JCTC Journal of Chemical Theory and Computation

Docking Ligands on Protein Surfaces: The Case Study of Prion Protein‡

Agata Kranjc,[†,#] Salvatore Bongarzone,[†,#] Giulia Rossetti,[†,#] Xevi Biarnés,[†,#,||] Andrea Cavalli,[§,⊥] Maria Laura Bolognesi,[§] Marinella Roberti,[§] Giuseppe Legname,[¶,#] and Paolo Carloni*[,†,#,||]

Statistical and Biological Physics Sector, Neurobiology Sector, International School for Advanced Studies (SISSA), SISSA-Unit, Italian Institute of Technology, 34014 Trieste, Italy, Department of Pharmaceutical Sciences, Alma Mater Studiorum, University of Bologna, 40126 Bologna, Italy, Department of Drug Discovery and Development, Italian Institute of Technology, 16163 Genova, Italy, and CNR-INFM-DEMOCRITOS Modeling Center for Research in Atomistic Simulation, 34014 Trieste, Italy

Received May 20, 2009

Abstract: Molecular docking of ligands targeting proteins undergoing fibrillization in neurodegenerative diseases is difficult because of the lack of deep binding sites. Here we extend standard docking methods with free energy simulations in explicit solvent to address this issue in the context of the prion protein surface. We focus on a specific ligand (2-pyrrolidin-1-yl-*N*-[4-[4-(2-pyrrolidin-1-yl-acetylamino)-benzyl]-phenyl]-acetamide), which binds to the structured part of the protein as shown by NMR (Kuwata, K. et al. *Proc Natl Acad Sci U.S.A.* **2007**, *104*, 11921−11926). The calculated free energy of dissociation (7.8 ± 0.9 kcal/mol) is in good agreement with the value derived by the experimental dissociation constant ($K_d = 3.9\ \mu$M, corresponding to $\Delta G^0 = -7.5$ kcal/mol). Several binding poses are predicted, including the one reported previously. Our prediction is fully consistent with the presence of multiple binding sites, emerging from NMR measurements. Our molecular simulation-based approach emerges, therefore, as a useful tool to predict poses and affinities of ligand binding to protein surfaces.

Introduction

Recent developments in molecular docking protocols (MDPs) allow one to predict accurately ligand poses in their target binding sites.[1] In several cases, the reason for their success lays in the coupling of traditional scoring function-based approaches with molecular simulation approaches[2] (such as soft harmonic modes,[3] molecular dynamics simulations,[4,5] and relaxed complex method[6,7]). The latter introduces conformational flexibility of the target, accounting for the fact that proteins are in constant motion between different conformational states.[8] These may be locally altered when a ligand is bound.[8]

In spite of these successes, there are still many important cases for which MDPs are challenged. These include the prediction of the following poses: transition metal and/or alkylating drugs, ligands causing large structural changes, and ligands *not* binding to specific pockets. The latter is common in proteins undergoing fibrillization in neurodegenerative diseases. Here we propose an enhanced molecular docking protocol (EMD, Scheme 1) that extends MDPs with

* Corresponding author. Telephone: +39 0403787407. Fax: +39 0403787528. E-mail: carloni@sissa.it.

‡ Agata Kranjc, Salvatore Bongarzone, Giulia Rossetti, and Xevi Biarnés contributed equally to this work.

† Statistical and Biological Physics Sector, International School for Advanced Studies (SISSA).

¶ Neurobiology Sector, International School for Advanced Studies (SISSA).

SISSA-unit, Italian Institute of Technology.

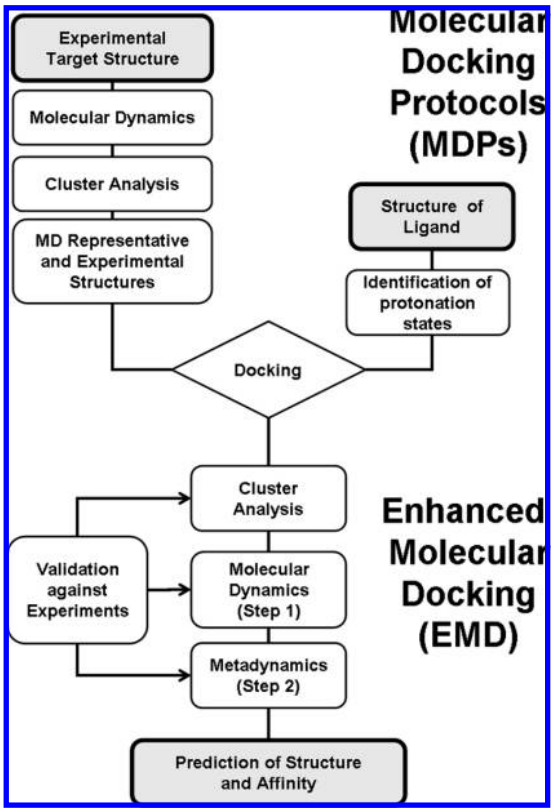
§ Department of Pharmaceutical Sciences, University of Bologna.

⊥ Department of Drug Discovery and Development, Italian Institute of Technology.

|| CNR-INFM-DEMOCRITOS Modeling Center for Research in Atomistic Simulation.

10.1021/ct900257t CCC: $40.75
Published on Web 08/14/2009

Scheme 1. MDPs are used to guess putative ligand binding regions on target surfaces based on structural information of the two separated moieties. Structural information of the target may come from the experiment and, in some cases, also the molecular simulation. Ligands may be docked on the entire structure (like in this work) or a putative binding site. Cluster analysis is used to group molecular dynamics (MD) conformers and/or ligand/target adducts into representative structures. In the EMD approach, MD simulations may be used to relax the structures and investigate the role of hydration. Enhanced sampling simulation techniques in explicit solvent (here metadynamics) allow the exploration of the ligand binding space and the prediction of the free energy of the binding. Comparison against experimental data, in this case structural information inferred by NMR chemical shift perturbations as well as with affinity measurements, allows the protocol to be validated[17]



free energy simulations in explicit solvent to predict the structure and the energetics of ligands binding to protein surfaces.

Protein misfolding, followed by self-association and subsequent deposition, has been observed in the brain tissues of patients affected by different neurodegenerative disorders. Diverse proteins have been shown to follow this process, including amyloid-β (in Alzheimer's disease), α-synuclein (in Parkinson's disease), huntingtin (in Huntington's disease), and prion protein (in prion diseases).[9–11] The protocol we propose here may be exploited in these cases for the design of ligands that, by stacking onto protein surfaces, may disrupt protein−protein interactions and, thus, inhibit protein self-assembly.

In this study, we apply our EMD protocol in the context of the cellular form of human prion protein ($HuPrP^C$).

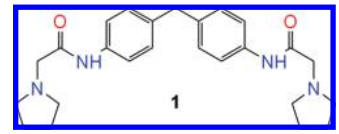


Figure 1. Chemical structure of 2-pyrrolidin-1-yl-*N*-[4-[4-(2-pyrrolidin-1-yl-acetylamino)-benzyl]-phenyl]-acetamide (GN8) considered in this work.

$HuPrP^C$ may convert into a pathogenic form (PrP^{Sc}, scrapie prion protein),[12] which is involved in the epidemics of the bovine spongiform encephalopathy (BSE) and the new variant Creutzfeldt-Jakob disease (nvCJD).[13,14] For these diseases, neither an early diagnosis nor a cure is currently available.[15] Therefore, there is great interest in designing ligands binding to $HuPrP^C$, which may interfere with its conversion and interaction with other self-aggregating proteins, such as amyloid-β.[16]

Recently, the ligand GN8 (2-pyrrolidin-1-yl-*N*-[4-[4-(2-pyrrolidin-1-yl-acetylamino)-benzyl]-phenyl]-acetamide, **1** in Figure 1)[17] has been shown to bind $MoPrP^C$ in the μM range. $MoPrP^C$ is highly similar to the $HuPrP^C$. The sequence similarity is as high as 98%, and the root-mean-square difference (rmsd) of the backbone between the molecular dynamics (MD) structures of $HuPrP^C$ (PDB code: 1HJM)[18] is the same (0.26 ± 0.02 nm) as that between the NMR structures of $HuPrP^C$ and $MoPrP^C$ (PDB code: 1AG2)[19] (0.27 nm). Therefore, significant changes of the structure on passing from the mouse to the human protein are not expected. For the **1**-$MoPrP^C$ adduct, NMR chemical shift perturbations of $MoPrP^C$ on protein residues induced by ligand binding have been reported.[17] These affect most significantly amino acid residues on one side of the protein surface (Arg156, Asn159 @ H1−S2 loop, Lys194 @ H2, Glu196, Thr199 @ H2−H3 loop, and Val210 @ H3). In addition, Val189 and Thr192, located on the other side of the PrP^C surface, are also perturbed (Figure 2A). All these residues are conserved on passing from $MoPrP^C$ to $HuPrP^C$. These perturbations have been ascribed to ligand binding, suggesting that multiple binding sites may be present. An ad hoc model of the **1**-$MoPrP^C$ adduct, constructed by docking and energy minimization, exhibited a single binding mode of GN8 connecting Asn159 and Glu196.[17] Subsequent quantum mechanical studies,[20] based on this model, pointed out that these two residues, along with Gln160 and Lys194, are important for the binding. However, such a single binding mode was not consistent with the presence of the contacts between the ligand and Val189, Thr192, and Thr199. A similar NMR study to map prion protein binding sites has been made only for one other ligand, quinacrine (Supporting Information, Figure S1).[21] The latter, however, has been later suggested not to bind to PrP^C but rather to PrP^{Sc} or other chemical chaperons involved in the prion propagation.[22]

The key ingredient of the protocol proposed in Scheme 1 is given by the type of free energy approach used. Several powerful methods are available for predicting ligand binding free energies by means of molecular simulation.[23,24] However, predictions have been made so far to targets with binding sites well characterized by X-ray crystallography or NMR experiments. Here we use the metadynamics[25] ap-

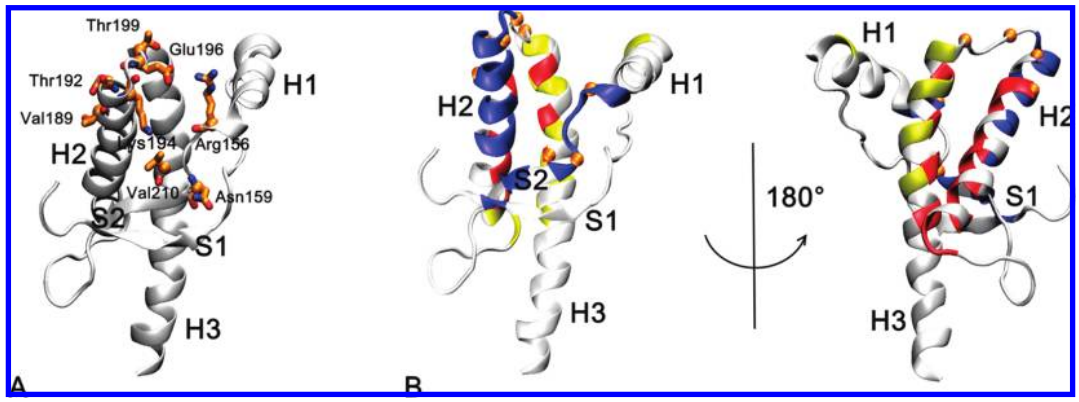


Figure 2. (A) Residues involved in GN8 binding to the prion protein (in licorice), as emerging from chemical shift changes.[17] (B) Three different binding regions (I, II, and III shown in blue, red, and yellow, respectively), as obtained after MDP procedures (see Scheme 1). Orange spheres represent compound **1** binding aminoacids defined by the NMR chemical shift study. The figure shows HuPrP^C. This is very similar to the MoPrP^C (sequence similarity = 98%) for which experiments have been carried out.

proach in its bias exchange variant.[26] This approach provides the free energy as a function of several reaction coordinates (such as geometrical distances, polar contacts, and water-mediated interactions), which characterize the ligand both binding to its target and dissociating from it.[27–30] Although GN8-PrP^C interaction energies have been provided by quantum chemical methods,[20] no calculation of free energy has been so far reported. As with several other techniques,[24] it may allow also simulating the whole molecular recognition process. This in turn may allow characterizing multiple binding sites of the ligand onto the proteins surface, such as those emerging from NMR in the **1**-MoPrP^C complex.

The proposed EMD protocol turns out to provide structural prediction consistently with the NMR data and affinity, which is in agreement with experimental data. The EMD protocol emerges, therefore, as a useful approach to investigate ligands sticking on protein surfaces.

Results and Discussion

In this study, we focus on the binding of the compound **1** to the surface of the HuPrP^C protein. Compound **1** is a symmetric molecule, composed of two pyrrolidine rings connected by acetamides to a diphenylmethane core (Figure 1). Two different conditions are considered: at neutral pH, where experimental affinity has been measured,[17] and at acidic pH, where NMR chemical shift perturbations[17] have been used to provide information on the amino acids involved in the binding.

We use the computational protocol summarized in Scheme 1: (i) Identification of the ligand protonation state at neutral and acidic pH. (ii) Use of MDPs to provide a first guess of the putative binding regions (iii). Use of MD simulations to relax the structure in an aqueous solution (step 1 in EMD). (iv) Use of metadynamics to predict the energetics of the binding of **1** to the protein. (v) Use of metadynamics to predict the binding poses of the compound (step 2 in EMD).

1. Protonation State of Compound 1. This compound can exist in different protonation states in which none, one, or both the tertiary nitrogen atoms of the pyrrolidine rings are protonated (Supporting Information, Figure S2). At pH = 7.4, at which the K_d has been measured, approximate pK_a calculations based on ref 31 (see Methods Section for details) allow us to suggest that, in water, **1** is present not only in the neutral form ($\mathbf{1}^0$) but also in the monoprotonated one ($\mathbf{1}^+$) (Supporting Information, Figure S2). In the latter, one of the two pyrrolidine nitrogen atoms is protonated.

At pH = 4.5, at which the NMR experiments were performed, the same calculations lead us to the conclusion that, in water, the ligand is mainly diprotonated ($\mathbf{1}^{2+}$), with both pyrrolidine nitrogen atoms protonated. Small amounts of $\mathbf{1}^+$ are also present. The calculated concentration of $\mathbf{1}^0$ in bulk water is very low (Supporting Information, Figure S2). However, one should keep in mind that the ligand−protein binding does not occur in pure water and the influence of the electrostatic field of the protein should be accounted for. Indeed, simple electrostatic potential calculations (see Methods Section for further details) show an increase of the positive charge density in the region of the protein defined by the NMR contacts (Supporting Information, Figure S6). This suggests that the protein environment will favor the accumulation of neutral $\mathbf{1}^0$. Therefore, binding poses involving the neutral form should be considered even at an acidic pH. Based on these results, we performed calculations on all of the three protomers.

2. Binding Regions of HuPrP^C Emerging from MDP. The three protomers were docked independently to the HuPrP^C NMR structure and to 20 different conformers obtained from a 20 ns MD simulations of the protein in aqueous solution. The putative binding regions I, II, and III were identified (Figure 2B). I is defined by the H2 helix and the loop connecting β-sheet S2 and helix H1. II consists of the H2−H3 helices. III is defined by the H3 helix, the N-term of H2 helix, and the loop between H1 helix and S1 β-sheet.

Binding region I is the only site which involves residues changing chemical shifts upon binding with compound **1**, and it is closer to all the other residues involved in the binding.[17] It was, therefore, the only one selected for subsequent free energy studies.

The adducts for each of the three protomers docked at the binding region I underwent 10 ns of MD calculations in an

aqueous solution. The ligands maintained completely ($\mathbf{1}^+$ and $\mathbf{1}^{2+}$) or partially ($\mathbf{1}^0$) the pose identified in the docking (see Supporting Information for details). Most importantly, the structural determinants of the three protomers turned out to be consistent with most ligand−protein contacts identified by NMR (Supporting Information, Table S1). However, the ligand−protein contacts with Val189, Thr192, Thr199, and Val210 could not be predicted. The simpler docking approach, combined with the energy minimization of protomer $\mathbf{1}^0$, provided similar results (Figure 3 in Kuwata et al.[17]).

Free energy calculations were used to explore the ligand binding space in an explicit solvent. These simulations identified alternative binding poses for each protomer of the ligand and predicted the dissociation free energy for $\mathbf{1}^0$ and $\mathbf{1}^+$. The free energy simulations were performed as a function of six collective variables that took into account rearrangements of the ligand and the protein, the hydrogen-bond contacts and the water bridges (see Methods Section). These variables have been already used to characterize ligand-target molecular recognition processes using the metadynamics approach.[27–30]

3. The HuPrP^C−$\mathbf{1}^0$ Complex. In the lowest free energy cluster identified by the metadynamics calculations, $\mathbf{1}^0$ is located in the wide cleft formed by helices H1, H2, and H3 (1°.B1 in Figure 3A), similar to the model proposed by Kuwata et al. for MoPrP^C.[17] The contacts $\mathbf{1}^0$ forms with the HuPrP^C are consistent with the reported chemical shift changes on Glu196, upon GN8 binding (Supporting Information, Table S1), as well as with a recent quantum chemical study.[20] They are also consistent with the chemical shift changes on Arg156, Thr199, and Val210 upon GN8 binding (Supporting Information, Table S1). The phenyl groups of $\mathbf{1}^0$ form a π-cation interaction with Arg156, and a water-mediated hydrogen-bond is present between Thr199 and the pyrrolidine nitrogen (N1; Figure 1). The pyrrolidine ring forms hydrophiobic interactions with Val210 (as well as with Pro158 and Thr183). The HuPrP^C−$\mathbf{1}^0$ complex is further stabilized with a direct hydrogen-bond between Thr190 and the carbonyl group of $\mathbf{1}^0$ (O2; Figure 1). The unbound state of HuPrP^C−$\mathbf{1}^0$ system corresponds to a conformation in which the ligand has no contact with the protein. The conformation of Lys194 changes upon ligand dissociation (Supporting Information, Figure S3). This is consistent with the significant chemical shift change reported for this amino acid.[17] This contrasts with a recent quantum chemical investigation, which points to the role of Lys194 for the bridging conformation of the GN8/PrP^C complex.[20] This discrepancy may be due to the fact that here we consider free energies in solution, while ref 20 presents interaction energies *in vacuo*. Smaller conformational changes were also observed for other residues present in the H2−H3 loop (res. 195−199, Supporting Information, Figure S3). These rearrangements were not observed with the MD calculations (see Supporting Information), possibly because they are induced during the ligand binding process simulated here.

The unbound state of HuPrP^C−$\mathbf{1}^0$ is 5.5 kcal/mol ± 0.9 higher in energy with respect to the bound state described previously. The ligand is not completely detached from the protein, although it is already separated by five layers of water molecules between the two moieties. The remaining free energy for the complete dissociation was estimated as the mean electrostatic interaction between the two molecules in implicit solvent (see Methods Section). This turned out to be −0.7 kcal/mol. Thus, we estimated the dissociation free energy to be 4.7 kcal/mol in our simulation conditions. Considering also the concentration of the species in the simulation box (see Methods Section), the standard free energy of dissociation is estimated to be 7.8 kcal/mol. This is in very good agreement with the experimental value of 7.5 kcal/mol (corresponding to $K_d = 3.9$ μM) reported by Kuwata et al.[17]

4. The HuPrP^C−$\mathbf{1}^+$ Complex. Four different stable conformations of $\mathbf{1}^+$ were identified on HuPrP^C surface (Figure 3B). In the lowest free energy state, $\mathbf{1}^+$ lays along the loop connecting helices H2 and H3 (1^+.B1 in Figure 3B). It forms a remarkable hydrophobic interaction with Thr199, consistently with the chemical shift changes reported for this residue.[17] The amidic nitrogen atoms of $\mathbf{1}^+$ (N3 and N4; Figure 1) are hydrogen-bonded to Thr201 and Asn197, respectively. This induces a subtle conformational change of the Glu196 and Asn197 backbone upon ligand binding, which may be the reason for the chemical shift displacement reported experimentally for Glu196 (Supporting Information, Table S1). Additionally, the neutral pyrrolidine ring of $\mathbf{1}^+$ is kept by the hydrophobic cleft formed by Ile184, Thr188, Phe198, Val203, and Met206 further stabilizing the complex. No water-mediated interactions were observed between $\mathbf{1}^+$ and HuPrP^C.

The free energy difference between the bound state of HuPrP^C−$\mathbf{1}^+$ (1^{1+}.B1 in Figure 3B) and the corresponding unbound state, with the corrections described above, turns out to be 8.6 kcal/mol. This is similar to that predicted for $\mathbf{1}^0$ and is in good agreement with the experimental data.

5. The HuPrP^C−$\mathbf{1}^{2+}$ Complex. Five different stable conformations of $\mathbf{1}^{2+}$ were identified on HuPrP^C surface (Figure 3C). In the most stable conformation, $\mathbf{1}^{2+}$ binds yet in another position of HuPrP^C, laying along helix H2 (1^{2+}.B1 in Figure 3C). Half of a part of $\mathbf{1}^{2+}$ is in close contact with the HuPrP^C surface in the cleft formed by Val189, Thr192, and Thr193. Indeed, these positions were reported to interact directly with the ligand according to NMR experiments (Supporting Information, Table S1). Two layers of water molecules are present between the protein surface and the rest of the molecule, presumably due to the presence of Lys185. In the other accessible conformations, $\mathbf{1}^{2+}$ covers different regions of the protein surface (Supporting Information, Figure S4). The interaction with Lys194 is conserved in the majority of them. This result is consistent with the chemical-shift changes of this residue upon ligand binding.[17] The dissociation free energy of $\mathbf{1}^{2+}$ was not calculated as, according to our calculations based on pK_a estimations, this protomer is not present at the conditions in which the K_d was measured.[17]

In summary, the EMD protocol enables to identify binding poses of the **1** protomers to the HuPrP^C surface

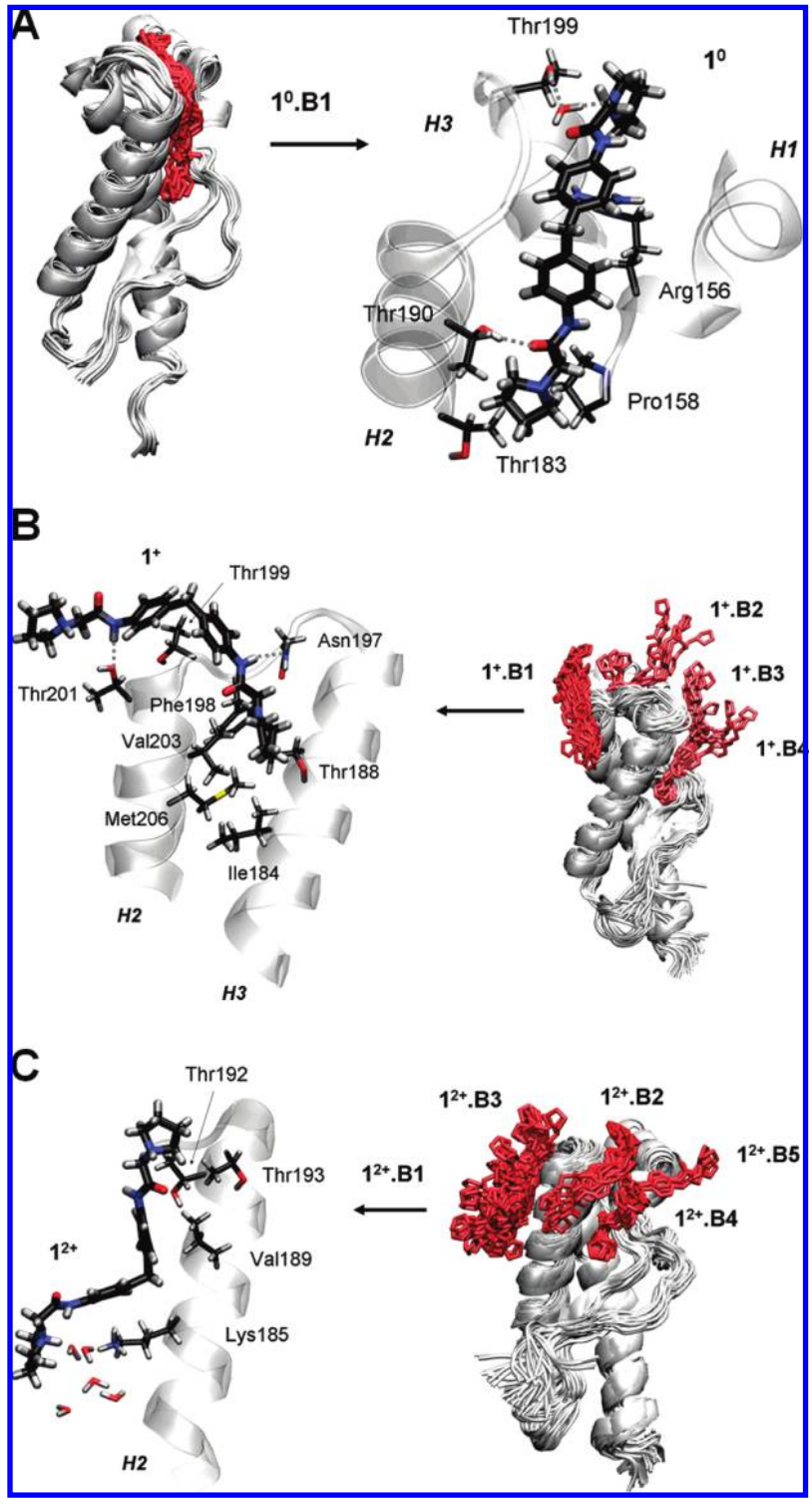


Figure 3. Three dimensional structures of HuPrP^C in complex with **1** protomers: (A) $\mathbf{1}^0$, (B) $\mathbf{1}^+$, and (C) $\mathbf{1}^{2+}$. These structures correspond to the bound-state free energy minima (1°.B1, 1^+.B1−B4, 1^{2+}.B1−B5), as calculated with the metadynamics method (see Methods Section). Close-ups on the ligands and the binding sites are also shown.

other than that proposed by Kuwata et al.[17] The multiple binding sites pattern that we observed from our simulations (Figure 4A) provides a structural basis for the NMR contacts (Figure 4B). The NMR contacts are observed in distant positions of HuPrP^C; in fact, they are located in opposite parts of the protein surface.[17] This result could only be obtained when we extended MDPs with enhanced sampling simulations. The predicted value of the K_d was in good agreement with experiments.[17]

Conclusions

The integration of docking algorithms with MD based simulations has been shown to be convenient in computer-assisted drug design.[2] We suggest here that its extension with

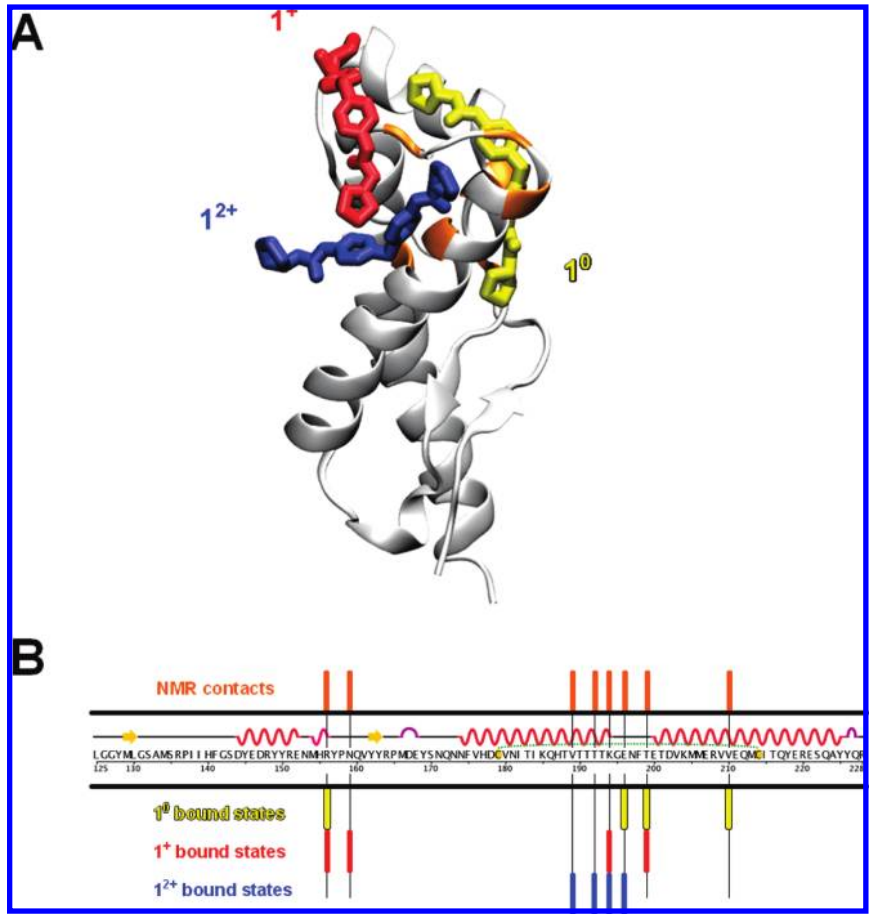


Figure 4. (A) Superimposition of the most populated binding poses of the three GN8 protomers: $\mathbf{1}^0$ (yellow), $\mathbf{1}^+$ (red), and $\mathbf{1}^{2+}$ (blue). (B) HuPrP^C sequence. The residues experimentally found to be involved in binding are highlighted in orange bars.[17] Those emerging from the calculations are shown with the same color code as (A).

free energy simulations, based on metadynamics, is useful to predict the adducts and the energetics of small ligands binding to cavity-less proteins. This approach enables the identification of the preferential ligand binding poses on protein surfaces with their corresponding binding affinity. This information is helpful for the subsequent improvement of the lead compounds in drug design. Docking simulations at protein surface sites are still in their infancy, and therefore, our study could provide a step towards the development of a computational protocol able to identify small organic molecules that interfere with protein−protein interactions occurring in the fibrillization process.

Here, we have focused on the ligand **1** (Figure 1) that targets the cellular form of the prion protein (HuPrP^C), the main agent involved in prion diseases.[17] Given the lack of deep binding pockets along the protein structure, it is reasonable to assume that a small molecule will not bind specifically to a single site. NMR data indicates that this is indeed the case.[17] This data shows that, for compound **1**, there are few hot spots far away from each other on the prion protein surface. This is a clear indication that **1** cannot bind to a unique position. Our study shows that compound **1** can adopt different protonation states at physiological and acid pH, at which experiments have been carried out.

Several poses of the species most present at an acidic pH, the diprotonated state ($\mathbf{1}^{2+}$), are found at the protein surface (Figure 3C). Similar binding sites are observed for $\mathbf{1}^+$ (Figure 3B), which is present in smaller content at an acidic pH. The neutral form ($\mathbf{1}^0$), which also may be present at an acidic pH, binds to a specific region of HuPrP^C surface along the shallow cleft formed by helices H1, H2, and H3 (Figure 3A).

Taken altogether, these multiple binding modes (Figure 4A) are consistent with all of the experimentally predicted contacts between the compound **1** and the different parts of the protein (Figure 4B).[17] The lack of a unique binding site is also coherent with the fact that different PrP^C binders, such as PrP^C antibodies,[32,33] molecular chaperones,[15,34] and DNA aptamers,[35] may interact with different regions of PrP^C surface. The multiple-site binding pattern that arises from our simulations might be an important facet for the anti-fibrillization potency of compound **1**, as it shows different points for disrupting the protein−protein interactions among HuPrP. This result could not be obtained by applying only standard MDP protocols (Scheme 1). In fact, it was necessary to simulate the whole binding process of compound **1** from the solution to the protein surface by means of enhanced sampling MD techniques like metadynamics. Most importantly, the predicted dissociation free energy turned out to be in very good agreement with the experimental data. Based on these encouraging results, the EMD protocol may now be used to predict the potency of ligands interacting with protein surfaces or target proteins without a unique binding

site. These include oligomers,[65] β-amyloid, as well as α-sinucleyn[66] and other unstructured proteins.

Methods

Identification of Binding Sites. The residues interacting with **1** are located in HuPrP^C C-term, for which the NMR structure is available (residues 125−228, PDB ID: 1HJM).[18] Protonation states were assigned by the web server H++[36] assuming pH 7.4. Putative binding sites were identified by (i) molecular simulations (using the GROMACS package[37] and (ii) docking procedure (using the GOLD[38,39] and the Autodock programs[40]).

1. Molecular Simulations. The protein was inserted into a cubic box of water molecules, ensuring that the solvent shell would extend for at least 0.8 nm around the system. Three sodium counterions were added. The AMBER99 force field[41,42] was used for the protein. Sodium ions were modeled with the AMBER-adapted Aqvist potential.[43] The water molecules were described by the TIP3P model.[44] The system was minimized imposing harmonic position restraints of 1000 kJ·mol^{-1}·nm^{-2} on solute atoms, allowing the equilibration of the solvent without distorting the solute structure. After an energy minimization of the solvent and the solute without harmonic restraints, the temperature was gradually increased from 0 to 298 K. This was performed by increasing the temperature from 0 to 298 K in 12 steps in which the temperature was increased by 25 K in 100 ps of MD.

Constant temperature−pressure (T = 298 K, P = 1 bar) 20-ns dynamics was then performed through the Nosé-Hoover[45,46] and Andersen-Parrinello-Rahman[47,48] coupling schemes. Periodic boundary conditions were applied. The final simulation box equilibrated at around 6.69 × 6.69 × 6.69 nm. Long-range electrostatic interactions were treated with the particle mesh Ewald (PME)[49,50] method, using a grid with a spacing of 0.12 nm combined with a fourth-order B-spline interpolation to compute the potential and forces in between grid points. The cutoff radius for the Lenard-Jones interactions as well as for the real part of PME calculations was set to 0.9 nm. The pair list was updated every 2 steps, and the LINCS algorithm[51] was used to constrain all bond lengths involving hydrogen atoms allowing us to use a time step of 2 fs.

The MD trajectory of prion protein alone was clustered with the gromos method[52] and as result 20 different conformations were obtained, which were used along with the NMR structure for docking of compound **1**.

2. Docking. Titration curves for compound **1** in bulk solution were calculated by the ChemAxon software[31] showing that this molecule is present in two protonation states at pH = 7.4: neutral (**1**^{0}) and monoprotonated (**1**$^+$); while at pH = 4.5 it exists mostly in the diprotonated form (**1**$^{2+}$) (Figure 1 and S2). This method has been used because it appears to be rather reliable: in a calculation of pK_a of 1000 molecules, less than 0.5% calculations turned out to differ by more than 0.5 pH unit from the experimental value.[31] Since the acidity of compound **1** is expected to change in the proximity of the protein, electrostatic potential calculations in implicit solvent[64] were addressed for the prion protein in two different conditions: pH = 4.5 for NMR measurements conditions and pH = 7 for affinity measurements conditions. The electrostatic potential surfaces were calculated using the APBS package[64] and visualized with VMD[68]. The calculation parameters were 0.3 Å grid spacing, 129^3 meshes, solvent and protein dielectrics of 78.54 and 40, respectively.

All three protomers underwent geometry optimization at the B3LYP/6-31G** level of theory by means of the Gaussian03 software (g03).[53]

The optimized structures, **1**^{0}, **1**$^+$ and **1**$^{2+}$, were docked to the NMR structure of HuPrP^C and to its 20 different conformations, as obtained after the cluster analysis of the MD trajectory.

The GOLD 3.1[38,39] and Autodock 3.0.5[40] programs were used. In GOLD, the docking area was defined as a sphere of 3.5 nm radius around the His187, so that the whole protein was screened. The ChemScore (CS)[54] and GoldScore (GS)[38] scoring functions were used for ranking. For each protomer and scoring function, 100 docking runs were performed.

In Autodock,[40] a Lamarckian genetic search algorithm was used to identify low energy binding sites and orientations of **1** protomers. Binding modes were ranked by a scoring function implemented in the Autodock. A point grid with a spacing 0.0475 nm was used. A point grid was centered to the center of mass of the protein, its dimensions were 12.6 × 12.6 × 12.6 nm. Gasteiger atom charges were assigned to the protein atoms using AutoDock tools. Water molecules were excluded from the protein before docking. One hundred randomly seeded runs were performed. The binding poses were identified by the AClAP 1.0 clustering procedure.[55]

3. **Hydration and Thermal Stability of 1−HuPrP^C Adducts.** A 10 ns MD simulation of the adducts (HuPrP^C−**1**^{0}, HuPrP^C−**1**$^+$ and HuPrP^C−**1**$^{2+}$) allowed for proper hydration of the system and identification of the collective motions that may be essential for PrP^C−ligand interactions. The protomers were bound to the binding region I (Figure 2B). The simulation protocol was the same as for the free protein. For the three ligands, the gaff force field[42,56] was used. The atomic restrained electrostatic potential (RESP) charges[57,58] were calculated by using the respectively module of AMBER after geometry optimization and electrostatic potential calculations of each protomer at the B3LYP/6-31G** level of theory by means of the g03 software.[53]

4. **Dissociation Free Energy Calculations.** The dissociation free energies of **1**^{0}, **1**$^+$, and **1**$^{2+}$ were calculated using metadynamics[25] in its bias-exchange variant[26] as a function of collective variables (CVs), which should be relevant for describing the dissociation process. CVs used in this work are: (i) the distance between the center of mass of the ligand and the protein binding region; (ii) the number of polar contacts between the ligand and one portion of the protein binding region I; (iii) the number of polar contacts between the ligand and the other portion of the protein binding region I; (iv) the number of water bridge contacts between the ligand and the protein binding region I; (v) the rmsd difference of the system with respect to an equilibrated MD structure taken from the previous section; and (vi) the rmsd fluctuation of the residues defining the protein binding region I. The choice of these CVs was based on previous ligand−target interaction metadyamics studies[27–30] as well as by observations based on the former MD simulations (see Supporting Information for more details). The calculations

do not require in principle the previous knowledge of the protein−ligand adduct structure. However, for computational efficiency we exploit the fact that all the target regions detected from NMR are in the close proximity of region I. Therefore, here we explored only this region.

Six independent metadynamics simulations were run in parallel. Each replica was biased by different one-dimensional time-dependent potentials, which were built as a function of each of the collective variables defined above. Exchanges among replicas were attempted every 10 ps using a metropolis acceptance criterion.[26] Similar setup was shown to improve the sampling of the configurational space and the convergence of the results.[26,59–63] At the end of the different replica simulations, the explored phase space, in terms of the six collective variables used in this study, was clustered using the gromos method.[52] The clustering radius for each collective variable was set to 0.1, 0.2, 0.4, 2.5, 0.05, and 0.02 nm, respectively. The free energy corresponding to each cluster was then reconstructed from the populations of clusters observed during the simulations. The free energy value was corrected by the corresponding bias potentials acting on that cluster as in a usual weighted histogram analysis.[62] Details on this procedure can be found in references,[59–63] and are summarized in Supporting Information together with the converged free energy profiles.

Two reference states of the ligand-protein system, bound and unbound, needed to be defined to provide the corresponding dissociation free energy value. The bound state was considered as the lowest free energy cluster. The unbound state was considered to be a cluster showing no contacts with the binding regions I (lowest values of CVs *ii* and *iii*) and at the same time with a higher rmsd with respect to the initial docked structure (highest value of CV *v*). Given the size of the simulation box, the ligand is not fully detached from the protein in its unbound state. Therefore, the residual dissociation energy of the unbound state was roughly estimated in implicit solvent using an adaptive Poisson−Boltzmann solver. The APBS package[64] was used with the same parameters described previously. It was estimated as the difference in solvation energy of the complex minus the solvation energy of each component plus the intermolecular Coulomb interaction. The standard free energy of dissociation was obtained by applying the following relationship: $\Delta G^0 = \Delta G - RT \ln([L])$, where ΔG is the total dissociation free energy as a result of our simulation, R is the molar constant, and $[L]$ is the concentration of the ligand in our simulation box (i.e., 5.5 mM, corresponding to one molecule in 6.69^3 nm^3). The standard free energy is related to the dissociation equilibrium constant (K_d) by $\Delta G^0 = -RT \ln(K_d)$.

Acknowledgment. The authors acknowledge A. Laio, F. Marinelli, and F. Pietrucci for fruitful discussions and for providing the necessary information to perform the analysis of the bias-exchange metadynamics simulations. X.B. acknowledges the financial support from the Government of Catalonia through a Beatriu de Pinós fellowship (BP-A 2007 http://ww.gencat.cat/agaur). P.C. acknowledges financial support from IIT (http://www.iit.it).

Supporting Information Available: Details of the MD simulations of compound **1** protomers in complex with HuPrP^C and additional information of the bias-exchange metadynamics procedure. The relevant experimental NMR data referenced in this work are summarized. Parameters of the biasing potential in the metadynamics calculations are provided. Furthermore, six supplementary figures, as referenced in the main text, are reported. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. *Nat. Rev. Drug. Discov.* **2004**, *3*, 935–49.

(2) Alonso, H.; Bliznyuk, A. A.; Gready, J. E. *Med. Res. Rev.* **2006**, *26*, 531–68.

(3) May, A.; Zacharias, M. *Biochim. Biophys. Acta* **2005**, *1754*, 225–31.

(4) Mangoni, M.; Roccatano, D.; Di Nola, A. *Proteins* **1999**, *35*, 153–62.

(5) Pak, Y.; Wang, S. *J. Phys. Chem. B* **2000**, *104*, 354–9.

(6) Lin, J. H.; Perryman, A. L.; Schames, J. R.; McCammon, J. A. *J. Am. Chem. Soc.* **2002**, *124*, 5632–3.

(7) McCammon, J. A. *Biochim. Biophys. Acta* **2005**, *1754*, 221–4.

(8) Carlson, H. A. *Curr. Opin. Chem. Biol.* **2002**, *6*, 447–52.

(9) Bucciantini, M.; Giannoni, E.; Chiti, F.; Baroni, F.; Formigli, L.; Zurdo, J.; Taddei, N.; Ramponi, G.; Dobson, C. M.; Stefani, M. *Nature* **2002**, *416*, 507–11.

(10) Cavalli, A.; Bolognesi, M. L.; Minarini, A.; Rosini, M.; Tumiatti, V.; Recanatini, M.; Melchiorre, C. *J. Med. Chem.* **2008**, *51*, 347–72.

(11) Soto, C. *Nat. Rev. Neurosci.* **2003**, *4*, 49–60.

(12) Prusiner, S. B. *Science* **1982**, *216*, 136–44.

(13) Caughey, B.; Baron, G. S. *Nature* **2006**, *443*, 803–10.

(14) Prusiner, S. B. *Trends Biochem. Sci.* **1996**, *21*, 482–7.

(15) Trevitt, C. R.; Collinge, J. *Brain* **2006**, *129*, 2241–65.

(16) Laurén, J.; Gimbel, D. A.; Nygaard, H. B.; Gilbert, J. W.; Strittmatter, S. M. *Nature* **2009**, *457*, 1128–1132.

(17) Kuwata, K.; Nishida, N.; Matsumoto, T.; Kamatari, Y. O.; Hosokawa-Muto, J.; Kodama, K.; Nakamura, H. K.; Kimura, K.; Kawasaki, M.; Takakura, Y.; Shirabe, S.; Takata, J.; Kataoka, Y.; Katamine, S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11921–6.

(18) Calzolai, L.; Zahn, R. *J. Biol. Chem.* **2003**, *278*, 35592–6.

(19) Riek, R.; Hornemann, S.; Wider, G.; Billeter, M.; Glockshuber, R.; Wuthrich, K. *Nature* **1996**, *382*, 180–2.

(20) Ishikawa, T.; Ishikura, T.; Kuwata, K. *J Comput Chem* 2009 [Online early access]. DOI: 10.1002/jcc.21265. Published Online: Apr 30, 2009. http://www3.interscience.wiley.com/cgi-bin/fulltext/122370880/HTMLSTART. Accessed June 28, 2009.

(21) Vogtherr, M.; Grimme, S.; Elshorst, B.; Jacobs, D. M.; Fiebig, K.; Griesinger, C.; Zahn, R. *J. Med. Chem.* **2003**, *46*, 3563–4.

(22) Kirby, L.; Birkett, C. R.; Rudyk, H.; Gilbert, I. H.; Hope, J. *J. Gen. Virol.* **2003**, *84*, 1013–20.

(23) Gilson, M. K.; Zhou, H. X. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21–42.

(24) Rodinger, T.; Pomes, R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 164–70.

(25) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562–6.

(26) Piana, S.; Laio, A. *J. Phys. Chem. B* **2007**, *111*, 4553–9.

(27) Branduardi, D.; Gervasio, F. L.; Cavalli, A.; Recanatini, M.; Parrinello, M. *J. Am. Chem. Soc.* **2005**, *127*, 9147–55.

(28) Fiorin, G.; Pastore, A.; Carloni, P.; Parrinello, M. *Biophys. J.* **2006**, *91*, 2768–2777.

(29) Gervasio, F. L.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2005**, *127*, 2600–2607.

(30) Vargiu, A. V.; Ruggerone, P.; Magistrato, A.; Carloni, P. *Nucleic Acids Res.* **2008**, *36*, 5910–21.

(31) Calculator Plugins were used for structure property prediction and calculation. *Marvin 5.0.0*; ChemAxon: Budapest, Hungary, 2008; http://www.chemaxon.com; (accessed May 29, 2008).

(32) Campana, V.; Zentilin, L.; Mirabile, I.; Kranjc, A.; Casanova, P.; Giacca, M.; Prusiner, S. B.; Legname, G.; Zurzolo, C. *Biochem. J.* **2009**, *418*, 507–15.

(33) Williamson, R. A.; Peretz, D.; Pinilla, C.; Ball, H.; Bastidas, R. B.; Rozenshteyn, R.; Houghten, R. A.; Prusiner, S. B.; Burton, D. R. *J. Virol.* **1998**, *72*, 9413–8.

(34) Kaneko, K.; Zulianello, L.; Scott, M.; Cooper, C. M.; Wallace, A. C.; James, T. L.; Cohen, F. E.; Prusiner, S. B. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10069–74.

(35) Takemura, K.; Wang, P.; Vorberg, I.; Surewicz, W.; Priola, S. A.; Kanthasamy, A.; Pottathil, R.; Chen, S. G.; Sreevatsan, S. *Exp. Biol. Med.* **2006**, *231*, 204–214.

(36) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. *Nucleic Acids Res.* **2005**, *33*, W368–71.

(37) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43–56.

(38) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727–48.

(39) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. *Proteins* **2003**, *52*, 609–23.

(40) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639–62.

(41) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668–88.

(42) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2005**, *26*, 114–114.

(43) Aaqvist, J. *J. Phys. Chem.* **1990**, *94*, 8021–8024.

(44) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

(45) Hoover, W. G. *Phys. Rev. A:At., Mol., Opt. Phys.* **1985**, *31*, 1695–1697.

(46) Nose, S. *Mol. Phys.* **1984**, *52*, 255–268.

(47) Nosé, S.; Klein, M. L. *Mol. Phys.* **1983**, *50*, 1055–76.

(48) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(49) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–92.

(50) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(51) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

(52) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angew. Chem., Int. Ed.* **1999**, *38*, 236–40.

(53) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc.: Wallingford, CT, 2004.

(54) Nissink, J. W.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C.; Taylor, R. *Proteins* **2002**, *49*, 457–71.

(55) Bottegoni, G.; Rocchia, W.; Recanatini, M.; Cavalli, A. *Bioinformatics* **2006**, *22*, e58−65.

(56) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049–74.

(57) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(58) Cornell, W.; Cieplak, P.; Bayly, C. I.; Kollmann, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620–9631.

(59) Leone, V.; Lattanzi, G.; Molteni, C.; Carloni, P. *PLoS Comput. Biol.* **2009**, *5*, e1000309.

(60) Piana, S.; Laio, A.; Marinelli, F.; Van Troys, M.; Bourry, D.; Ampe, C.; Martins, J. C. *J. Mol. Biol.* **2008**, *375*, 460–70.

(61) Todorova, N.; Marinelli, F.; Piana, S.; Yarovsky, I. *J. Phys. Chem. B* **2009**, *113*, 3556–3564.

(62) Marinelli, F.; Pietrucci, F.; Laio, A.; Piana, S. *PloS Comp. Biol.* **2009**, *5*, e1000452.

(63) Pietrucci, F.; Marinelli, F.; Carloni, P.; Laio, A. *J. Am. Chem. Soc.* **2009**; in press. Epub ahead of print, available at http://pubs.acs.org/doi/abs/10.1021/ja903045y.

(64) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037–41.

(65) Esteras-Chopo, A.; Morra, G.; Moroni, E.; Serrano, L.; Lopez de la Paz, M.; Colombo, G. *J. Mol. Biol.* **2008**, *383*, 266–280.

(66) Herrera, F. E.; Chesi, A.; Paleologou, K. E.; Schmid, A.; Munoz, A.; Vendruscolo, M.; Gustincich, S.; Lashuel, H. A.; Carloni, P. *PLoS One* **2008**, *3*, e3394.

(67) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14.1*, 33–38.

CT900257T