JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Gating and Intermolecular Interactions in Ligand-Protein Association: Coarse-Grained Modeling of HIV-1 Protease

Myungshim Kang,[†] Christopher Roberts,[†] Yuhui Cheng,[‡] and Chia-en A. Chang[†,*]

[†]Department of Chemistry, University of California, Riverside, California, United States

[‡]Pacific Northwest National Laboratory, Richland, Washington, United States

Supporting Information

ABSTRACT: Most biological processes are initiated or mediated by the association of ligands and proteins. This work studies multistep, ligand−protein association processes by Brownian dynamics simulations with coarse-grained models for HIV-1 protease (HIVp) and its neutral ligands. We report the average association times when the ligand concentration is 100 μM. The influence of crowding on the simulated binding time was also studied. HIVp has flexible loops that serve as a gate during the ligand binding processes. It is believed that the flaps are partially closed most of the time in its free state. To accelerate our simulations, we fixed a part of the HIVp and reparameterized our coarse-grained model, using atomistic molecular dynamics simulations, to reproduce the "gating" motions of HIVp. HIVp−ligand interactions changed the gating behavior of HIVp and helped ligands diffuse on HIVp surface to accelerate binding. The structural adjustment of the ligand toward its final stable state was the limiting step in the binding processes, which is highly system dependent. The intermolecular attraction between the ligands and crowder proteins contributes the most to the crowding effects. The results highlight broader implications in recognition pathways under more complex environment that considers molecular dynamics and conformational changes. This work brings insights into ligand−protein associations and is helpful in the design of targeted ligands.

INTRODUCTION

In biological systems, many processes such as the immune response, signal transduction, and metabolism are initiated or mediated by ligand−protein associations.[1−3] Several factors affecting ligand−protein associations include conformational changes of a molecule, intermolecular interactions, gating effects, molecular crowding, and solvent effects.[4] Notably, proper biomolecular function depends on a balance between the timing and duration of the ligand−protein interactions. The association rate constant is one of the most important properties that describe the movement of molecules in solvent. The diffusion-limited rate is the upper limit of the ligand−protein association rate,[5] and many ligand−protein systems exhibit association rates slower than the diffusion-limited association rate. Understanding ligand−protein associations provides insight into regulating protein functions and can lead to practical applications such as drug design.

If a ligand and a protein are each approximated as spheres, and their surfaces are perfectly reactive, then the diffusion-limited rate constant is given by the classical Smoluchowski equation $k_{on} = 4\pi RD$, where R is the sum of the radii and D is the relative translational diffusion constant.[6] The expression thus yields rates of 10^9-10^{10} $M^{-1}s^{-1}$ for most ligand−protein associations. However, proteins commonly have a highly anisotropic binding site, and molecules are not simple spheres, resulting in slower associations in a complex cellular environment. Although the ligand−protein association can be complicated, it may be simplified into 3 steps:[7−9] the initiation step, characterized by the free diffusion of unbound ligand and protein molecules; a second step involving several intermediates to rearrange ligand, protein, and water molecules when the two molecules meet; and the third step finalizing the binding process and resulting in the bound complex.

The free molecules and bound complex may be captured by high-resolution experimental structures, but in most cases, the second step in the process is poorly understood. During the second step, the ligand and protein optimize their interactions by changing their conformational preference or inducing new conformations.[10−12] Protein structural fluctuations also affect binding, and the binding sites of proteins can be occluded by motions of protein loops. Such flexible loops or flaps may be described as a "gate" that can open and close to influence ligand binding. An early analysis from McCammon and Northrup revealed two limiting cases: fast and slow gating.[13−15] In fast gating, the gate opens and closes much faster than the rate of escape of the ligand from the protein proximity into the solvent. The protein can thus be viewed as "always-open". In slow gating, the gate opens and closes much slower than the rate of escape of the ligand. Therefore, the successful binding rate is that of the always-open protein multiplied by the opening fraction of the gate. As compared with the second step, the initiation step is relatively well studied in the dilute environment, so theories and analytical models are available for diffusion-controlled ligand−protein associations.[4] However, cells are not dilute environments, and approximately 20%−40% of the cyctoplasmic volume is occupied by biomolecules or small compounds.[16] Simple analytical models may not describe the large macromolecular crowding effects.

Received: July 13, 2011
Published: August 26, 2011

To study the molecular encounter process, we employed the model system of human immunodeficiency virus type 1 protease (HIVp). HIVp has been one of the major targets for AIDS treatments because it is essential in the viral replication cycle. The enzyme cleaves the viral polyproteins at the active site to make the replicated virus mature and become infectious.[17,18] HIVp has flexible flaps over the binding site that open and close to "gate" ligand binding.[19] Some HIVp inhibitors have similar binding affinities, i.e., the dissociation constant K_d but quite different kinetic behaviors in the association and dissociation rate constants k_{on} and k_{off}.[20,21] Some mutations may be more sensitive to the kinetics, and elucidation of the HIVp−ligand association may aid in the drug discovery process.

In this work, we simulated ligand−protein association times using a coarse-grained (CG) model and the Brownian dynamics (BD) algorithm to connect the initial and final steps of binding processes. Each residue of HIVp is represented by one bead and the ligands are modeled by 5 or 6 beads. To accelerate our simulations, we fixed a part of HIVp and reparameterized our CG model by atomistic molecular dynamics simulations to resemble the motions of a fully flexible HIVp. To study the environmental crowding effects, we coarse-grained glutathione S-transferase P (GSTP1) as our crowder protein and as in HIVp−ligand interactions, our ligand had both attractive and repulsive interactions with the crowders. We found that the ligand−protein interactions alter the flap dynamics from slow gating to fast one. The structural adjustment of the ligand toward its final bound state was the limiting step in the binding process, which was highly system dependent. The intermolecular attractions between the ligand and crowders contributed the most to the crowder effects, as well.

METHODS

Protein and Ligand Structure. The HIVp structural coordinates were obtained from the protein data bank (PDB code 1HHP).[22] One of the most abundant proteins in the human cell, GSTP1, (PDB code 3DGQ) was selected as the crowding protein.[23] The structures of the ligands XK263, ritonavir, and saquinavir were also obtained from the protein databank (PDB codes 1HVR, 2B60, and 3EKQ, respectively).[24−26]

Coarse-Grained Model of Proteins. Various CG approaches have been developed to overcome the limitation of atomistic molecular dynamics simulations in terms of time scales as well as the system sizes.[27,28] Here, we used CG models to represent HIVp, the crowding protein, and the ligands. In the CG model, a single interaction center, placed on the C_α position, was used to represent each amino acid.[29−31] For the five charged amino acids, appropriate formal charges were assigned. The detailed procedure for the coarse graining from a collection of pdb files and the all-atom representations to the one-bead representation was previously described.[32]

The coarse-grained potential energy function is defined as follows:

$$U = U_{bond} + U_{angle} + U_{dihe} + U_{elec} + U_{vdw}^{intra} + U_{vdw}^{inter} \quad (1)$$

When one of a pair of beads was fixed, the force between the pair was applied only on the flexible bead to avoid a possible collision. Although no detailed solvent model was used, a distance-dependent dielectric constant ($\varepsilon = 4r_{ij}$) was used for compensation to avoid unrealistic Coulombic interactions.

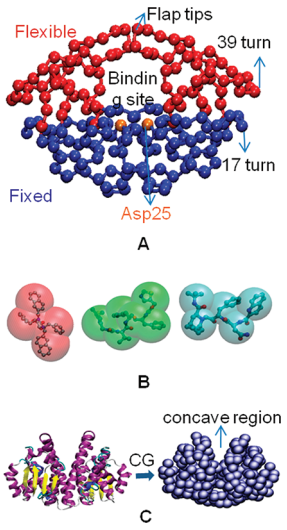

Figure 1. Coarse-grained models of HIV protease (HIVp), ligands, and GSTP1. (A) Model of HIVp. The flexible part and fixed part are in red and blue, respectively. The binding site has 2 aspartic acids (orange). (B) Models of ligands. XK263 (red, left), saquinavir (green, middle), and ritonavir (blue, right). (C) Cartoon representation (left) and model of glutathione S-transferase P (GSTP1) (right). GSTP1 has a concave region on the surface.

CG models for the ligands were also generated. The ligands were manually parametrized with interaction centers on the various functional groups in each structure (Figure 1). The net charge of each ligand is zero, but partial charges were applied to each bead (for details, see the Supporting Information).

For further acceleration of the simulation, the 108 beads distant from the highly mobile flaps of HIVp were fixed in their position. However, after fixation, the flaps did not open as widely as the free protein, which suggests possible correlations between the fixed and flexible parts. In fact, previous BD simulations suggested a correlation between the 17 and 39 turns of HIVp.[33] To properly reproduce the intramolecular motions of the flaps, we introduced modifications in angles and dihedrals. From previous atomistic molecular dynamics (MD) simulations,[34] the angles and the dihedrals in the open conformations were compared with those in the closed as well as transition conformations. The angles and dihedrals with significant differences were selected and their parameters were modified to fit the distributions from the atomistic MD simulation, based on the Boltzmann inversion[31−33,35] (see Supporting Information for further details).

Brownian Dynamics. Simulations involved the use of the modified UHBD software package as previously described.[36,37] The maximum simulation time for each run, without crowding, was 30 μs because more than 99.5% of runs bind within this time. With crowding, the maximum simulation time was set to 100 μs for a similar success rate of binding. The time step was 50 fs for all runs and the trajectory was saved every 1 ns without crowding and 10 ns with crowding. The viscosity of water η was set to 1 cP, which corresponds to a water temperature of 293 K. The total system is spherical in nature, with a radius of 160 Å. This represents the volume necessary to obtain a single ligand concentration of 100 μM, which corresponds to experimental concentrations used for comparison.

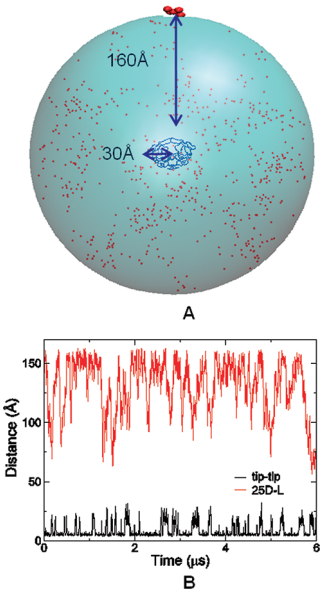

Figure 2. Diffusion association model. (A) HIVp, with a radius of about 30 Å, is centered in the spherical system of radius 160 Å. A ligand starts to diffuse in a random position on the surface of the sphere. Red dots represent the trajectory of an example ligand. (B) The tip to tip distances and distances between a ligand and the active site by simulation time.

Each simulation began with a single ligand placed in a random position on the surface of the bounding sphere, 160 Å from the point of origin to a central bead of the ligand (Figure 2). The ligand follows a Brownian trajectory over the time step given. If the trajectory were such that the ligand escaped the bounding sphere, a periodic boundary condition-like conversion was used to place the ligand at a diagonally opposite direction of the 160 Å sphere, with an additional pull of 2% total distance toward the origin to ensure that the ligand was inside the sphere. This condition ensures that the ligand concentration remains constant. Additionally, a binding event occurred if 2 conditions were satisfied: the ligand diffuses within 11 Å of the active site, and the protease conformation is in the closed position. These criteria were defined by the distance between the ligand and Asp25 and the distance between the tips of each flap of the protease, respectively.

Calculation of Translational Diffusion Coefficients. The Einstein formula was used to calculate the translational diffusion coefficients for the ligand molecules:

$$D_{trans} = \frac{\langle \delta r^2 \rangle}{6\delta t} \qquad (2)$$

where δr is the distance traveled in 3 dimension space during the time interval δt. The equation describes the ensemble average of the center of mass Cartesian trajectory over the time interval δt. The time interval selected for the ligand diffusions were 1 and 10 ns. The longer time step represents the time needed to average out anomalous diffusion, and the shorter time step was used to ensure statistical accuracy of the method.

Crowding Model. To assess the degree to which crowding in a cytoplasm can affect the ligand−protein association, we built a crowding model. We selected GSTP1 as a crowder and built a CG model for it as we did for HIVp. We fixed the crowder protein models during the simulation. To reduce the computational penalty of additional structures in the system, we removed all nonsurface beads, because they would have negligible nonbonded interactions with the ligand. Full attractive and repulsive van der Waals and electrostatic interactions were enabled. We inserted the crowder proteins in our sphere randomly but reserved a radius around HIVp to ensure that the flaps had full range of motion. Two crowding models were simulated, with crowder protein concentrations of 20% and 40% of total volume (see Figures 3 and S8). These concentrations represent the cytoplasmic conditions.

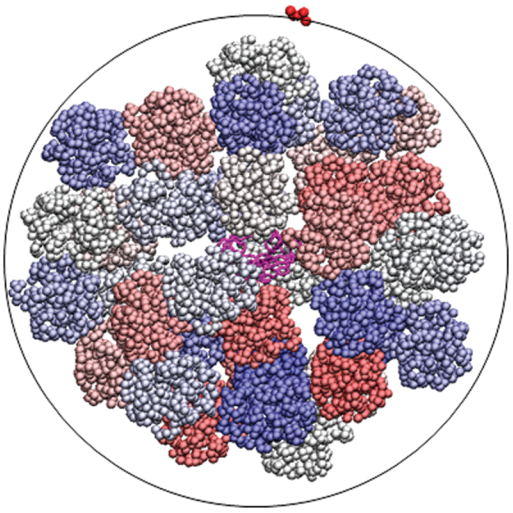

Figure 3. 40% crowding model. HIVp (magenta) is centered in the sphere and each run begins with a ligand (red) on the surface of the sphere. The crowder proteins were inserted in our sphere randomly, but we reserved a radius around HIVp to ensure that the flaps had full range of motion.

RESULTS

Validation of the Coarse-grained Brownian Dynamics Model. We analyzed the trajectories from CGBD simulations of HIVp−ligand association processes to study gated diffusion-controlled encounters. The protein's two polypeptide "flaps" that cover the active site at the dimer interface were in the open position for most of the binding process. To keep the correct flap open/closed fraction and simulation time scale, the HIVp was flexible even when the ligand was far from the protein (Figure 2). Because the simulations are computationally intensive (2 h/1 μs using 1 CPU), we fixed the atoms in the bottom half of HIVp structure. However, fixing half of the protein diminished the possible correlation between the 17 and 39 turns (Figure 1),[33] resulting in a more rigid protein with a small open fraction, as compared with the same CG model with a fully flexible protein. Therefore, we restored the protein's internal motion by modifying a few parameters of angles and dihedral angles based on the open conformations, as detailed in the Methods section. Our partially fixed wild-type HIVp has about 15% open fraction, with mean open and closed times of ~30 ns and ~150 ns, respectively (see Supporting Information for details).

In our BD simulation, the Ermak-McCammon algorithm did not include the hydrodynamic tensor for N particles/beads in order to speed up the calculations, as computing the tensor results in a run time of $O(N^3)$.[38] Because the 3N × 3N diffusion tensor determines the correlated random displacements for N particles/beads, we examined the effect of neglecting this

Table 1. Average Association Times (μs) of the Ligands XK263, Ritonavir and Saquinavir to HIV Protease (HIVp) in Simplified Coarse-Grained (CG) Models[a]

HIVp model	binding criteria	XK263	ritonavir	saquinavir
(A) no interaction	surface of 30-Å sphere	4.49 (1000)	4.67 (1000)	3.97 (1000)
			4.73 (1600)	
(B) fixed closed CG	surface of 30-Å sphere	3.18 (1000)	3.83 (1000)	3.32 (1000)
(C) fixed open CG	$D_{\text{protein-ligand}} \leqslant 11$ Å	5.37 (131)	6.31 (131)	5.04 (131)
(D) Partially flexible CG	$D_{\text{protein-ligand}} \leqslant 11$ Å,	4.70 (2140)	5.49 (2140)	4.63 (2140)
	$D_{\text{tip}} \leqslant 7$ Å	4.60 (6879)		
exp. k_{on}[b] (M^{-1} s^{-1})		$\sim 10^9$	$\sim 10^6$	$\sim 10^5$
exp. ΔG[c] (kcal/mol)		−13.5	−9.3∼ −14.4	−8.5∼ −13.0

[a] The numbers of runs are in the brackets. The standard deviations are in Table S2 of the SI. [b] Ref 55. [c] Ref 21.

hydrodynamic interaction when diffusing our ligands. When we freely diffused the ligands using only one bead with the approximated radius as in eq 2, the translational diffusion coefficients from our BD algorithm of XK263, ritonavir and saquinavir were 2.2×10^{-6}, 2.1×10^{-6}, and 2.2×10^{-6} cm^2/s, respectively, in the dilute solution. These coefficients are in good agreement with the analytical values, thus our algorithm accurately diffuses a sphere solute in free solution. When using the multibead models (see Figure 1), the diffusion coefficients are 0.9×10^{-6}, 0.8×10^{-6}, and 0.9×10^{-6} cm^2/s, respectively. In addition, using 1 or 10 ns as the observed time interval (δt in eq 2) did not affect the computed diffusion constants. Although the simulated diffusion coefficients are 2 to 3 times slower than the analytical values because of the more complicated molecular shape and simplified hydrodynamic interactions, the coefficients keep the same order as those computed from the simple sphere models. Neglecting the hydrodynamic interactions may reduce the diffusion coefficients; however, because our ligands have only 5 to 6 beads, the effects are not significant.

Association Time for Ligand Binding to Simplified HIVp. On the basis of experimental assays, our ligand concentration was set to 100 μM, which is equivalent to the diffusion of one free ligand within a sphere of 160 Å while HIVp is located in the center of the sphere. Instead of computing the association rate constant with a probability-based algorithm such as Northrup-Allison-McCammon (NAM),[39] we computed the association time, the time it takes each ligand to reach the binding site of HIVp. Each run began with a ligand on the surface of the 160-Å sphere, and the simulation stopped when the binding criteria was satisfied. We studied factors such as the size of the active site, intermolecular interactions between the protein and ligands, and intramolecular "gating" motion of the protein, so several simplified models were also tested (Table 1 and Figure 4). We started with the simplest of these models, in which HIVp is approximated by a 30-Å radius sphere without any interactions with the ligands, and the binding is considered successful when a ligand reaches anywhere on its surface. As shown in Table 1, all of the ligands needed nearly 4 μs, on average, to associate with the sphere protein, and the standard deviation was as large as the association time (see Supporting Information for details). With eq 2 and a diffusion coefficient of $\sim 2 \times 10^{-6}$ cm^2/s, the time required for diffusing 160 Å in one dimension takes only ∼0.6 μs, with no analytical association time available. Therefore, we used the association time obtained from this simplest model as our reference.

To study how the intermolecular interactions may affect the association time, we replaced the 30-Å sphere by HIVp with a

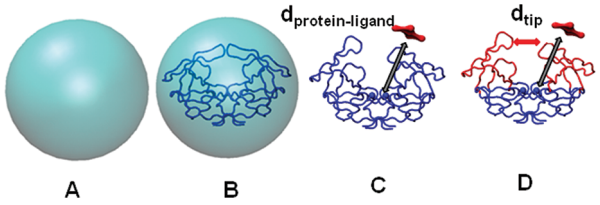

Figure 4. Simplified models for HIVp. (A) 30 Å radius sphere. Ligands "bind" at the surface of the sphere. There is no intermolecular interaction between the protein and the ligand. (B) Fixed HIVp CG model with closed conformation in a 30-Å radius sphere. The intermolecular interaction is introduced, but ligands still "bind" at the surface of the 30-Å radius sphere. (C) Fixed HIVp CG model with open conformation. The ligands bind when they are within 11 Å of the Asp25 of the active site. (D) Partially flexible HIVp CG model with both fixed (blue) and flexible parts (red). The ligands bind when they are within 11 Å of the Asp25 of the active site and the flaps are closed. (also see Table 1).

closed conformation and again considered the binding successful when the ligand reached 30 Å from the center of the protein. Because our ligands are neutral, without long-range electrostatic steering, van der Waals attractions slightly decreased the association time by ∼17% to 20%. We further examined the geometric effects by obtaining the association time for ligands binding to the active site of HIVp with an open conformation. Because the protein is fixed, once the ligand is within 11 Å of the Asp25 in the active site, it is considered bound. This simulation also provides insight into how the ligands would bind to HIVp if there were no "gating" motion. The geometry constraint leads to an increase in mean binding times by 1.5 to 1.7-fold, although the binding site occupies only about 10% of the protein surface. Purely probabilistic models account for such steric constraints by multiplying the Smoluchowski rate for uniform spheres by the probability that, in a random encounter, the two molecules are properly bound. Although the binding site is decreased 90%, the mean binding times did not increase as much, which suggests that the association is accelerated by surface diffusion. The initial contact of the ligand can be anywhere on the surface of HIVp. Instead of immediately leaving HIVp, the ligands diffused laterally along the HIVp surface to search for the correct orientation and protein binding site. The diffusion constants were substantially reduced (see Table 2), but the lateral diffusion largely decreases the association time because the ligand does not need several hundred attempts to reach the binding site of HIVp. Interestingly, diffusion along the z-axis, corresponding to the longest horizontal dimension across the HIVp, was faster for all ligands.

Table 2. Translational Diffusion Coefficients, D_T ($\times 10^{-6}$ cm^2/s), 1000 Samples, for XK263, Ritonavir and Saquinavir with HIVp

model		XK263	ritonavir	saquinavir	dt (ns)
		diffusion through space			
1 bead	approximate radius (Å)	7.3	8.0	7.5	
	D_T	2.25	2.16	2.29	10
		2.24	2.06	2.18	1
multibead	D_T	0.93	0.75	0.93	10
		0.90	0.76	0.90	1
		lateral diffusion on the surface of the protein			
multibead	D_T	0.34	0.26	0.34	1
		0.08	0.07	0.09	10
	D_X[a]	0.45	0.09	0.19	1
		0.11	0.02	0.06	10
	D_Y[a]	0.13	0.34	0.39	1
		0.02	0.11	0.07	10
	D_Z[a]	0.42	0.34	0.44	1
		0.11	0.08	0.13	10

[a] D_X, D_Y, and D_Z represent 1-D translational diffusion coefficients in x, y, and z axes, respectively. See also Figure S4.

Of note, use of a different observation time interval (δt in eq 2) results in different translation coefficients when the ligand diffuses along the protein surface.

Association Pathways and Motions of HIVp. As summarized in Table 1, our CG models for HIVp and ligands did not show large differences between the ligand−protein association times. During the simulations, the flaps of HIVp could open and close, and the final bound state was defined by two distances: the distance between the tip of the flaps and that between the center bead of a ligand and Asp25 of the protein. After the gating motions of the flaps were introduced, the average bindings were approximately 14% faster than those with a fixed open CG model, while the gate only opened ∼15% in the free protein. For example, as shown in Figure 2, before the ligand approached the protein, the flaps opened many times, with the open fraction ∼15%. According to previous studies, these flap motions may be considered "slow gating"; thus, the rate for successful binding should be reduced significantly because only 15% of the attempted bindings can be successful. Therefore, the association time was expected to increase considerably, since the ligand may leave the protein easily when the flaps are not open. However, the ligand can wander around the enzyme until it reaches the neighborhood of the binding site (see Figure S4), and the ligand−protein interactions modify the flap dynamics, making the opening more frequent to facilitate ligand binding.

Because the initial positions of the ligands are far from HIVp, the ligand's initial position and orientation did not affect the general binding pathway. Figure 2 shows the density of the ligands from simulations with different starting positions on the 160-Å sphere, which confirms that all ligands explore the surface of the enzyme. Although studies have suggested that the surface of the "hinge" region of the flaps of HIVp may be a potential ligand binding site,[40,41] all three ligands did not show higher density near the "hinge", but tended to stay longer near the bottom of HIVp, where two cavities can be observed. To determine whether the trapping is an artifact from the fixation of the bottom portion of the enzyme, we also carried out 25 BD simulations with a fully flexible HIVp. In these simulations, ligands also stayed longer in the bottom part of the protein (see Figure S4 of the SI for the density plot), thus the rigid part of HIVp does not produce artificial intermolecular attractions. Although our ligands formed a very stable ligand−protein complex in our coarse-grained model, occasionally the ligands left the binding sites and returned again, as seen in the atomistic simulation.

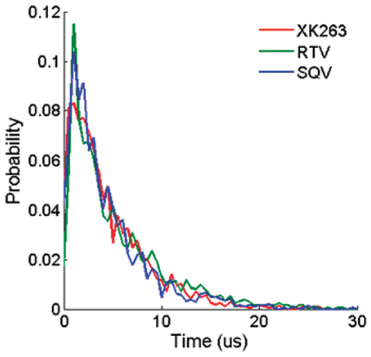

Figure 5. Histogram for the association time for each ligand to bind to HIVp.

With our binding criteria, the association time for XK263 was the shortest, but the times were similar for all three ligands (see Table 1). Note that although instruments for measuring association or dissociation rate constants, such as BIAcore, have a data collection rate of about 1−100 Hz, we set a maximum of 30 μs for our simulations. Figure 5 illustrates the probability of binding at different cutoff times, and XK263 shows slightly higher probability of binding. For example, within 0.6 μs, the analytical time for diffusing a 7.3-Å spherical ligand over 160 Å, the probability of binding for XK263 is only about 20% larger than that for ritonavir and saquinavir. In contrast, the experimental association rate constant for XK263 is about 1000 times larger than those for ritonavir and saquinavir. However, although HIVp−ligand interaction energies and association time are similar for all ligands (see Figure 6), XK263 has much smaller fluctuations in the active site of the protein. Ritonavir and saquinavir did not leave the binding site after they reached the bound state, but they tumbled more frequently and the flaps opened more often. In a few cases, ritonavir and saquinavir show smaller ligand fluctuations and very few flap openings in the bound state, so some may reach the final bound state, and our CG model can form the final bound state. The movement of both ritonavir and saquinavir suggest that they need significantly longer time than XK263 to adjust the conformation to find the final bound state.

Our CG model uses simple criteria to determine successful binding; however, the final bound conformation may need atomistic details, including waters, to fully optimize the final bound complex. Of note, XK263 does not need water molecules, but other peptidomimic ligands such as ritonavir and saquinavir have one crystal water in the binding site in their experimental structures. This work focused on only a wild-type protein. Mutants that change the flap motions of HIVp may affect the ligand−protein association, which is not discussed here.

Ligand Binding in Crowded Environment. To study ligand−protein encounters in the crowded environment of cells, we selected

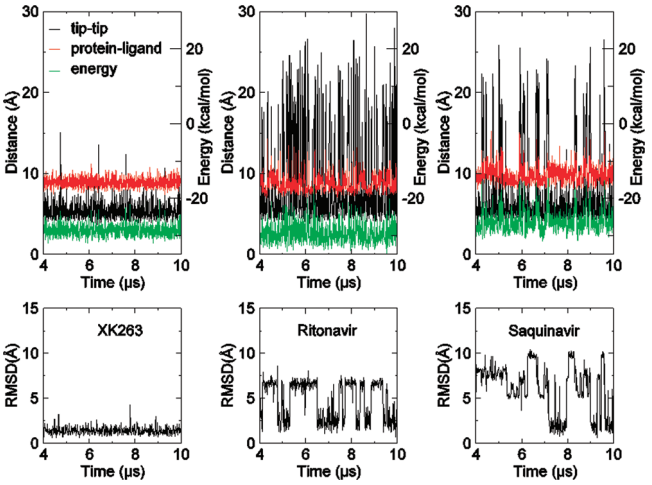

Figure 6. The different behaviors of the ligands at the binding site. The distance between the protein (Asp25) and the center of each ligand, the tip to tip distance, and the ligand−protein interaction energy (top), and the root-mean-square deviation (rmsd) of each ligand after they "bind" (bottom).

Table 3. Mean Association Times, t_{on}, and Diffusion Coefficients, D_T, of XK263 in Crowding Models

interaction	repulsion + attraction		repulsion only	
concentration (%)	20	40	20	40
$t_{on,\ c\%}$ (μs)	11.0	19.0	4.8	4.9
$t_{on,c\%}/t_{on,0\%}$ [a]	2.3	4.0	1.0	1.1
$D_T(10^{-6}cm^2/s)$ [b]	0.41	0.24	0.74	0.75
no. of samples	300	300	300	300

[a] The ratio of the mean association time with c% crowding, $t_{on,\ c\%,}$ to that without crowding, $t_{on,\ 0\%}$. [b] The dt for the translational diffusion coefficient is 10 ns.

and coarse-grained one of the most abundant proteins in human cells, GSTP1, as a crowder agent. All GSTP1 molecules were held rigid and had intermolecular attractive and repulsive interactions with our ligands (eqs S6 and S7 of the SI). Typically, 20% to 40% of the cytoplasmic volume is occupied by macromolecules, and we had two crowding systems: GSTP1 in 20% (20 proteins) and 40% (38 proteins) of excluded volume concentration in our 160-Å sphere, where GSTP1 were randomly placed in the sphere (see Figures 3 and S8 of the SI). GSTP1 is approximately 2.2- times larger than HIVp, and a few small shallow cavities can be observed on the protein surface (see Figure 1C). Previous work showed that crowded conditions can affect the internal dynamics of HIVp.[42] Therefore, we placed GSTP1 far enough, with the closest distance between HIVp and GSTP1 ~26 Å, to avoid any influence of crowding on HIVp internal dynamics, and the intermolecular interactions between HIVp and GSTP1 were also turned off. We also ran BD simulations for HIVp with the crowders to ensure that the fractions of open/closed flaps remain unchanged. Since we fixed GSTP1 proteins during our simulations, two different crowder configurations with various protein orientations were used, and the ligand binding time did not show noticeable difference from the simulations (data not shown).

Table 3 shows the average binding time of different ligands for both crowding environments. The binding is slowed by two- to 4-fold with crowding, which is similar to other simulations.

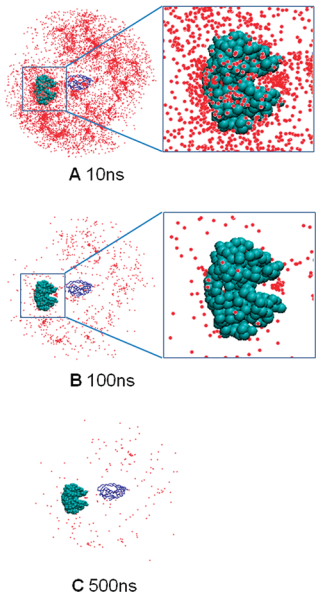

Figure 7. The density of XK263 with 20% crowding. Snapshots at 10 (A), 100 (B), and 500 ns (C). One GSTP1 is shown in the extended illustration (right) to show increased density near the concave region on the surface of GSTP1.

For example, the association time for XK263−HIVp binding increased from 4.7 to 19.0 μs with 40% of the volume occupied by GSTP1. No simple relation was found between binding time and crowder concentration—the attraction between the ligand and GSTP1 slowed the HIVp−ligand association time, but the reduction in ligand diffusion space could also accelerate HIVp−ligand binding. Figure 7 illustrates the density of XK263 when encountering the protease from various starting positions on the 160 Å sphere. The snapshots saved every 10, 100, and 500 ns show an increased density near the concave regions of the GSTP1 surface. With snapshots saved every 500 ns, the clusters of ligand distribution in those regions are less noticeable, thus the drug may be temporarily trapped by the crowders for up to a few hundred ns. To test the effects of intermolecular attractions to HIVp−ligand encounters, we turned off the intermolecular attractive forces between the ligand and GSTP1 and kept only the repulsive potential. The association time was similar to that in dilute solution.

Crowders occupying space may effectively decrease the volume available for the ligand when it diffuses toward the enzyme, which can also be considered as increasing the ligand concentration. We therefore reduced the size of the sphere to 149 and 135 Å to make the ligand concentrations equivalent to the effective concentrations when the space is 20% and 40% excluded by GSTP1, respectively. The association time is reduced 33% and 50%, which showed a positive correlation between the binding time and the space reduced.

DISCUSSION

General Comments. This work aimed to provide a more complete description of ligand−protein association behavior with the occurrence of significant protein conformational changes during the ligand−protein encounter. We also studied the influence of crowder proteins on the simulated binding time, which provides insights into molecular binding in the crowded

environment. Instead of computing the association rate constants, we reported the mean association times (conceptually similar to inverse rates) for ligands in a dilute solvent and crowded environment. The ligand concentration was set to 100 μM. As previously stated, a 1 M ligand concentration is equivalent to 1 molecule/1660 Å^3, so our model freely diffuses one ligand in a volume inside a 160 Å sphere. If a ligand diffuses outside the sphere, then we ensured a consistent concentration by placing a ligand in the opposite direction of the 160 Å sphere, as mentioned in the Methods. One can also place multiple ligands in a larger volume for a 100 μM ligand concentration to simulate the effects of ligand−ligand interactions on association times. Because our ligands are neutral, without noticeable long-range electrostatic interactions, we assume that the influence of these interactions can be ignored. However, for charged ligands, or in substantially high ligand concentration, the ligands may interact with each other or compete for the binding site of the protein.[43]

Unlike a natural substrate that will be cleaved and then the product released, the chemical compound stays in the binding site. If the HIVp concentration is not very low and the dissociation rate is slow (e.g., $< 10^{-2}$ s^{-1}), then a ligand may already bind in the binding site of HIVp when our ligand diffuses close to the binding site. This study assumes that the HIVp concentration is significantly lower than the ligand concentration, so when the ligand diffuses near HIVp, the enzyme is in a free state and no compound occupies the binding site. The simulation times are sufficiently long for the ligands to have enough time to reach the binding site. More than 99.5% of ligands could bind to the protein within 30 μs in the dilute system. However, to reach the same success rate, we needed a 100-μs simulation time when crowders are present.

At the given ligand concentration, the association rate constant can conceptually be converted to the approximate association time and vice versa. For example, the association time of XK263 estimated from the experimental association rate constant at the ligand concentration of 100 μM is ∼10 μs. The association rate constant computed by the Smoluchowski expression using the diffusion coefficients from a single sphere, $D_T = 2.25 \times 10^{-6}$cm^2/s for XK263, yields $k_{on} \approx 6 \times 10^9$ M^{-1}s^{-1}. The approximate association time based on this k_{on} is ∼1.7 μs. Notably, the diffusion-control association rate constant for two molecules without intermolecular interactions is 10^9-10^{10} M^{-1}s^{-1} (see Smoluchowski equation in the introduction). Our calculations of XK263 suggest that the gating effects do not remarkably slow down the association time, and the association is close to diffusion-control processes. Ligands may diffuse away when they reach the nonbonding site regions of HIVp or when the flaps are closed. Nevertheless, in most cases, XK263 can diffuse on the HIVp surface to wait for the flaps to open or bind from the side when the flaps are not fully open. In addition, the ligand−protein interactions can facilitate the opening of the flaps, which alters the flap dynamics from "slow gating" to "fast gating" for HIVp. Previous modeling showed that the flap dynamics leads to "slow gating".[19] According to the gating theory, a slow gating effect suggests that the gated association rate constant is simply the ungated rate constant times the probability of the gate opening.[14,15] Although the opening of the flaps is not large, the space is spacious enough for small drug binding. The flaps can open fast enough to not affect the binding of the ligand (fast gating). Note that although our ligands are small molecules, natural substrates are large polypeptides that may need fully open flaps. Consequently, the slow gating feature may remain depending on the ligand. The modifications of flap motions due to intermolecular interactions are consistent with previous studies.[33,34,44] In addition, the facilitation of binding by diffusion of ligands along the protein surface has been previously suggested.[45]

As expected, the analytical results from eq 2 are similar for all ligands tested because they have similar sizes and diffusion coefficients. Surprisingly, the mean association times from our simulations were also similar for all ligands, which is inconsistent with experimental data. While the three ligands have similar binding affinities, the association rate constants are a few orders slower in saquinavir and ritonavir (Table 1). Before the ligands reach the binding site, their association processes were similar, and they can all induce flap opening. All ligands may become temporarily trapped in the bottom of the HIVp structure, with subsequent diffusion along the protein surface. The attraction energies for nonbonded van der Waals interactions (eq S6 of the SI) may be overestimated, so once a ligand forms an attractive potential with HIVp, it stays with the protein, thus resulting in a rate constant near that of diffusion-limit association. However, our ligands do not stay with the crowders, which use the same nonbonded potential energy function as those used in HIVp. Therefore, overstabilizing ligand−protein interactions is not likely. During our simulations, sometimes a ligand left the enzyme after encountering the surface but returned to attempt binding two or three times. As a result, the standard deviations of the association times are all large (see Figure S5 of the SI).

After the ligand diffused into the binding site and the flaps of HIVp were fully closed, XK263 showed less fluctuation in the protein pocket, but saquinavir and ritonavir kept tumbling in the binding site. We also carried out several 200 to 300 μs long simulations, and saquinavir and ritonavir could reduce their fluctuation after >100 μs in a few runs. Therefore, if we include the ligand stability in the HIVp binding site as a criterion to define the bound state, then saquinavir and ritonavir require relatively long association times as compared with XK263. Despite the similar intermolecular interactions between HIVp and different ligands in our CG model, a few details crucial to stabilize the ligand-bound state cannot be fully captured by the model. As previously mentioned, one detail is the role of water molecules in the binding-site to stabilize the ligands. During the binding processes, the association time spent to replace and recruit the bound water can differ greatly between cyclic urea ligands (e.g., XK263) and peptidomimetic drugs (e.g., saquinavir and ritonavir). As a result, to thoroughly study the second step of the ligand binding mechanism, an atomistic force field with an explicit water model may be necessary. Nevertheless, our CG model with BD simulations is an efficient method to gain insight into and an overview of the binding processes in both dilute and crowded environments.

Binding in Crowded Environments. Under crowding conditions, as in cytoplasm, the effective molar concentration of a ligand is increased with decreased volume of the solution. The macromolecules may interact nonspecifically with the ligand and the target protein in vivo. Previous publications showed that crowders can affect protein dynamics and folding.[11,42,46−51] Even though the hydrodynamic interactions have been suggested to be significant in modeling crowding and macromolecular motion,[52,53] the hydrodynamic interaction is insignificant because the ligands are small and crowders are held fixed. This work focused on the crowding effects contributed by the excluded diffusion volume and crowder−ligand interactions. We first increased the ligand

concentration by reducing the diffusion volume to a space in the 145 and 139 Å sphere, and the association time decreased with higher ligand concentrations. Without intermolecular attraction forces, we expected that the association time may also be shorter because the available space for ligand diffusion was cut by 20% to 40% (volume occupied by the crowders).[54] However, the association time is the same with and without crowders, although the diffusion space for XK263 was reduced by 20% and 40%. The volume was reduced because of the volume excluded by the crowders, and the translational diffusion coefficients were also reduced, so the crowding actually slows down the diffusion, but it is canceled out by the excluded volume effect. Thus, when only repulsive forces are considered, excluded volume effects are not significant for diffusing small molecules to a protein target. The distance that the ligand needs to travel may be the main determinant. In contrast, if we consider all intermolecular attractions, then the association times increase 2.3-fold with 20% crowders and nearly 4-fold with 40% crowders. The translational diffusion coefficients are consistent with the association times (Table 3). A series of BD simulations show that intermolecular attractions are the major factors that might account for the large increase in association time of small molecules binding to HIVp. Notably, our ligands have only small partial charges in each bead and are neutral; thus, without long-range and strong electrostatic attractions, nonspecific van der Waals attractions contribute significantly to crowding effects. More studies with various sizes and charges of ligands are needed to study the crowding effects.

CONCLUSIONS

Ligand−protein complex formation can be viewed as a multistep process. The initial step is a collision between two molecules in a solution. In the absence of electrostatic steering forces, this step is driven mainly by translational diffusion, and a crowded environment slows down the small-molecule diffusion process by 2- to 4-fold in our simulations. The second step involves several intermediates, before the third step, the final formation of the bound complex. This work investigated the molecular recognition processes that connect the initial and final steps of the binding process. We found that the ligand−protein interactions alter the flap dynamics from slow to fast gating. The interactions induced the flaps to open more frequently, although not sizably. Therefore, the gating effects became less significant in our work. Three ligands bind to HIVp with similar association times, but saquinavir and ritonavir fluctuated considerably in the bound state, so the two drugs may need a longer time to adjust their contact interface with HIVp. Even though our CG model cannot reveal all details, including water rearrangement, during ligand binding, the model provides valuable insights into the binding processes and guides us to further studies using detailed atomistic models. CGBD simulations can also simulate ligand−protein associations in the physiological ligand concentration.

ASSOCIATED CONTENT

Supporting Information. Additional data, coarse-grained parameters, and potential functions are included. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: 951-827-7263; E-mail: chiaenc@ucr.edu.

ACKNOWLEDGMENT

We thank Dr. Joanna Trylska for valuable suggestions and discussion. This research was supported by the U.S. National Science Foundation (MCB-0919586). Additional support from the University of California, Riverside, Computer and Communications, and ShaRCS, the University of California Shared Research Computing Services Cluster, which is technically supported by multiple U.C. information technology divisions and managed by the University of California, Office of the President, is gratefully acknowledged.

REFERENCES

(1) Mollerup, J. M. *Chem. Eng. Technol.* **2008**, *31*, 864–874.
(2) Janin, J. *Prog. Biophys. Mol. Biol.* **1995**, *64*, 145–166.
(3) Lundberg, S.; Backman, L. *Methods Enzymol.* **1994**, *228*, 241–254.
(4) Mereghetti, P.; Kokh, D.; McCammon, J. A.; Wade, R. C. *BMC Biophys.* **2011**, *4*, 2.
(5) Gabdoulline, R. R.; Wade, R. C. *Methods* **1998**, *14*, 329–341.
(6) Smoluchowski, M. V. *Z. Phys. Chem.* **1917**, *92*, 129–168.
(7) Schreiber, G.; Haran, G.; Zhou, H. X. *Chem. Rev.* **2009**, *109*, 839–860.
(8) Furfine, E. S.; Dsouza, E.; Ingold, K. J.; Leban, J. J.; Spector, T.; Porter, D. J. T. *Biochemistry* **1992**, *31*, 7886–7891.
(9) Luty, B. A.; McCammon, J. A. *Mol. Simul.* **1993**, *10*, 61–65.
(10) Horn, J. R.; Sosnick, T. R.; Kossiakoff, A. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 2559–2564.
(11) Kozer, N.; Kuttner, Y. Y.; Haran, G.; Schreiber, G. *Biophys. J.* **2007**, *92*, 2139–2149.
(12) Wang, J.; Zhang, K.; Lu, H. Y.; Wang, E. *Biophys. J.* **2006**, *91*, 866–872.
(13) McCammon, J. A.; Northrup, S. H. *Nature* **1981**, *293*, 316–317.
(14) Northrup, S. H.; McCammon, J. A. *J. Am. Chem. Soc.* **1984**, *106*, 930–934.
(15) Szabo, A.; Shoup, D.; Northrup, S. H.; McCammon, J. A. *J. Chem. Phys.* **1982**, *77*, 4484–4493.
(16) Ellis, R. J. *Curr. Opin. Struct. Biol.* **2001**, *11*, 114–119.
(17) Kohl, N. E.; Emini, E. A.; Schleif, W. A.; Davis, L. J.; Heimbach, J. C.; Dixon, R. A. F.; Scolnick, E. M.; Sigal, I. S. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4686–4690.
(18) Krausslich, H. G.; Ingraham, R. H.; Skoog, M. T.; Wimmer, E.; Pallai, P. V.; Carter, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 807–811.
(19) Chang, C. E.; Shen, T.; Trylska, J.; Tozzini, V.; McCammon, J. A. *Biophys. J.* **2006**, *90*, 3880–3885.
(20) Markgren, P. O.; Lindgren, M. T.; Gertow, K.; Karlsson, R.; Hamalainen, M.; Danielson, U. H. *Anal. Biochem.* **2001**, *291*, 207–218.
(21) Shuman, C. F.; Markgren, P. O.; Hamalainen, M.; Danielson, U. H. *Antiviral Res.* **2003**, *58*, 235–242.
(22) Spinelli, S.; Liu, Q. Z.; Alzari, P. M.; Hirel, P. H.; Poljak, R. J. *Biochimie* **1991**, *73*, 1391–1396.
(23) Ang, W. H.; Parker, L. J.; De Luca, A.; Juillerat-Jeanneret, L.; Morton, C. J.; Lo Bello, M.; Parker, M. W.; Dyson, P. J. *Angew. Chem., Int. Ed.* **2009**, *48*, 3854–3857.
(24) Lam, P. Y. S.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N.; Chang, C. H.; Weber, P. C.; Jackson, D. A.; Sharpe, T. R.; Ericksonviitanen, S. *Science* **1994**, *263*, 380–384.
(25) Clemente, J. C.; Stow, L. R.; Janka, L. K.; Jeung, J. A.; Desai, K. A.; Govindasamy, L.; Agbandje-McKenna, M.; McKenna, R.; Goodenow, M. M.; Dunn, B. M., In vivo, kinetic, and structural analysis of the development of ritonavir resistance. DOI: 10.2210/pdb2b60/pdb.
(26) King, N. M.; Prabu-Jeyabalan, M.; Bandaranayake, R. M.; Nalam, M. N. L.; Ozen, A.; Hailloglu, T.; Schiffer, C. A., Extreme entropy-enthalpy compensation correlates with a drug resistant variant of HIV-1 protease. DOI: 10.2210/pdb3ekq/pdb.
(27) Gopal, S. M.; Mukherjee, S.; Cheng, Y. M.; Feig, M. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1266–1281.

(28) Chu, J. W.; Izvekov, S.; Voth, G. A. *Mol. Simul.* **2006**, *32*, 211–218.
(29) Muller-Plathe, F. *Chemphyschem* **2002**, *3*, 754–769.
(30) Tozzini, V. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144–150.
(31) Trylska, J.; Tozzini, V.; McCammon, J. A. *Biophys. J.* **2005**, *89*, 1455–1463.
(32) Tozzini, V.; McCammon, J. A. *Chem. Phys. Lett.* **2005**, *413*, 123–128.
(33) Tozzini, V.; Trylska, J.; Chang, C. E.; McCammon, J. A. *J. Struct. Biol.* **2007**, *157*, 606–615.
(34) Chang, C. E. A.; Trylska, J.; Tozzini, V.; McCammon, J. A. *Chem. Biol. Drug Design* **2007**, *69*, 5–13.
(35) Reith, D.; Putz, M.; Muller-Plathe, F. *J. Comput. Chem.* **2003**, *24*, 1624–1636.
(36) Ermak, D. L.; McCammon, J. A. *J. Chem. Phys.* **1978**, *69*, 1352–1360.
(37) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. *Comput. Phys. Commun.* **1991**, *62*, 187–197.
(38) Geyer, T.; Winter, U. *J. Chem. Phys.* **2009**, 130.
(39) Northrup, S. H.; Allison, S. A.; McCammon, J. A. *J. Chem. Phys.* **1984**, *80*, 1517–1526.
(40) Judd, D. A.; Nettles, J. H.; Nevins, N.; Snyder, J. P.; Liotta, D. C.; Tang, J.; Ermolieff, J.; Schinazi, R. F.; Hill, C. L. *J. Am. Chem. Soc.* **2001**, *123*, 886–897.
(41) Brynda, J.; Rezacova, P.; Fabry, M.; Horejsi, M.; Stouracova, R.; Soucek, M.; Hradilek, M.; Konvalinka, J.; Sedlacek, J. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2004**, *60*, 1943–1948.
(42) Minh, D. D. L.; Chang, C. E.; Trylska, J.; Tozzini, V.; McCammon, J. A. *J. Am. Chem. Soc.* **2006**, *128*, 6006–6007.
(43) Senapati, S.; Wong, C. F.; McCammon, J. A. *J. Chem. Phys.* **2004**, *121*, 7896–7900.
(44) Li, D. C.; Liu, M. S.; Ji, B. H.; Hwang, K.; Huang, Y. G. *J. Chem. Phys.* **2009**, 130.
(45) Marashi, S. A.; Behrouzi, R. *Biochem. Biophys. Res. Commun.* **2005**, *333*, 1–4.
(46) Długosz, M.; Trylska, J. *BMC Biophys.* **2011**, 4.
(47) Qin, S. B.; Minh, D. D. L.; McCammon, J. A.; Zhou, H. X. *J. Phys. Chem. Lett.* **2010**, *1*, 107–110.
(48) Zaid, I. M.; Lomholt, M. A.; Metzler, R. *Biophys. J.* **2009**, *97*, 710–721.
(49) Cheung, M. S.; Klimov, D.; Thirumalai, D. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 4753–4758.
(50) Stagg, L.; Christiansen, A.; Wittung-Stafshede, P. *J. Am. Chem. Soc.* **2011**, *133*, 646–648.
(51) Dix, J. A.; Verkman, A. S. *Annu. Rev. Biophys.* **2008**, *37*, 247–263.
(52) McGuffee, S. R.; Elcock, A. H. *PLoS Comp. Biol.* **2010**, *6*.
(53) Ando, T.; Skolnick, J. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 18457–18462.
(54) Wieczorek, G.; Zielenkiewicz, P. *Biophys. J.* **2008**, *95*, 5030–5036.
(55) Markgren, P. O.; Schaal, W.; Hamalainen, M.; Karlen, A.; Hallberg, A.; Samuelsson, B.; Danielson, U. H. *J. Med. Chem.* **2002**, *45*, 5430–5439.