exist at z.

With the energy barriers summarized in Fig. 4, we can calculate for each z the rate $R_{ij} = R_{ij}(z)$ of the transition from one state i to another j. If a MEP from i to j passes through another state k (cf. Fig. 3), then we set $R_{ij}(z) = 0$. If there is only one MEP connecting i and j (see, e.g., $z < 4$ in Fig. 2), then $R_{ij} = R_0 e^{-B_{ij}(z)/k_{\rm B}T}$ with $B_{ij}(z)$ the energy barrier from i to j and R_0 a constant prefactor, describing the intrinsic time scale of water dynamics in the pocket. Finally, if there are two MEPs (axisymmetric and axiasymmetric) connecting i and j, we use the same formula but with B_{ij} an effective barrier. For instance, consider i and j the states (I) and (IV) in Fig. 3, respectively. The two transition states are II and III, respectively. We set $B_{ij}(z) = B_{\rm I,IV}(z) = p(G_{\rm II} - G_{\rm I}) + (1-p)(G_{\rm III} - G_{\rm I})$, where $p = e^{-(G_{\rm II}-G_{\rm I})/k_{\rm B}T}/(e^{-(G_{\rm II}-G_{\rm I})/k_{\rm B}T} + e^{-(G_{\rm III}-G_{\rm I})/k_{\rm B}T})$ and G_A is the VISM solvation free energy at state $A \in \{\rm I, II, III\}$. To determine the prefactor R_0, we calculate the equilibrium (i.e., the large z limit) energy barriers $B_{\rm dw}$ and $B_{\rm wd}$ in the pocket dry-wet and wet-dry transitions, respectively, and equate $[R_0(e^{-B_{\rm dw}/k_{\rm B}T} + e^{-B_{\rm wd}/k_{\rm B}T})]^{-1}$ with the time scale for the relaxation of water fluctuation of 10 ps as predicted by explicit-water MD simulations (18). See SI for discussions on the sensitivity of the results on R_0.

Continuous-time Markov chain (CTMC) Brownian dynamics (BD) simulations and the mean first-passage time (MFPT). To include explicitly the dry-wet fluctuations, we introduce a position-dependent, multi-state, random variable $\eta = \eta(z)$: $\eta(z) = i$ ($i \in \{0,1,2\}$) if the system is in the ith hydration state when the ligand is located at z, with the transition rates $R_{ij}(z)$ given above. We define the potential $V_{\rm fluc}(\eta, z) = V_i(z)$ (cf. Eq. [3]) if $\eta(z) = i$. (52). The random position $z = z(t) = z_t$ of the ligand is now determined by our CTMC BD simulations in which we solve the stochastic differential equation

$$dz_t = \left[-\frac{D(z_t)}{k_{\rm B}T} \frac{\partial V_{\rm fluc}(\eta(z_t), z_t)}{\partial z} + D'(z_t) \right] dt + \sqrt{2D(z_t)}\, d\xi_t.$$

Here, the partial derivative of $V_{\rm fluc}$ is with respect to its second variable, $D(z)$ is an effective diffusion coefficient that smoothly interpolates the diffusion coefficients $D_{\rm in}$ and $D_{\rm out}$ inside and outside the pocket, respectively, and ξ_t is the standard Brownian motion. Solutions to this equation are constrained by $z_t \in [z_{\rm L}, z_{\rm R}]$ for some $z_{\rm L}$ and $z_{\rm R}$. For the simulation of a binding process, we reset the value of z_t to be $2z_{\rm R} - z_t$ if $z_t \geq z_{\rm R}$, and we stop the simulation if $z_t \leq z_{\rm L}$. For the simulation of an unbinding process, we reset the value of $z(t)$ to be $z_{\rm L}$ if $z_t \leq z_{\rm L}$, and we stop the simulation if $z_t \geq z_{\rm R}$. The distribution of $\eta(z_0)$ for an initial ligand position z_0 is set based on the equilibrium probabilities $e^{-G_i/k_{\rm B}T}/\sum_{j=0}^2 e^{-G_j/k_{\rm B}T}$ ($i = 0, 1, 2$), where G_i is the solvation free energy of the ith hydration state at z_0.

We run our CTMC BD simulation for the ligand starting at a position $z_0 = z_{\rm init}$, and record the time at which the ligand reaches $z_{\rm L}$ (or $z_{\rm R}$) for the first time for a binding (or unbinding) simulation. We run simulations for 3,000 times and average these times to obtain the corresponding MFPTs.

Fokker–Planck equations (FPE) and the MFPT. The probability densities $P_i = P_i(z,t)$ for the ligand at location z at time t with the system in the ith hydration state are determined by the generalized FPEs (25, 52):

$$\frac{\partial P_i}{\partial t} = \frac{\partial}{\partial z}\left\{D(z)\left[\frac{\partial P_i}{\partial z} + \frac{1}{k_{\rm B}T}V_i'(z)P_i\right]\right\} + \sum_{0\le j\le 2, j\neq i} R_{ji}(z)P_j - \left(\sum_{0\le j\le 2, j\neq i} R_{ij}(z)\right)P_i$$

for $i = 0, 1, 2$, where V_i is defined in Eq. [3]. These equations are solved for $z_{\rm L} < z < z_{\rm R}$, with the boundary conditions $P_i(z_{\rm L}, t) = 0$ and $\partial_z P_i(z_{\rm R}, t) = 0$ for binding, and $\partial_z P_i(z_{\rm L}, t) + (1/k_{\rm B}T)V_i'(z_L)P_i(z_{\rm L}, t) = 0$ and $P_i(z_{\rm R}, t) = 0$ for unbinding, respectively. The initial conditions are $P_i(z, 0) = \delta(z - z_{\rm init})$ if the ligand is initially at $z_{\rm init}$. We obtain the MFPT as the double integral of $\sum_{i=0}^{2} P_i(z,t)$ over $(z,t) \in [z_{\rm L}, z_{\rm R}] \times [0, \infty)$.

Parameters. We set the temperature $T = 298$ K, bulk water density $\rho_0 = 0.033$ Å^{-3}, the solute-water surface tension constant $\gamma_0 = 0.143\ k_{\rm B}T/$Å^2 ($k_{\rm B}$ is the Boltzmann constant), and the Tolman length $\tau = 0.8$ Å. We set $\Delta P{\rm vol}\,(\Omega_{\rm m}) = 0$ as it is relatively very small. The LJ parameters for the wall particles, ligand, and water are $\varepsilon_{\rm wall} = 0.000967\ k_{\rm B}T$ and $\sigma_{\rm wall} = 4.152$ Å, $\varepsilon_{\rm ligand} = 0.5\ k_{\rm B}T$ and $\sigma_{\rm ligand} = 3.73$ Å, and $\varepsilon_{\rm water} = 0.26\ k_{\rm B}T$ and $\sigma_{\rm water} = 3.154$ Å, respectively. The interaction LJ parameters are determined by the Lorentz–Berthelot mixing rules. The prefactor $R_0 = 0.13$ ps^{-1}. The diffusion constants are $D_{\rm out} = 0.26$ Å^2/ps (18), and $D_{\rm in} = 1$ Å^2/ps. The cut-off position distinguishing the inside and outside of the pocket is $z_{\rm c} = -0.5$ Å. BD simulations and FPE calculations are done for $z_{\rm L} \le z \le z_{\rm R}$ with $z_{\rm L} = -4$ Å and $z_{\rm R} = 15.5$ Å.

ACKNOWLEDGMENTS. SZ was supported in part by NSF of Jiangsu Province, China, through grant BK20160302, NSFC through grant NSFC 21773165 and NSFC 11601361, and Soochow University through a start-up grant Q410700415. RGW and JD thank the DFG for financial support. JD also acknowledges funding from the ERC within the Consolidator Grant with Project No. 646659–NANOREACTOR. Work in the McCammon group is supported in part by NIH, NBCR, and SDSC. LTC and BL were supported in part by the NSF through the grant DMS-1620487. SZ thanks Dr. Yanan Zhang for helpful discussions on the string method.

1. Babine RE, Bender SL (1997) Molecular recognition of protein-ligand complexes: Applications to drug design. *Chem. Rev.* 97(5):1359–1472.
2. Pan AC, Borhani DW, Dror RO, Shaw DE (2013) Molecular determinants of drug–receptor binding kinetics. *Drug Discov. Today* 18:668–673.
3. Levy Y, Onuchic J (2006) Water mediation in protein folding and molecular recognition. *Ann. Rev. Biophys. Biomol. Struct.* 35:389–415.
4. Baron R, McCammon JA (2013) Molecular recognition and ligand association. *Annu. Rev. Phys. Chem.* 64:151–175.
5. Hummer G, Garde S, Garcia S, Pratt LR (2000) New perspectives on hydrophobic effects. *Chem. Phys.* 258:349–370.
6. Chandler D (2005) Interfaces and the driving force of hydrophobic assembly. *Nature* 437:640–647.
7. Berne BJ, Weeks JD, Zhou R (2009) Dewetting and hydrophobic interaction in physical and biological systems. *Annu. Rev. Phys. Chem.* 60:85–103.
8. Ben-Amotz D (2016) Water-mediated hydrophobic interactions. *Annu. Rev. Phys. Chem.* 67:617–638.
9. Gilson MK, Given JA, Bush BL, McCammon JA (1997) The statistical-thermodynamic basis for computation of binding affinities: A critical review. *Biophys. J.* 72:1047–1069.
10. Gilson MK, Zhou HX (2007) Calculation of protein-ligand binding affinities. *Annu. Rev. Biophys. Biomol. Struct.* 36:21–42.
11. Bernetti M, Cavalliab A, Mollica L (2017) Protein–ligand (un)binding kinetics as a new paradigm for drug discovery at the crossroad between experiments and modelling. *Med. Chem. Commun.* 8:534–550.
12. Schuetz DA, et al. (2017) Kinetics for drug discovery: An industry-driven effort to target drug residence time. *Drug Discov. Today* 22:896–911.
13. Young T, Abel R, Kim B, Berne B, Friesner R (2007) Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc. Natl. Acad. Sci. USA* 104(3):808–813.
14. Qvist J, Davidovic M, Hamelberg D, Halle B (2008) A dry ligand-binding cavity in a solvated protein. *Proc. Natl. Acad. Sci. USA* 105(17):6296–6301.
15. Wang L, Berne BJ, Friesner RA (2011) Ligand binding to protein-binding pockets with wet and dry regions. *Proc. Natl. Acad. Sci. USA* 108:1326–1330.
16. Beuming T, et al. (2012) Thermodynamic analysis of water molecules at the surface of proteins and applications to binding site prediction and characterization. *Proteins* 80:871–883.
17. Setny P, et al. (2009) Dewetting-controlled binding of ligands to hydrophobic pockets. *Phys. Rev. Lett.* 103:187801.
18. Setny P, Baron R, Kekenes-Huskey P, McCammon JA, Dzubiella J (2013) Solvent fluctuations in hydrophobic cavity-ligand binding kinetics. *Proc. Natl. Acad. Sci., USA* 110:1197–1202.
19. Huang X, Margulis CJ, Berne BJ (2003) Dewetting-induced collapse of hydrophobic particles. *Proc. Natl. Acad. Sci., USA* 100(21):11953–11958.
20. Liu P, Huang X, Zhou R, Berne BJ (2005) Observation of dewetting transition in the collapse of the melittin tetramer. *Nature* 437:159–162.
21. Baron R, Setny P, McCammon JA (2010) Water in cavity-ligand recognition. *J. Am. Chem. Soc.* 132:12091–12097.
22. Weiß RG, Setny P, Dzubiella J (2017) Principles for tuning hydrophobic ligand-receptor binding kinetics. *J. Chem. Theory Comput.* 13(6):3012–3019.
23. Morrone JA, Li J, Berne BJ (2012) Interplay between hydrodynamics and the free energy surface in the assembly of nanoscale hydrophobes. *J. Phys. Chem. B* 116(1):378–389.
24. Li J, Morrone JA, Berne BJ (2012) Are hydrodynamic interactions important in the kinetics of hydrophobic collapse? *J. Phys. Chem. B* 116:378–389.
25. Mondal J, Morrone J, Berne BJ (2013) How hydrophobic drying forces impact the kinetics of molecular recognition. *Proc. Natl. Acad. Sci., USA* 110(33):13277–13282.
26. Tiwary P, Mondal J, Morrone JA, Berne BJ (2015) Role of water and steric constraints in the kinetics of cavity-ligand unbinding. *Proc. Natl. Acad. Sci., USA* 112(39):12015–12019.
27. Weiß RG, Setny P, Dzubiella J (2016) Solvent fluctuations induce non-Markovian kinetics in hydrophobic pocket-ligand binding. *J. Phys. Chem. B* 120(33):8127–8136.
28. Dror RO, et al. (2011) Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl Acad. Sci. USA* 108(32):13118–13123.
29. Plattner N, Noé F (2015) Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models. *Nat. Commun.* 6:7653.
30. Tiwary P, Limongelli V, Salvalaglio M, Parrinello M (2015) Kinetics of protein-ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proc. Natl. Acad. Sci., USA* 112(5):E386–E391.
31. Jagger BR, Lee CT, Amaro RE (2018) Quantitative ranking of ligand binding kinetics with a multiscale milestoning simulation approach. *J. Phys. Chem. Lett.* 9(17):4941–4948.
32. Miao Y, Huang Y, Walker RC, McCammon JA, Chang C (2018) Ligand binding pathways and conformational transitions of the HIV protease. *Biochemistry* 57:1533–1541.
33. Bruce NJ, Ganotra GK, Kokh DB, Sadiq SK, Wade RC (2018) New approaches for computing ligand-receptor binding kinetics. *Curr. Opin. Struct. Biol.* 49:1–10.
34. Dzubiella J, Swanson JMJ, McCammon JA (2006) Coupling hydrophobicity, dispersion, and electrostatics in continuum solvent models. *Phys. Rev. Lett.* 96:087802.
35. Dzubiella J, Swanson JMJ, McCammon JA (2006) Coupling nonpolar and polar solvation free energies in implicit solvent models. *J. Chem. Phys.* 124:084905.
36. Cheng LT, Dzubiella J, McCammon JA, Li B (2007) Application of the level-set method to the implicit solvation of nonpolar molecules. *J. Chem. Phys.* 127:084503.
37. Wang Z, et al. (2012) Level-set variational implicit solvation with the Coulomb-field approximation. *J. Chem. Theory Comput.* 8:386–397.
38. Zhou S, Cheng LT, Dzubiella J, Li B, McCammon JA (2014) Variational implicit solvation with Poisson–Boltzmann theory. *J. Chem. Theory Comput.* 10(4):1454–1467.
39. Lum K, Chandler D, Weeks JD (1999) Hydrophobicity at small and large length scales. *J. Phys. Chem. B* 103:4570–4577.
40. Ramirez R, Borgis D (2005) Density functional theory of solvation and its relation to implicit solvent models. *J. Phys. Chem. B* 109:6754–6763.
41. Bates PW, Chen Z, Sun YH, Wei GW, Zhao S (2009) Geometric and potential driving formation and evolution of biomolecular surfaces. *J. Math. Biol.* 59:193–231.
42. Cheng LT, Li B, Wang Z (2010) Level-set minimization of potential controlled Hadwiger valuations for molecular solvation. *J. Comput. Phys.* 229:8497–8510.
43. E W, Ren W, Vanden-Eijnden E (2007) Simplified and improved string method for computing the minimum energy paths in barrier-crossing events. *J. Chem. Phys.* 126:164103.
44. Ren W, Vanden-Eijnden E (2013) A climbing string method for saddle point search. *J. Chem. Phys.* 138:134105.
45. Setny P, Geller M (2006) Water properties inside nanoscopic hydrophobic pockets. *J. Chem. Phys.* 125:144717.
46. Setny P (2007) Water properties and potential of mean force for hydrophobic interactions of methane and nanoscopic pockets studied by computer simulations. *J. Chem. Phys.* 127:054505.
47. Cheng LT, et al. (2009) Interfaces and hydrophobic interactions in receptor-ligand systems: A level-set variational implicit solvent approach. *J. Chem. Phys.* 131:144102.
48. Maragliano L, Fischer A, Vanden-Eijnden E, Ciccotti G (2006) String method in collective variables: Minimum free energy paths and isocommittor surfaces. *J. Chem. Phys.* 125:024106.
49. Sittel F, Stock G (2018) Perspective: Identification of collective variables and metastable states of protein dynamics. *J. Chem. Phys.* 149:150901.
50. Cheng LT, et al. (2009) Coupling the level-set method with molecular mechanics for variational implicit solvation of nonpolar molecules. *J. Chem. Theory Comput.* 5:257–266.
51. Tolman RC (1949) The effect of droplet size on surface tension. *J. Chem. Phys.* 17:333–337.
52. Doering CR, Gadoua JC (1992) Resonant activation over a fluctuating barrier. *Phys. Rev. Lett.* 69(16):2318–2321.

Supplementary Information for

Variational Implicit-Solvent Predictions of the Dry-Wet Transition Pathways for Ligand-Receptor Binding and Unbinding Kinetics

Shenggao Zhou, R. Gregor Weiß, Li-Tien Cheng, Joachim Dzubiella, J. Andrew McCammon, and Bo Li

J. Andrew McCammon (Email: jmccammon@ucsd.edu) and Bo Li (Email: bli@math.ucsd.edu)

This PDF file includes:

Supporting Information Text

Table of contents.

Abbreviations. BD: Brownian dynamics. CTMC: continuous-time Markov chain. FPE: Fokker–Planck equation. LJ: Lennard-Jones. MD: molecular dynamics. MEP: minimum energy path. MFPT: mean first-passage time. PMF: potential of mean force. vdW: van der Waals. VISM: variational implicit-solvent model.

1. The Level-Set Method for Minimizing the VISM Solvation Free-Energy Functional

We consider the solvation of solute molecules with all the solute atoms located at $\mathbf{r}_1, \ldots, \mathbf{r}_N$ in an aqueous solvent. A solute-solvent interface Γ is a closed surface that encloses all the solute atoms but no solvent molecules. The interior and exterior of such a surface Γ, denoted by $\Omega_{\rm m}$ and $\Omega_{\rm w}$, are termed the solute and solvent regions, respectively. In the variational implicit-solvent model (VISM), we minimize the solvation free-energy functional (cf. Eq. [2] in the main text) (1, 2)

$$G[\Gamma] = \Delta P \,\mathrm{vol}\,(\Omega_{\rm m}) + \gamma_0 \int_\Gamma (1 - 2\tau H)\, dS + \rho_0 \sum_{i=1}^{N} \int_{\Omega_{\rm w}} U_i(|\mathbf{r} - \mathbf{r}_i|)\, dV + G_{\rm e}[\Gamma] \qquad [1]$$

among all the solute-solvent interfaces Γ. The parameters ΔP, γ_0, τ, and ρ_0 are the difference of pressures across Γ, the surface tension constant for a planar solute-solvent interface, the curvature correction coefficient (i.e., the Tolman length), and the bulk solvent density, respectively. In Eq. [1], H is the local mean curvature and each U_i is a 12-6 Lennard-Jones (LJ) potential with parameters σ_i and ε_i. We shall set the electrostatic part $G_{\rm e}[\Gamma] = 0$ in this study. But we will make a remark at the end of this section on the full VISM with the electrostatics. We call a solute-solvent interface a VISM surface if it minimizes (locally) the VISM functional Eq. [1], i.e., if it is a stable equilibrium. A VISM surface is dry, representing a dry hydration state, if it loosely wraps up all the solute atoms with enough space for a few solvent molecules, or wet, representing a wet hydration state, if it tightly wraps up all the solute atoms without extra space for a solvent molecule.

We have designed and implemented a robust level-set method to numerically minimize the VISM solvation free-energy functional Eq. [1] in the three-dimensional setting (3–8). Beginning with an initial solute-solvent interface that may have a large value of solvation free energy, our level-set method moves the interface in the direction of steepest descent of the VISM solvation free energy step by step until a VISM surface is reached. The (normal component of the) boundary force that moves the interface is given by the negative first variation, $F_n = -\delta_\Gamma G[\Gamma]$, of the VISM solvation free-energy functional Eq. [1] (with $G_{\rm e}[\Gamma] = 0$) (3, 7):

$$F_n(\mathbf{r}) = -\Delta P - 2\gamma_0[H(\mathbf{r}) - \tau K(\mathbf{r})] + \rho_0 \sum_{i=1}^{N} U_i(|\mathbf{r} - \mathbf{r}_i|) \qquad \forall \mathbf{r} \in \Gamma, \qquad [2]$$

where $K(\mathbf{r})$ is the Gaussian curvature at $\mathbf{r}$. As our level-set method is an optimization method of the steepest descent type, different initial interfaces are relaxed to different VISM surfaces, often representing different hydrations states. We often use the following two types of initial interfaces: a tight wrap that is a surface of the union of van der Waals (vdW) spheres centered at solute atoms with reduced radii; and a loose wrap that is a large surface loosely enclosing all the solute atoms.

To apply the level-set method (9–11) to minimizing the functional Eq. [1], we represent a solute-solvent interface Γ as the zero level set (i.e., level surface) of a function $\phi = \phi(\mathbf{r})$ (called a level-set function), i.e., $\Gamma = \{\mathbf{r} : \phi(\mathbf{r}) = 0\}$. We keep a level-set function to be negative and positive inside and outside the interface Γ, respectively. The unit normal $\mathbf{n}$ pointing from the solute to solvent region, the mean curvature H, and the Gaussian curvature K at a point

$\mathbf{r}$ on the interface can be readily expressed as $\mathbf{n} = \nabla\phi/|\nabla\phi|$, $H = (1/2)\nabla \cdot \mathbf{n}$, and $K = \mathbf{n} \cdot \text{adj}\,(\nabla^2\phi)\mathbf{n}$, respectively. Here, $\nabla^2\phi$ is the Hessian matrix of the function ϕ with entries being the second order partial derivatives $\partial^2_{ij}\phi$ of the level-set function ϕ, and adj $(\nabla^2\phi)$ is the adjoint matrix of the Hessian $\nabla^2\phi$. The motion of the interface $\Gamma = \Gamma(t)$, where t denotes the relaxation time, is then tracked by locating the level set of the corresponding level-set function $\phi = \phi(\mathbf{r}, t)$ that solves the so-called level-set equation

$$\partial_t\phi + F_n|\nabla\phi| = 0, \qquad [3]$$

where the boundary force F_n, given in Eq. [2], is extended to the entire computational box or a band centered around the interface Γ. We start from an initial level-set function ϕ_0 at $t = 0$ and solve the equation by iteration in time until a steady-state solution is reached. To avoid the gradient $\nabla\phi$ being too small which can lead to numerical instability in locating the interface, we reinitialize the level-set function ϕ every few time steps in iteration. The reinitialization is done by solving

$$\partial_t\phi + \text{sign}\,(\phi_0)(|\nabla\phi| - 1) = 0, \qquad [4]$$

where ϕ_0 is the level-set function before reinitialization, sign (ϕ_0) is the sign of ϕ_0, and the time t can be different from that in the original level-set equation [3]. See (3, 5–7) for more details.

We remark that the electrostatic part of the solvation free energy, $G_e[\Gamma]$, can be included as the Coulomb-field approximation (CFA) (12, 13) or the dielectric-boundary Poisson–Boltzmann (PB) electrostatic free energy (14–16). The CFA does not include the ionic effect but is efficient as it requires no numerical solution of partial differential equations. The PB free energy is determined by the electrostatic potential that is the unique solution to a boundary-value problem of the dielectric-boundary PB equation. Explicit formula of the (normal component of the) dielectric-boundary force, defined as the negative variation $-\delta_\Gamma G_e[\Gamma]$, has been obtained (17–19). We have implemented both CFA and PB electrostatics; cf. (6–8).

2. The Level-Set Implementation of the VISM-String Method

Let us fix all the solute atoms $\mathbf{r}_i$ $(i = 1, \ldots, N)$ and consider two different VISM surfaces Γ_0 and Γ_1, represented by two level-set functions ϕ_0 and ϕ_1, respectively. We use the string method (20–22) to find minimum energy paths (MEPs) that connect these two states. A string or path here is a family of solute-solvent interfaces $\{\Gamma_\alpha\}_{\alpha\in[0,1]}$, or their corresponding level-set functions $\{\phi_\alpha\}_{\alpha\in[0,1]}$, that connect the two states Γ_0 and Γ_1, or their level-set functions ϕ_0 and ϕ_1. A MEP here is a string that is orthogonal to the level surfaces of the VISM solvation free-energy functional. In the level-set formulation, a MEP can be obtained by solving for a steady-state solution of the equation for the level-set function $\phi_\alpha = \phi_\alpha(x, t)$

$$\partial_t\phi_\alpha = -F_n(\phi_\alpha)|\nabla\phi_\alpha| + \lambda_\alpha \frac{\partial_\alpha\phi_\alpha}{\|\partial_\alpha\phi_\alpha\|} \qquad \text{for each } \alpha \in (0, 1),$$

together with a given initial string $\{\phi^{(0)}_\alpha\}_{\alpha\in[0,1]}$ that connects ϕ_0 and ϕ_1, Here, the normal component of the boundary force $F_n(\phi_\alpha) = -\delta_\Gamma G[\Gamma_\alpha]$ (with ϕ_α being a zero level-set of Γ_α) is given in Eq. [2], $\partial_\alpha\phi_\alpha/\|\partial_\alpha\phi_\alpha\|$ is the unit vector tangential to the string, the constant λ_α is a Lagrange multiplier for enforcing particular parameterization (e.g., equal arc-length or energy weighted arc-length parameterization) of the string, and $\|\cdot\|$ denotes the $L^2(\Omega)$-norm.

Let us focus now on the model ligand-pocket system (cf. Fig. 1 in the main text) with a fixed reaction coordinate z. We implement a simplified version of the string method (21) to numerically find a MEP connecting two hydration states Γ_0 and Γ_1, with their level-set functions ϕ_0 and ϕ_1, respectively. To do so, we select some integer $M \geq 2$ and discretize the parameter $\alpha \in [0, 1]$ by $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_M < \alpha_{M+1} = 1$, and consider the corresponding level-set functions ϕ_{α_j} $(j = 0, 1, \ldots, M + 1)$ that represent some solute-solvent interfaces. Each ϕ_{α_j} is called an image. These images are discrete points of a string or path connecting ϕ_0 and ϕ_1. They are updated iteratively to reach a stable steady state, representing a MEP. We set the initial images for the iteration to be

$$\phi^{(0)}_{\alpha_j} = \phi_0 + \alpha_j(\phi_1 - \phi_0) \qquad (j = 1, \ldots, M). \qquad [5]$$

Each iteration is a two-step process: relaxation and redistribution. Suppose we know all the interior images $\phi^{(k)}_{\alpha_j}$ $(j = 1, \ldots, M)$ after the kth iteration. In the first step, we solve the level-set equation [3] for each j $(1 \leq j \leq M)$ with the initial function $\phi^{(k)}_{\alpha_j}$ but only for one time step, followed by the reinitialization (cf. Eq. [4]), and obtain a solution $\phi^*_{\alpha_j}$. These images $\phi^*_{\alpha_j}$ $(j = 1, \ldots, M)$ should make the new string "closer" to being normal to the free-energy level surfaces, but may also cluster around the two states ϕ_0 and ϕ_1, as they are local minimizers of the VISM solvation free-energy functional. In the second step, we redistribute these intermediate images by linear interpolation

to generate new and well-separated images $\phi_{\alpha_j}^{(k+1)}$. More precisely, we set $s_0 = 0$ and $s_j = s_{j-1} + \|\phi^*_{\alpha_j} - \phi^*_{\alpha_{j-1}}\|$ $(j = 1, \ldots, M+1)$, where $\phi^*_{\alpha_0} = \phi_0$ and $\phi^*_{\alpha_{M+1}} = \phi_1$. We also set $\alpha_j^* = s_j/s_M$ $(j = 0, 1, \ldots, M+1)$. For each j $(1 \le j \le M)$, we find the unique i $(1 \le i \le M+1)$ that depends on j such that $\alpha^*_{i-1} \le \alpha_j < \alpha^*_i$. We then calculate $\phi_{\alpha_j}^{(k+1)}$ by the linear interpolation

$$\phi_{\alpha_j}^{(k+1)} = \phi^*_{\alpha_{i-1}} + \frac{\alpha_j - \alpha^*_{i-1}}{\alpha^*_i - \alpha^*_{i-1}}(\phi^*_{\alpha_i} - \phi^*_{\alpha_{i-1}}). \qquad [6]$$

Once the iteration converges to a MEP, we find an interior image that has the largest VISM solvation free energy among all the images, and identify it as a saddle point. Note that different initial images may lead to different MEPs; cf. Fig. 3 in the main text.

Algorithm of a Simplified String Method.

Step 1. Input all the parameters ΔP, γ_0, τ, ρ_0, and $\mathbf{r}_i$, σ_i, and ε_i for all $i = 1, \ldots, N$. Input the level-set functions ϕ_0 and ϕ_1 for the two states. Input M, the number of (interior) images in the string, the parameters α_j $(j = 0, 1, \ldots, M+1)$ for the string images, and the initial (interior) image level-set functions $\phi_j^{(0)}$ $(j = 1, \ldots, M)$; cf. Eq. [5]. Input the time step Δt. Set the iteration counter $k = 0$.

Step 2. Given the interior images $\phi_{\alpha_j}^{(k)}$ $(j = 1, \ldots, M)$. For each j $(1 \le j \le M)$, solve the level-set equation [3] using the initial solution $\phi_{\alpha_j}^{(k)}$ for one time step to obtain the image $\bar{\phi}_{\alpha_j}$. Compute the image $\phi^*_{\alpha_j}$ by solving the reinitialization equation [4] with $\bar{\phi}_{\alpha_j}$ as the initial solution.

Step 3. Compute the arc lengths $s_0 = 0$ and $s_j = s_{j-1} + \|\phi^*_{\alpha_j} - \phi^*_{\alpha_{j-1}}\|$ $(j = 1, \ldots, M+1)$ and the parameters $\alpha_j^* = s_j/s_M$ $(j = 0, 1, \ldots, M+1)$. Generate the images $\phi_{\alpha_j}^{(k+1)}$ $(j = 1, \ldots, M)$ by Eq. [6].

Step 4. Check the stopping criteria. If failed, set $k := k + 1$ and go to Step 2.

To find possible multiple MEPs connecting the two states ϕ_0 and ϕ_1, we can alternatively apply the climbing string method (23) to first find saddle points near ϕ_0. In implementation, we fix the first image ϕ_0 but allow the last image to climb uphill in the direction tangental to the string. The string converges when the last image approaches a saddle point close to the starting state ϕ_0. Usually, we use more images close to the last one to more efficiently find a saddle point. Once a saddle point is found, we then relax it to a level-set function representing a VISM surface. If this function is ϕ_1, then we can use the simplified string method described above, in which we keep the saddle point as an image during the iteration, to find an MEP that connects these two states ϕ_0 and ϕ_1, and that passes through the found saddle point. Otherwise, we start over with different initial images. Since we usually have at most three significant hydration states for each reaction coordinate, we can efficiently find multiple MEPs (if exist) connecting these states.

Algorithm of a Climbing String Method.

Step 1. Input all the parameters ΔP, γ_0, τ, ρ_0, and $\mathbf{r}_i$, σ_i, and ε_i for all $i = 1, \ldots, N$. Input a level-set function ϕ_0 for a VISM surface. Input M, with $M+2$ the number of images in the string, the parameters $\{\alpha_j\}_{j=0}^{M+1}$ for the string images with $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_{M+1} < 1$, and the initial image level-set functions $\{\phi_{\alpha_j}^{(0)}\}_{j=1}^{M+1}$. Input the time step Δt. Set the iteration counter $k = 0$.

Step 2. Given the images $\phi_{\alpha_j}^{(k)}$ $(j = 1, \ldots, M+1)$. For each j $(1 \le j \le M+1)$, solve the level-set equation [3] using the initial solution $\phi_{\alpha_j}^{(k)}$ for one time step to obtain an image $\bar{\phi}_j$. Solve the reinitialization equation [4] using the initial solution $\bar{\phi}_j$ for one time step to obtain an image ϕ_j^*.

Step 3. Update the last image

$$\phi_{\alpha_{M+1}}^{(k+1)} = \phi^*_{M+1} - 2\langle \phi^*_{M+1} - \phi^{(k)}_{\alpha_{M+1}}, \hat{\tau}_{M+1}\rangle \hat{\tau}_{M+1} \quad \text{with} \quad \hat{\tau}_{M+1} = \frac{\phi^{(k)}_{\alpha_{M+1}} - \phi^{(k)}_{\alpha_M}}{\|\phi^{(k)}_{\alpha_{M+1}} - \phi^{(k)}_{\alpha_M}\|},$$

where $\langle \cdot, \cdot \rangle$ denotes the $L^2(\Omega)$-inner product.

Step 4. Compute the arc lengths $s_0 = 0$ and $s_j = s_{j-1} + \|\phi^*_{\alpha_j} - \phi^*_{\alpha_{j-1}}\|$ $(j = 1, \ldots, M+1)$, and set $\alpha_j^* = s_j/s_M$ $(j = 0, 1, \ldots, M+1)$. Update the other images to obtain $\phi_{\alpha_j}^{(k+1)}$ $(j = 1, \ldots, M)$ by Eq. [6].

Step 5. Check the stopping criteria. If failed, set $k := k + 1$ and go to Step 2.

3. Algorithms for Brownian Dynamics Simulations of the Ligand Stochastic Motion

In the absence of the pocket dry-wet fluctuations, the random position $z = z(t)$ (also denoted z_t) can be determined by the standard Brownian dynamics (BD) simulations that solve numerically the stochastic differential equation

$$dz_t = \left[-\frac{1}{k_{\rm B}T}D(z_t)V'(z_t) + D'(z_t)\right]dt + \sqrt{2D(z_t)}\,d\xi_t, \qquad [7]$$

together with a given initial position $z(0)$, where $V(z)$ is the equilibrium potential of mean force (PMF) (defined in Eq. [1] and plotted in Fig. 2 (B), both in the main text), ξ_t is the standard Brownian motion, and a prime stands for derivative. The effective and position-dependent diffusion coefficient $D = D(z)$ is a smooth interpolation of the diffusion constants $D_{\rm in}$ and $D_{\rm out}$ for the ligand inside and outside the pocket, respectively. It is given by

$$D(z) = \frac{D_{\rm in} + D_{\rm out}}{2} - \frac{D_{\rm in} - D_{\rm out}}{2}\tanh\left[\nu\left(z - z_{\rm c}\right)\right], \qquad [8]$$

where $\nu > 0$ is a parameter that controls the width of the transition from $D_{\rm in}$ to $D_{\rm out}$ and $z_{\rm c}$ is a threshold reaction coordinate distinguishing the ligand being inside or outside the pocket. Solutions to Eq. [7] are constrained by $z(t) \in [z_{\rm L}, z_{\rm R}]$ for all t for some boundaries $z_{\rm L}$ and $z_{\rm R}$, with $z_{\rm L}$ close to the pocket and $z_{\rm R}$ far away from the pocket, respectively. For the binding simulation (i.e., the simulation of a binding process), we reset the value of $z(t)$ to be $2z_{\rm R} - z(t)$ if $z(t) \geq z_{\rm R}$, and we stop the simulation if $z(t) \leq z_{\rm L}$. For the unbinding simulation (i.e., the simulation of an unbinding process), we reset the value of $z(t)$ to be $z_{\rm L}$ if $z(t) \leq z_{\rm L}$, and we stop the simulation if $z(t) \geq z_{\rm R}$.

Algorithm for BD Simulations without the Dry-Wet Fluctuations.

Step 1. Input the diffusion constants $D_{\rm in}$ and $D_{\rm out}$, the controlling parameter ν, the threshold position $z_{\rm c}$, the total PMF $V(z)$, an initial ligand position $z_{\rm init}$, and the simulation time step δt. Set Time $= 0$, $z^{(0)} = z_{\rm init}$, and $k = 0$.

Step 2. Given a ligand position $z^{(k)}$. Calculate $z^{(k+1)}$ by

$$z^{(k+1)} - z^{(k)} = -\left[\frac{1}{k_{\rm B}T}D(z^{(k)})V'(z^{(k)}) + D'(z^{(k)})\right]\delta t + \sqrt{2D(z^{(k)})\delta t}\,\xi,$$

where ξ is a random number with the standard normal distribution.

Step 3. Set Time := Time $+\delta t$.

(a) For binding simulations: If $z^{(k+1)} \geq z_R$, set $z^{(k+1)} := 2z_R - z^{(k+1)}$; If $z^{(k+1)} \leq z_L$, then stop.

(b) For unbinding simulations: If $z^{(k+1)} \leq z_L$, set $z^{(k+1)} := z_L$; If $z^{(k+1)} \geq z_R$, then stop.

Step 4. Set $k := k + 1$ and go to Step 2.

To study the effect of dry-wet fluctuations on the kinetics of ligand-pocket binding/unbinding, let us define a position-dependent, three-state, random variable $\eta = \eta(z) \in \{0, 1, 2\}$ by $\eta(z) = 0$, 1, or 2, if the hydration state of the system at a given reaction coordinate z is 1s-dry, 2s-dry, or 2s-wet, respectively. The (discrete) probability density of $\eta(z)$ is defined by the equilibrium probabilities $P_i^{\rm eq}(z)$ $(i = 0, 1, 2)$:

$$\text{Prob}\left(\{\eta(z) = i\}\right) = P_i^{\rm eq}(z) = \frac{e^{-G[\Gamma_i(z)]/k_{\rm B}T}}{\sum_{j=0}^{2} e^{-G[\Gamma_j(z)]/k_{\rm B}T}}, \qquad i = 0, 1, 2,$$

where $G[\Gamma_i(z)]$ is the VISM solvation free energy at the ith hydration state represented by the VISM surface $\Gamma_i(z)$, and the sum runs over all the hydration states, at the given reaction coordinate z. To account for the fluctuations among the three states at each reaction coordinate, we further define a potential $V_{\rm fluc} = V_{\rm fluc}(\eta, z)$ by $V_{\rm fluc}(\eta, z) = V_i(z)$ if $\eta = i$ for $i \in \{0, 1, 2\}$, where the potential functional $V_i(z)$, defined in Eq. [3] in the main text, is the sum of the solvation free energy of the ith hydration state and the ligand-pocket vdW interaction energy at the reaction coordinate z. If at a given coordinate z, there is only one or two hydration states, then we set $V_i(z) = 0$ for the other states i.

We perform our continuous-time Markov chain (CTMC) BD simulations, i.e., numerically solve the following stochastic differential equation for the ligand position $z = z(t) = z_t$ (same as that in the CTMC BD simulations part

of section Theory and Methods in the main text):

$$\begin{cases} dz_t = \left[-\dfrac{1}{k_{\mathrm{B}}T}D(z_t)\dfrac{\partial V_{\mathrm{fluc}}(\eta(z_t),z_t)}{\partial z} + D'(z_t)\right]dt + \sqrt{2D(z_t)}\,d\xi_t, \\ \eta(z_t) \in \{0,1,2\} \text{ is a CTMC with the transition rate matrix} \\ \begin{pmatrix} -\left[R_{01}(z_t)+R_{02}(z_t)\right] & R_{01}(z_t) & R_{02}(z_t) \\ R_{10}(z_t) & -\left[R_{10}(z_t)+R_{12}(z_t)\right] & R_{12}(z_t) \\ R_{20}(z_t) & R_{21}(z_t) & -\left[R_{20}(z_t)+R_{21}(z_t)\right] \end{pmatrix}, \end{cases} \quad [9]$$

together with a given initial position $z_0 = z_{\mathrm{init}}$. Here, the partial derivative of V_{fluc} is with respect to its second variable, ξ_t is the standard Brownian motion, and the rates of transitions $R_{ij}(z)$ from the ith state to the jth state for all $i, j = 0, 1, 2$ are defined in Theory and Methods in the main text. Solutions to Eq. [9] are constrained by $z(t) \in [z_{\mathrm{L}}, z_{\mathrm{R}}]$ for all t for some boundaries z_{L} and z_{R}. Again, for a binding simulation, we reset the value of $z(t)$ to be $2z_{\mathrm{R}} - z(t)$ if $z(t) \geq z_{\mathrm{R}}$, and we stop the simulation if $z(t) \leq z_{\mathrm{L}}$. For an unbinding simulation, we reset the value of $z(t)$ to be z_{L} if $z(t) \leq z_{\mathrm{L}}$, and we stop the simulation if $z(t) \geq z_{\mathrm{R}}$.

Algorithm for CTMC BD Simulations.

Step 1. Input the diffusion constants D_{in} and D_{out}, the controlling parameter ν, the threshold position z_{c}, the potential functions $V_0(z)$, $V_1(z)$, and $V_2(z)$, an initial position z_{init}, and the simulation time step δt. Initialize the hydration state $\eta(z_{\mathrm{init}})$ according to the probabilities $P_i^{\mathrm{eq}}(z_{\mathrm{init}})$ $(i = 0, 1, 2)$. Set Time $= 0$, $z^{(0)} = z_{\mathrm{init}}$, and $k = 0$.

Step 2. Given a ligand position $z^{(k)}$. Calculate $z^{(k+1)}$ by

$$z^{(k+1)} - z^{(k)} = -\left[\frac{1}{k_{\mathrm{B}}T}D(z^{(k)})V_i'(z^{(k)}) + D'(z^{(k)})\right]\delta t + \sqrt{2D(z^{(k)})\delta t}\,\xi \qquad \text{if } \eta(z^{(k)}) = i,$$

where ξ is a random number with the standard normal distribution.

Step 3. Update the hydration state η. If $z^{(k+1)} \leq z_{\mathrm{c}}$, set $\eta = 0$; else, determine η as follows:

For $\eta = i$, if $e^{-\delta t \sum_{j\neq i} R_{ij}(z^{(k+1)})} \geq \zeta$, keep $\eta = i$,; otherwise, determine the transition from state i to state j according to the probability $R_{ij}/\sum_{k\neq i} R_{ik}$ $(i \neq j)$, where ζ is a random number uniformly distributed between 0 and 1.

Step 4. Set Time := Time $+\delta t$.

(a) For binding simulations: If $z^{(k+1)} \geq z_R$, set $z^{(k+1)} := 2z_R - z^{(k+1)}$; If $z^{(k+1)} \leq z_L$, then stop.

(b) For unbinding simulations: If $z^{(k+1)} \leq z_L$, set $z^{(k+1)} := z_L$; If $z^{(k+1)} \geq z_R$, then stop.

Step 5. Set $k := k + 1$ and go to Step 2.

4. Generalized Fokker–Planck Equations and the Mean First-Passage Time

Let us denote by $\bar{P}(z,t)$ the probability density of the ligand random position $z = z(t) \in [z_{\mathrm{L}}, z_{\mathrm{R}}]$ in the absence of pocket dry-wet fluctuations. It is determined by the following Fokker–Planck equation (FPE) that is associated with the stochastic differential equation [7]:

$$\frac{\partial \bar{P}}{\partial t} = \frac{\partial}{\partial z}\left\{D(z)\left[\frac{\partial \bar{P}}{\partial z} + \frac{1}{k_{\mathrm{B}}T}V'(z)\bar{P}\right]\right\}, \quad [10]$$

where $V = V(z)$ is the equilibrium PMF defined in Eq. [1] in the main text. The initial condition for this equation is $\bar{P}(z,0) = \bar{P}^{(0)}(z)$ for some $\bar{P}^{(0)}(z)$ and the boundary conditions are designed separately for the simulation of binding and that of unbinding:

$$\begin{aligned} &\bar{P}(z_{\mathrm{L}},t) = 0 \quad \text{and} \quad \frac{\partial \bar{P}(z_{\mathrm{R}},t)}{\partial z} = 0 && \text{for binding,} \\ &\frac{\partial \bar{P}(z_{\mathrm{L}},t)}{\partial z} + \frac{1}{k_{\mathrm{B}}T}V'(z_{\mathrm{L}})\bar{P}(z_{\mathrm{L}},t) = 0 \quad \text{and} \quad \bar{P}(z_{\mathrm{R}},t) = 0 && \text{for unbinding.} \end{aligned} \quad [11]$$

The mean first-passage time (MFPT) of binding/unbinding is given by

$$\tau_{\mathrm{MFPT}}(z_{\mathrm{init}}) = \int_0^\infty \int_{z_{\mathrm{L}}}^{z_{\mathrm{R}}} \bar{P}(z,t)\,dz\,dt,$$

where z_{init} is the initial ligand position, or equivalently, the initial value of $\bar{P}$ is given by $\bar{P}(z,0) = \delta(z - z_{\text{init}})$, the Dirac mass concentrated at z_{init}. Integrating both sides of Eq. [10] with respect to time, we arrive at

$$-\bar{P}^{\text{init}}(z, z_{\text{init}}) = \frac{d}{dz}\left\{D(z)\left[\frac{d\bar{P}^I(z)}{dz} + \frac{1}{k_{\text{B}}T}\bar{P}^I(z)V'(z)\right]\right\}, \quad [12]$$

where $\bar{P}^{\text{init}}(z, z_{\text{init}}) = \delta(z - z_{\text{init}})$ is the initial probability density, and

$$\bar{P}^I(z) = \int_0^\infty \bar{P}(z,t)dt.$$

The solution to Eq. [12] can be obtained by integrating the equation twice with the boundary conditions Eq. [11]. For instance, the unbinding MFPT of a ligand starting at z_{init} without solvent fluctuations is given by

$$\begin{aligned}\tau_{\text{MFPT}}(z_{\text{init}}) &= \int_{z_{\text{L}}}^{z_{\text{R}}} \bar{P}^I(z)dz \\ &= \int_{z_{\text{init}}}^{z_{\text{R}}} \frac{e^{\beta V(z)}}{D(z)}dz \int_{z_{\text{L}}}^{z_{\text{init}}} e^{-\beta V(z)}dz + \int_{z_{\text{init}}}^{z_{\text{R}}} e^{-\beta V(z)}\left[\int_z^{z_{\text{R}}} \frac{e^{\beta V(z')}}{D(z')}dz'\right]dz,\end{aligned}$$

where $\beta = 1/(k_{\text{B}}T)$. To get an explicit analytical solution for the MFPT of the binding, we make an assumption that $V'(z_R) = 0$, which is often true when z_R is far from the pocket. Under such an assumption, the binding MFPT of a ligand starting at z_{init} without solvent fluctuations is obtained analogously:

$$\begin{aligned}\tau_{\text{MFPT}}(z_{\text{init}}) &= \int_{z_{\text{L}}}^{z_{\text{R}}} \bar{P}^I(z)dz \\ &= \int_{z_{\text{L}}}^{z_{\text{init}}} \frac{e^{\beta V(z)}}{D(z)}dz \int_{z_{\text{init}}}^{z_{\text{R}}} e^{-\beta V(z)}dz + \int_{z_{\text{L}}}^{z_{\text{init}}} e^{-\beta V(z)}\left[\int_{z_{\text{L}}}^{z} \frac{e^{\beta V(z')}}{D(z')}dz'\right]dz.\end{aligned}$$

We now consider the MFPT with dry-wet fluctuations (or the solvent fluctuations). We solve the following system of generalized FPEs for the probability densities, $P_0(z,t)$, $P_1(z,t)$, and $P_2(z,t)$, for the probabilities of finding the ligand at location z at time t with the system being in the states of 1s-dry, 2s-dry, and 2s-wet, respectively (24):

$$\frac{\partial P_i}{\partial t} = \frac{\partial}{\partial z}\left\{D(z)\left[\frac{\partial P_i}{\partial z} + \frac{1}{k_{\text{B}}T}V_i'(z)P_i\right]\right\} + \sum_{0\le j\le 2,\, j\ne i} R_{ji}(z)P_j - \left(\sum_{0\le j\le 2,\, j\ne i} R_{ij}(z)\right)P_i \qquad \text{for } i = 0,1,2. \quad [13]$$

This is the same equation as in section Theory and Methods in the main text.

These equations correspond to the stochastic differential equation [9] for our CTMC BD simulations. They are solved with some initial values and also for $z_{\text{L}} < z < z_{\text{R}}$, with the boundary conditions

$$\begin{aligned}&P_i(z_{\text{L}},t) = 0 \quad \text{and} \quad \frac{\partial P_i(z_{\text{R}},t)}{\partial z} = 0 && \text{for binding,} \\ &\frac{\partial P_i(z_{\text{L}},t)}{\partial z} + \frac{1}{k_{\text{B}}T}V'(z_{\text{L}})P_i(z_{\text{L}},t) = 0 \quad \text{and} \quad P_i(z_{\text{R}},t) = 0 && \text{for unbinding,}\end{aligned}$$

where $i = 0,1,2$.

To calculate the MFPT for the ligand-pocket binding/unbinding starting from z_{init}, we let

$$P_i^{\text{init}}(z, z_{\text{init}}) = P_i^{\text{eq}}(z_{\text{init}})\delta(z - z_{\text{init}})$$

be the initial probability densities for P_i with $i = 0,1,2$. Integrating both sides of the Eq. [13] with respect to time, we have

$$-P_i^{\text{init}}(z, z_{\text{init}}) = \frac{d}{dz}\left\{D(z)\left[\frac{dP_i^I(z)}{dz} + \frac{1}{k_{\text{B}}T}P_i^I(z)V_i'(z)\right]\right\} + \sum_{j\ne i} R_{ji}(z)P_j^I - \left(\sum_{j\ne i} R_{ij}(z)\right)P_i^I,$$

where

$$P_i^I(z) = \int_0^\infty P_i(z,t)\,dt. \quad \text{for } i = 0,1,2.$$

With certain boundary conditions, the boundary-value problem can be solved with the finite difference method. The MFPT is then given by

$$\tau_{\mathrm{MFPT}}(z_{\mathrm{init}}) = \sum_{i=0}^{2} \int_{z_{\mathrm{L}}}^{z_{\mathrm{R}}} P_i^I(z) dz.$$

This can be calculated with numerical integration.

5. Parameters

We list the values and units of all the parameters in our computations. These are the same as those described in the main text.

Symbol	Description	Units	Value
T	Temperature	Kelvin	298
ΔP	Pressure difference (cf. Eq. [1]) [a]	bar	0
γ_0	surface tension for a planar interface (cf. Eq. [1])	$k_{\mathrm{B}}T/\text{Å}^2$	0.143
τ	Tolman length (cf. Eq. [1])	Å	0.8
ρ_0	bulk solvent (i.e., water) density (cf. Eq. [1])	Å^{-3}	0.033
σ_{water}	LJ length parameter for a solvent molecule [b]	Å	3.154
σ_{wall}	LJ length parameter for a wall particle [b]	Å	4.152
σ_{ligand}	LJ length parameter for the ligand [b]	Å	3.73
$\varepsilon_{\mathrm{water}}$	LJ energy parameter for a solvent molecule [b]	$k_{\mathrm{B}}T$	0.26
$\varepsilon_{\mathrm{wall}}$	LJ energy parameter for a wall particle [b]	$k_{\mathrm{B}}T$	9.67E-4
$\varepsilon_{\mathrm{ligand}}$	LJ energy parameter for the ligand [b]	$k_{\mathrm{B}}T$	0.5
M	Number of (interior) images of a string	No units	10
z_{c}	The coordinate of the pocket entrance	Å	-0.5
z_{L}	Smallest value of the reaction coordinate [c]	Å	-4
z_{R}	Largest value of the reaction coordinate [c]	Å	15.5
R_0	Prefactor of transition rates [d]	ps^{-1}	0.13
D_{in}	Diffusion constant inside the pocket (cf. Eq. [8]) [e]	Å^2/ps	1
D_{out}	Diffusion constant outside the pocket (cf. Eq. [8]) [f]	Å^2/ps	0.26
ν	The control parameter in D (cf. Eq. [8])	1/Å	5

Table S1. Parameters.

[a] The term $\Delta P \,\mathrm{vol}\,(\Omega_{\mathrm{m}})$ is very small compared with the other terms in Eq. [1].
[b] The values are taken from (25, 26). We use the Lorentz–Berthelot mixing rules to determine the LJ parameters for the interaction of two particles.
[c] These values can vary.
[d] R_0 is estimated from the relaxation timescale $(R_{\mathrm{dw}} + R_{\mathrm{wd}})^{-1} \approx 10$ ps of water fluctuations in the pocket when the ligand is far away (27), where $R_{\mathrm{dw}} = R_0 e^{-B_{\mathrm{dw}}/k_{\mathrm{B}}T}$ and $R_{\mathrm{wz}} = R_0 e^{-B_{\mathrm{wd}}/k_{\mathrm{B}}T}$ with B_{dw} and B_{wd} the barriers in the pocket dry-wet and wet-dry transitions when the ligand is far away; cf. section Theory and Methods in the main text.
[e] This is a trial value. See subsection B in section 6. Additional Simulation Results.
[f] The value is taken from (27).

6. Additional Simulation Results

A. Minimum Energy Paths for $z = 2$ Å and $z = 10$ Å. At the reaction coordinate $z = 2$ Å, there are two hydration states: 2s-wet and 1s-dry, and only one MEP is found to connect these two states. Fig. S1 shows this MEP, together with the solute-solvent interfacial structures of the two hydration states (marked (I) and (III), respectively) and the only transition state (marked (II)). Note that the 1s-dry has a lower solvation free energy.

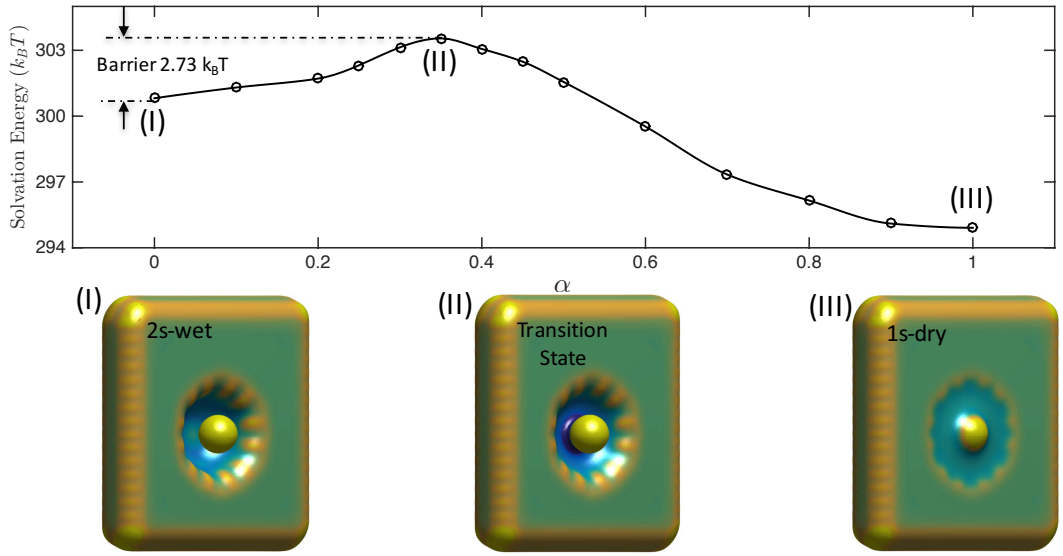

Fig. S1. The MEP connecting the only hydration states of 2s-wet (marked (I)) and 1s-dry (marked (III)) when the ligand is placed at $z = 2$ Å. The solute-solvent interfaces of these hydration states, and the transition state (marked (II)) are also shown. The energy barrier in the dewetting transition from 2s-wet to 1s-dry is $2.73\ k_{\rm B}T$.

Fig. S2 shows the MEP connecting the only hydration states 2s-wet and 2s-dry for the reaction coordinate $z = 10$ Å. The calculated activation energy barrier is about $0.68\ k_{\rm B}T$. In contrast to the dewetting energy barrier ($0.70\ k_{\rm B}T$) for $z = 6$ Å (cf. Fig. 3 in the main text), one finds that the presence of the ligand with a smaller ligand-pocket distance increases the dewetting energy barrier of the hydrophobic pocket. This is because that, when the ligand is close, part of the solvent region with the attractive solute-solvent vdW interaction is lost in such a dewetting transition. From an explicit-solvent point of view, the water molecules in the hydration shell of the methane particle hinders the evaporation of water molecules from the pocket.

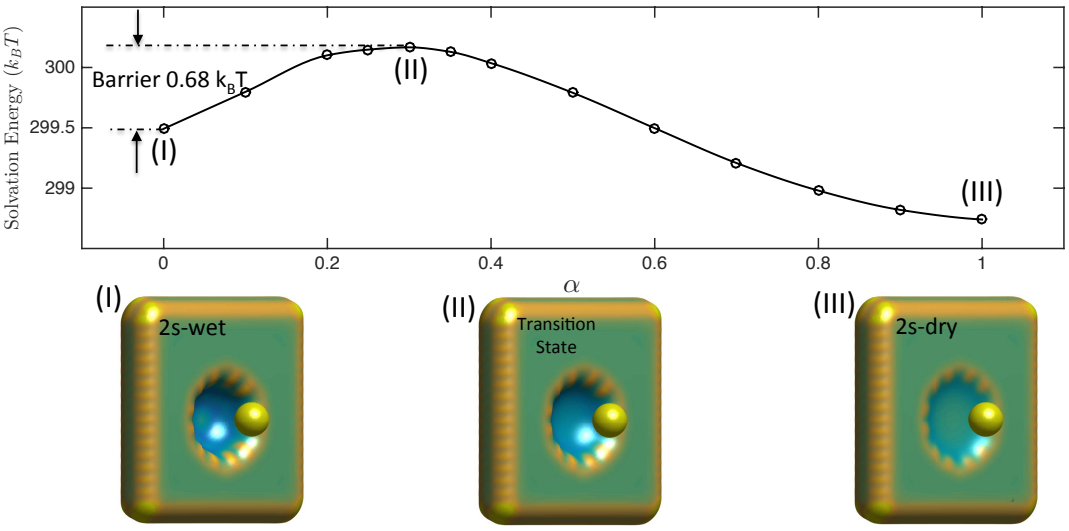

Fig. S2. The MEP connecting the hydration states 2s-wet and 2s-dry with the ligand is placed at $z = 10$ Å. The energy barrier in the dewetting transition from 2s-wet to 2s-dry is $0.68\ k_{\rm B}T$. The solute-solvent interfaces of the hydration states 2s-wet (marked (I)) and 2s-dry (marked (III)), and that of the transition state (marked (II)) are also shown.

B. Effect of $D_{\rm in}$. We choose two very different values of the diffusion constant $D_{\rm in} = 1$ Å^2/ps and $D_{\rm in} = 1,000$ Å^2/ps, and hence determine two, effective and position-dependent diffusion coefficient $D(z)$ by Eq. [8]. With these diffusion coefficients, we solve numerically Eq. [10], and Eqs [13], and then calculate the MFPT for the binding and unbinding process. Fig. S3 shows that the large difference in the diffusion constant $D_{\rm in}$ does not affect the MFPT with or without the dry-wet fluctuations.

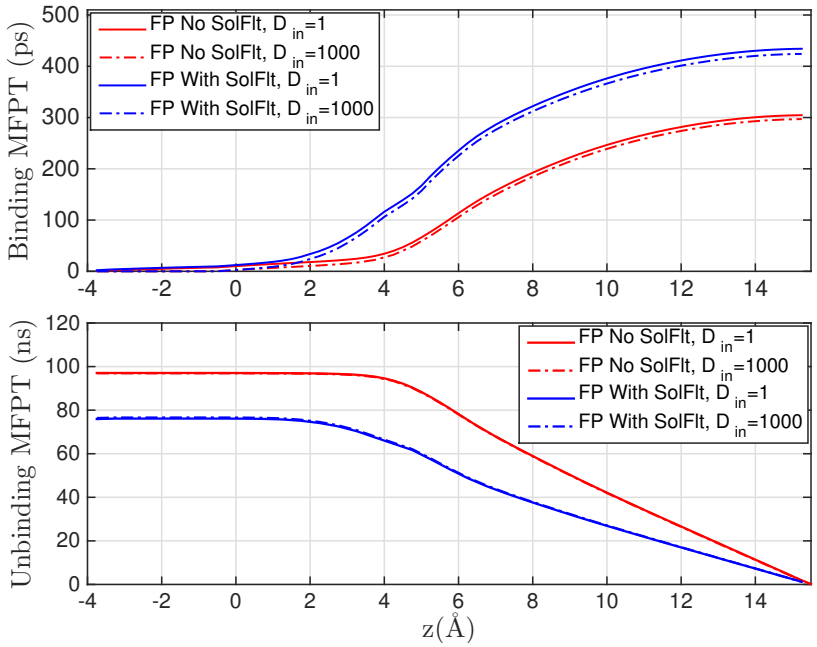

Fig. S3. The FPE calculations of the binding and unbinding MFPT of the ligand starting at z with two different values of the diffusion coefficient $D_{\rm in}$ for the ligand inside the pocket. SolFlt stands for the solvent fluctuations, i.e., the pocket dry-wet fluctuations.

C. Evolution of Probability Density of Ligand Position. To further understand the effect of solvent fluctuations, we investigate the decay rate of the probability densities $\bar{P}(z,t)$ and $P_{\rm tot}(z,t)=\sum_{i=0}^{2}P_i(z,t)$ in binding and unbinding processes Here, $\bar{P}(z,t)$ is the probability density for the ligand random position $z(t)$ in the absence of dry-wet fluctuations (cf. Eq. [10]), and each $P_i(z,t)$ ($i=0,1,$ or 2) is the probability density for the ligand random position $z(t)$ with the system being at the ith hydration state (cf. Eq. [13]). Fig. S4 displays the evolution of the probability densities normalized by the initial value at the positions $z=6$ and $z=-2$ in binding and unbinding simulations, respectively. In the binding processes, the normalized probability density decays slower when solvent fluctuations are included, because the pocket fluctuates between dry and wet states and the PMF of the wet branch is repulsive. On the contrary, the normalized probability density decays faster in unbinding processes, and hence a shorter residence time when solvent fluctuations are included. This is again due to the repulsive PMF of the wet branch. The pocket fluctuates to the wet state when the unbinding ligand approaches the entrance of the pocket.

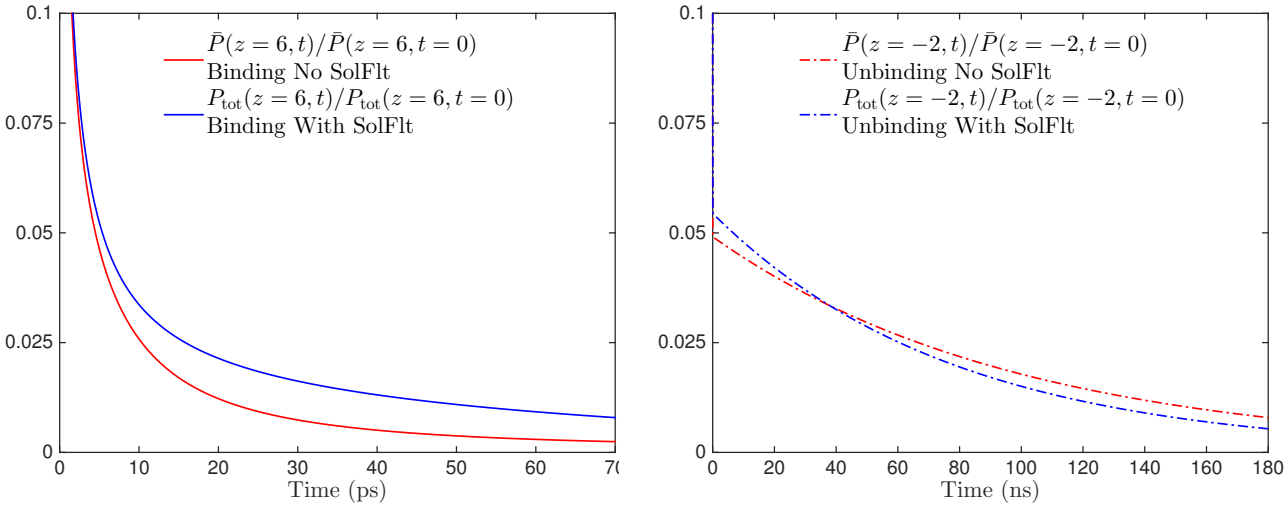

Fig. S4. Evolution of probability densities, $\bar{P}(z,t)$ (cf. Eq. [10]) and $P_{\rm tot}(z,t)=P_0(z,t)+P_1(z,t)+P_2(z,t)$ (cf. Eq. [13]) normalized by the initial values at $z=6$ (left) and $z=-2$ (right) in the binding and unbinding simulations with and without solvent fluctuations. SolFlt stands for the solvent fluctuations, i.e., the pocket dry-wet fluctuations.

D. Sensitivity of R_0. We now discuss the effect of R_0 on the binding and unbinding kinetics. Fig. S5 presents the MFPT of the binding and unbinding of ligand against the starting position $z_{\text{init}} = z$ with different values of R_0. We see that the results predicted by the CTMC BD simulations and FPE calculations agree with each other perfectly. As R_0 decreases, both binding and unbinding MFPTs increase. With a smaller R_0, the dewetting transition rate decreases and the ligand stays in the branch of 2s-wet for longer time in binding processes. This explains the longer binding MFPT with a smaller R_0. For unbinding, a smaller R_0 leads to a smaller wetting transition rate, restraining the transition starting from the 1s-dry state whose PMF is attractive. This explains the increasing unbinding MFPT with a decreasing value of R_0.

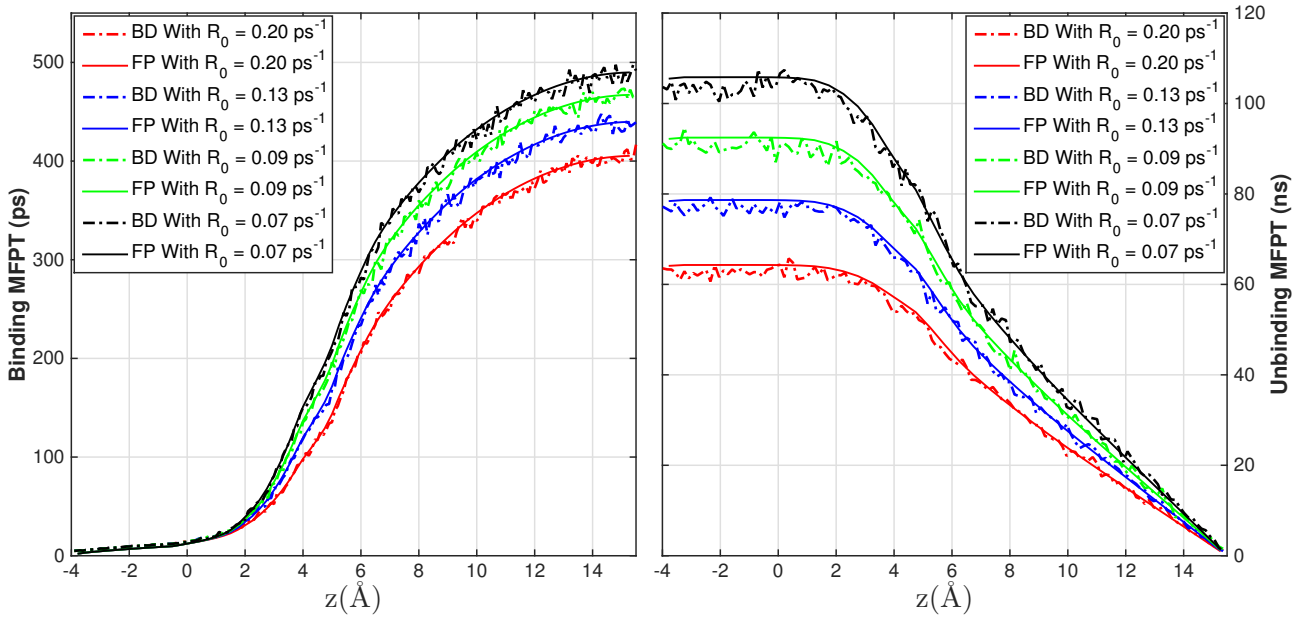

Fig. S5. The MFPT for the binding (left) and the unbinding of ligand that starts from $z_{\text{init}} = z$, predicted by the CTMC BD simulations and FPE calculations with different values of R_0.

References

1. Dzubiella J, Swanson JMJ, McCammon JA (2006) Coupling hydrophobicity, dispersion, and electrostatics in continuum solvent models. *Phys. Rev. Lett.* 96:087802.
2. Dzubiella J, Swanson JMJ, McCammon JA (2006) Coupling nonpolar and polar solvation free energies in implicit solvent models. *J. Chem. Phys.* 124:084905.
3. Cheng LT, Dzubiella J, McCammon JA, Li B (2007) Application of the level-set method to the implicit solvation of nonpolar molecules. *J. Chem. Phys.* 127:084503.
4. Cheng LT, et al. (2009) Coupling the level-set method with molecular mechanics for variational implicit solvation of nonpolar molecules. *J. Chem. Theory Comput.* 5:257–266.
5. Cheng LT, Li B, Wang Z (2010) Level-set minimization of potential controlled Hadwiger valuations for molecular solvation. *J. Comput. Phys.* 229:8497–8510.
6. Wang Z, et al. (2012) Level-set variational implicit solvation with the Coulomb-field approximation. *J. Chem. Theory Comput.* 8:386–397.
7. Zhou S, Cheng LT, Dzubiella J, Li B, McCammon JA (2014) Variational implicit solvation with Poisson–Boltzmann theory. *J. Chem. Theory Comput.* 10(4):1454–1467.
8. Zhou S, et al. (2015) LS-VISM: A software package for analysis of biomolecular solvation. *J. Comput. Chem.* 36:1047–1059.
9. Osher S, Sethian JA (1988) Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton–Jacobi formulations. *J. Comput. Phys.* 79:12–49.
10. Sethian JA (1999) *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision, and Materials Science.* (Cambridge University Press), 2nd edition.
11. Osher S, Fedkiw R (2002) *Level Set Methods and Dynamic Implicit Surfaces.* (Springer, New York).
12. Bashford D, Case DA (2000) Generalized Born models of macromolecular solvation effects. *Ann. Rev. Phys. Chem.* 51:129–152.

13. Cheng HB, Cheng LT, Li B (2011) Yukawa-field approximation of electrostatic free energy and dielectric boundary force. *Nonlinearity* 24:3215–3236.
14. Davis ME, McCammon JA (1990) Electrostatics in biomolecular structure and dynamics. *Chem. Rev.* 90:509–521.
15. Sharp KA, Honig B (1990) Electrostatic interactions in macromolecules: Theory and applications. *Annu. Rev. Biophys. Chem.* 19:301–332.
16. Che J, Dzubiella J, Li B, McCammon JA (2008) Electrostatic free energy and its variations in implicit solvent models. *J. Phys. Chem. B* 112:3058–3069.
17. Cai Q, Ye X, Wang J, Luo R (2011) Dielectric boundary forces in numerical Poisson–Boltzmann methods: Theory and numerical strategies. *Chem. Phys. Lett.* 514:368–373.
18. Li. B, Cheng XL, Zhang ZF (2011) Dielectric boundary force in molecular solvation with the Poisson–Boltzmann free energy: A shape derivative approach. *SIAM J. Applied Math.* 71:2093–2111.
19. Cai Q, Ye X, Luo R (2012) Dielectric pressure in continuum electrostatic solvation of biomolecules. *Phys. Chem. Chem. Phys.* 14:15917–15925.
20. E W, Ren W, Vanden-Eijnden E (2002) String method for the study of rare events. *Phys. Rev. B* 66:052301.
21. E W, Ren W, Vanden-Eijnden E (2007) Simplified and improved string method for computing the minimum energy paths in barrier-crossing events. *J. Chem. Phys.* 126:164103.
22. E. W, Vanden-Eijnden E (2010) Transition-path theory and path-finding algorithms for the study of rare events. *Annu. Rev. Phys. Chem.* 61:391–420.
23. Ren W, Vanden-Eijnden E (2013) A climbing string method for saddle point search. *J. Chem. Phys.* 138:134105.
24. Mondal J, Morrone J, Berne BJ (2013) How hydrophobic drying forces impact the kinetics of molecular recognition. *Proc. Natl. Acad. Sci., USA* 110(33):13277–13282.
25. Setny P, et al. (2009) Dewetting-controlled binding of ligands to hydrophobic pockets. *Phys. Rev. Lett.* 103:187801.
26. Cheng LT, et al. (2009) Interfaces and hydrophobic interactions in receptor-ligand systems: A level-set variational implicit solvent approach. *J. Chem. Phys.* 131:144102.
27. Setny P, Baron R, Kekenes-Huskey P, McCammon JA, Dzubiella J (2013) Solvent fluctuations in hydrophobic cavity-ligand binding kinetics. *Proc. Natl. Acad. Sci., USA* 110:1197–1202.