

Role of Urea–Aromatic Stacking Interactions in Stabilizing the Aromatic Residues of the Protein in Urea-Induced Denatured State

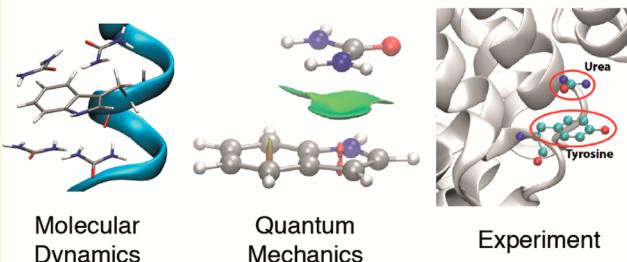
Siddharth Goyal,[†] Aditya Chattopadhyay,[†] Koushik Kasavajhala, and U. Deva Priyakumar^{*ID}

Center for Computational Natural Sciences and Bioinformatics, International Institute of Information Technology, Hyderabad 500 032, India

Supporting Information

ABSTRACT: A delicate balance of different types of intramolecular interactions makes the folded states of proteins marginally more stable than the unfolded states. Experiments use thermal, chemical, or mechanical stress to perturb the folding equilibrium for examining protein stability and the protein folding process. Elucidation of the mechanism by which chemical denaturants unfold proteins is crucial; this study explores the nature of urea–aromatic interactions relevant in urea-assisted protein denaturation. Free energy profiles corresponding to the unfolding of Trp-cage mini-protein in the presence and absence of urea at three different temperatures demonstrate the distortion of the hydrophobic core to be a crucial step. Exposure of the Trp6 residue to the solvent is found to be favored in the presence of urea. Previous experiments showed that urea has a high affinity for aromatic groups of proteins. We show here that this is due to the remarkable ability of urea to form stacking and NH–π interactions with aromatic groups of proteins. Urea–nucleobase stacking interactions have been shown to be crucial in urea-assisted RNA unfolding. Examination of these interactions using microsecond-long unrestrained simulations shows that urea–aromatic stacking interactions are stabilizing and long lasting. Further MD simulations, thermodynamic integration, and quantum mechanical calculations on aromatic model systems reveal that such interactions are possible for all the aromatic amino acid side-chains. Finally, we validate the ubiquitous nature of urea–aromatic stacking interactions by analyzing experimental structures of urea transporters and proteins crystallized in the presence of urea or urea derivatives.

Aromatic Amino Acid - Urea Stacking



Molecular Dynamics

Quantum Mechanics

Experiment

INTRODUCTION

Proteins exist in their functionally active folded state under ambient conditions. In the equilibrium between the folded and unfolded states, a combination of different types of intramolecular interactions makes the folded state of proteins moderately more stable than the unfolded states.^{1–5} Experiments that perturb normal environmental conditions, such as temperature, pressure, pH, and addition of cosolvents, have provided significant insights on the stability of proteins and protein folding pathways.^{6–11} Certain small molecules referred to as cosolvents are capable of shifting the protein folding equilibrium toward either of these states.^{6,7,12,13} Urea is one such cosolvent that is known to favor the unfolded state, a process known as chemical denaturation. Urea-assisted denaturation of proteins has been extensively employed to study protein folding both experimentally and computationally.^{14–17} Though similar to water in many aspects, urea has been shown to favorably interact with both polar and nonpolar groups, which enables protein denaturation in the presence of urea. The molecular mechanism by which urea assists denaturation of proteins is an area of intense research. Two different interaction models to explain this phenomenon have been proposed: indirect and direct mechanisms. The indirect mechanism suggests that urea disrupts

the structure of water, leading to weakening of the hydrophobic effect. Such an effect helps the solvation of hydrophobic groups of the amino acids in aqueous solution.^{18–22} On the other hand, the direct mechanism suggests that direct interactions between urea and the different functional groups of the protein compete with the native interactions present in the native structure, leading to denaturation. Now, it is generally accepted that the urea-assisted protein denaturation proceeds via the direct mechanism.

Studies reported in favor of the direct mechanism do not agree over the exact energetics and the selectivity demonstrated by urea. Few results posit that the interaction of urea with protein is primarily electrostatic in nature and related to formation of hydrogen bonds,^{23–27} whereas others suggest that preferential dispersion interaction is the main reason.^{28–31} Regarding selectivity, some investigations conclude interactions with the protein backbone as the major destabilizing effect,^{23,32} while others suggest that urea interacts preferentially with the amino acid side-chains.^{24,25,31,33,34} In the latter case, perspectives are further divided between those who favor the interaction with

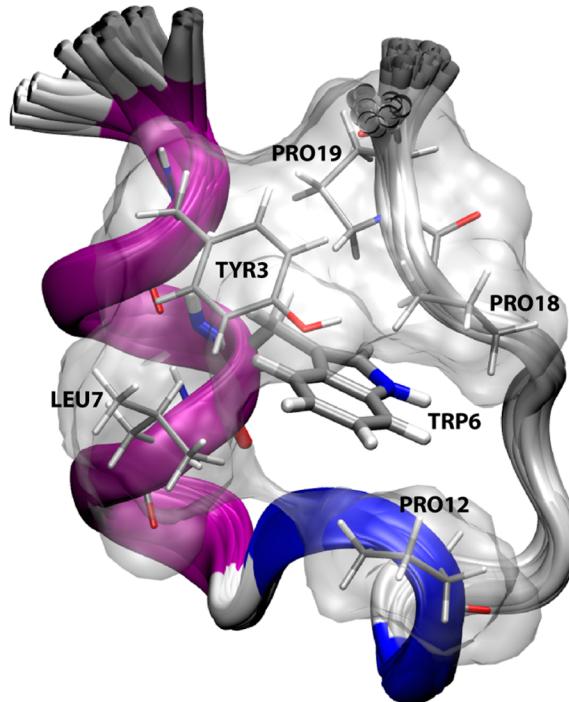
Received: May 26, 2017

Published: October 4, 2017

polar and charged²⁵ side-chains, those who emphasize interactions with apolar side-chains,^{29,31,33–35} and those who underline interactions with both polar and hydrophobic residues, as the major interactions.³⁶ Electrostatic interactions including the hydrogen-bonded interactions themselves are not capable of driving the denaturation process. Dispersion-type interactions involving urea are largely responsible for stabilizing the protein in their unfolded states.^{28,37}

Theoretical studies, specifically molecular dynamics (MD) simulations, have been instrumental in understanding the atomistic mechanism involved in the urea-assisted denaturation of proteins. Studies involving diverse model systems have been reported.^{28,31,38} For example, MD simulations have been done on small hydrophobic molecules and on hydrophobic polymers in the presence and absence of urea.^{25,39–42} The presence of urea minimizes the compactness of polymers and aggregation of small hydrophobic molecules that signify the role of dispersion-type interactions involving urea in stabilizing the unfolded states. Several studies involving MD simulations starting from the native folded states of proteins in the presence of aqueous urea have been reported. Experimental and computational studies on small model systems have enabled the understanding of thermodynamics of how each of the amino acid side-chains and the protein backbone interact with aqueous urea.^{31,43–46} The change in the accessible surface area (ΔASA) of a unified atom model when a protein unfolds due to addition of urea has been used to quantitatively understand the contribution of the interactions between urea and different chemical groups to the denaturation process.^{44–48} Thermodynamic and kinetic *m*-values calculated from the ΔASA values thus obtained have helped in the understanding of the protein denaturation mechanism in general. Guinn et al. have used osmometry to experimentally obtain the preferential interaction coefficients of urea with respect to different functional groups of the protein amino acids.^{44–46} They have found that urea prefers to interact with the aromatic carbon atoms followed by amide oxygen atoms compared to other functional groups in proteins. However, atomistic details of the nature of interactions that lead to such a remarkable phenomenon of interaction between urea and aromatic groups are not known. Though NH–π interactions have been suggested to be the mode of interaction,¹⁶ OH–π interactions involving water and aromatic groups are equally preferable, and hence this does not explain why there is a preferential interaction when it comes to water vs urea. A recent study by Cheng et al. examined the amide–aromatic interactions using osmometry and solubility and showed that amide–O interaction is more favorable than amide–N interaction with naphthalene.⁴⁹

Trp-cage miniprotein (Figure 1) is one of the smallest structures that have a protein-like fold and has been used as a prototypical model system in several protein folding studies both experimentally^{50–57} and theoretically.^{58–71} The Trp-cage miniprotein construct TCSb (PDB ID: 1L2Y) is a 20-residue globular protein designed by Neidigh et al. having a melting temperature of 315 K.^{72,73} The native Trp-cage miniprotein includes an α -helix from residues 2–8, a 3_{10} -helix from residues 11–14, and a polyproline II helix from residues 17–19. The tryptophan residue (Trp6), along with Tyr3, Leu7, Gly11, Pro12, Pro18, and Pro19, forms the hydrophobic core of the protein. The protein is further stabilized by the presence of a salt bridge between residues Asp9 and Arg16 and a hydrogen bond between H ϵ 1 of the Trp6 residue and the backbone carbonyl (C=O) of the Arg16 residue. It has been shown that the hydrophobic core



Trp-Cage Mini-Protein: NLYIQ WLKDGPSSGRPPPS

Figure 1. Structure of the Trp-cage miniprotein depicting the Trp6 residue buried in a hydrophobic cage and its primary sequence. Surface representation of select residues' part of the hydrophobic core is given (transparent gray). The backbone traces of all the conformations obtained using NMR (PDB ID: 1L2Y) showing the flexibility of the terminal residues are also shown.

with the central Trp6 is crucial for maintaining the integrity of the structure.

In the present study, an attempt is made to do a comprehensive investigation using several theoretical methodologies such as MD simulations, umbrella sampling calculations, thermodynamic integration free energy calculations, high-level quantum mechanical calculations, and database analysis of protein structures to (a) demonstrate the ease of the distortion of the hydrophobic core in the presence of urea, (b) show that aqueous urea optimally solvates aromatic groups, (c) identify novel interactions, namely, stacking and NH–π interactions involving aromatic groups and urea, (d) characterize the dynamics of such interactions, and (e) provide experimental evidence of these effects.

METHODOLOGY

All the MD simulations and free energy calculations were performed using the CHARMM⁷⁴ and NAMD⁷⁵ programs using the CHARMM36 protein force field⁷⁶ and the CHARMM general force field (CGenFF) for urea.⁷⁷ Though other widely used force fields such as the Kirkwood–Buff derived force field (KBFF)⁷⁸ are available for urea, the CHARMM force field was used so that it is compatible with the protein force field used here. The radial distribution functions calculated using the KBFF and CHARMM force fields are in near-quantitative agreement with each other (Figure S1 in the Supporting Information). Additionally, transfer free energies obtained using this force field agree well with experimental data (see later). Analysis of the MD trajectories was done using the CHARMM program and Python scripts developed in our lab; structural representations given here were rendered using the VMD program.⁷⁹ All the quantum mechanical calculations were performed using the GAMESS program.⁸⁰

Umbrella Sampling MD Simulations. The initial structure of the Trp-cage miniprotein was obtained from a previous NMR study (PDB ID: 1L2Y).⁷² The protein was adequately solvated using the TIP3P water model, energy minimized, and equilibrated with harmonic constraints on the non-hydrogen atoms for 10 ns. This was followed by a 50 ns long MD simulation without any restraints on the system other than constraining the covalent bonds involving hydrogen atoms. The resultant configuration at the end of this simulation was subjected to steered molecular dynamics (SMD) simulations for generating the initial structures for each of the windows for umbrella sampling simulations.⁸¹ The reaction coordinate was chosen to be the distance between the $\text{C}\alpha$ atoms of residues 2 and 19. The two terminal residues were not considered for defining the reaction coordinate because of the flexibility of these residues even in the native fold (Figure 1).⁷² Constant velocity SMD (0.01 Å/ps using 1 fs integration time step) was performed along the $\text{C}\alpha(\text{Leu}2)-\text{C}\alpha(\text{Pro}19)$ vector. From the SMD trajectory, structures corresponding to the reaction coordinate values varying between 3 and 42 Å with an interval of 3 Å between successive windows were chosen for the umbrella sampling MD simulations. These configurations of the protein were then resolved in each of the six pre-equilibrated solvent environments (water and 8 M aqueous urea at 300, 400, and 500 K). Previous studies have suggested that combination of two or more perturbations (chemical, thermal, and mechanical) helps examine different aspects of the protein folding problem that may not be accessible individually.^{48,82–86} The size of the solvation box for all the umbrella sampling simulations ($74 \times 74 \times 74 \text{ Å}^3$) was chosen based on the initial configuration of the last window so that the box extended at least 12 Å on all sides. Each of the windows was then subjected to short energy minimization and equilibration followed by 50 ns MD simulations, resulting in a cumulative sampling time of about 4.2 μs . During all these simulations, the SHAKE algorithm was used to constrain covalent bonds involving hydrogen atoms,⁸⁷ the nonbonded interactions were truncated by using a switching function between 10 and 13 Å, and the long-range electrostatic interactions were modeled using the particle mesh Ewald method. The umbrella sampling protocol used a harmonic potential with a force constant value of 1 kcal/mol/Å² to restrain the protein along the reaction coordinate. Free energy profiles corresponding to the unfolding along the distance between the second and penultimate residues of Trp-cage miniprotein were obtained using the weighted histogram analysis method (WHAM).^{88–90}

Unrestrained MD Simulations of the Trp-Cage Miniprotein at 300 K. Two configurations obtained from the umbrella sampling MD simulations of the last window in two environments (water and aqueous urea at 300 K) were further considered for unrestrained simulations to study the dynamic properties of urea–aromatic interactions in detail. Each of these two systems were subjected to 1 μs long MD simulations using the same protocol as above after removing the harmonic restraint along the reaction coordinate. This allowed the protein to access partially folded states starting from the completely unfolded state. A comparison of the propensities of the protein to sample the folded state and the nature of protein–solvent interactions in water and 8 M aqueous urea was performed by analyzing various geometric, energetic, and dynamic properties, details of which are presented in the *Results and Discussion* section.

MD Simulations and Thermodynamic Integration Calculations on Aromatic Model Systems. Analysis of the umbrella sampling MD trajectories and 1 μs long unrestrained MD trajectories revealed stacking and NH–π interactions between Trp6 and urea. A systematic study of five different model systems (viz., benzene, phenol, indole, imidazole, and imidazolium) was undertaken to study such interactions between all aromatic amino acids and urea. Each of these model systems were initially simulated in the presence of 0, 1, 2, 3, 4, 5, 6, 7, and 8 M aqueous urea for 20 ns after initial minimization and equilibration (cumulative production simulation time: 900 ns). This was followed by calculation of solvation free energies in 0, 4, and 8 M systems using the thermodynamic integration method. The coupling parameter was changed from 0 to 1 (vacuum to fully solvated system) at steps of 0.01. Each of these hundred windows of every system were simulated for 600 ps until the transfer free energies for the forward and reverse

processes (solvation and desolvation) were consistent with each other (cumulative simulation time: 1.8 μs).

Quantum Mechanical Calculations. Using the MD simulations and free energy calculations, it is shown that all the aromatic residues are capable of forming stacking and NH–π interactions with urea. These novel interactions were further validated by using high-level quantum mechanical calculations at the RI-MP2 level using the aug-cc-pVDZ basis set. Hundreds of putative binary complexes formed by each of the aromatic model systems and one urea molecule at different orientations corresponding to both stacking and NH–π interactions were conceived. Interaction energy calculations were calculated and were corrected for the basis set superposition error using the counterpoise correction method, and the dispersion contribution to the total interaction energies was assessed using the energy decomposition method given by Su and Li.⁹¹ Interaction energies corresponding to the most stable complexes thus obtained were also calculated at the RI-MP2/cc-pVQZ method. The two types of noncovalent interactions identified between the aromatic model systems and urea were further examined using NCIPLOT.^{92,93} The noncovalent interaction index is calculated from the electron density of the binary complex and its derivatives.

Database Analysis. Protein structures with bound urea and derivatives of urea from the Protein Data Bank were analyzed for possible occurrence of urea–aromatic stacking interactions. All structures that contain either urea or derivatives of urea were obtained from the PDB, and Python script is used to identify stacked orientations based on the below algorithm. An interatomic distance matrix was constructed for all nitrogen, carbon, and oxygen atoms that belong to the “HETATM” residues in the PDB files. Those sets of moieties where the distances between two nitrogens and a carbon atom are less than 1.6 Å, the distance between this carbon and an oxygen atom is less than 1.4 Å, and the N–N and N–O distances involving these four atoms are less than 3 Å were identified as urea or urea derivatives. The best fit plane corresponding to these four atoms and the best fit plane of the aromatic group whose atom(s) is present closer than 4.5 Å of one of the four atoms of the urea moiety were obtained. Based on the geometric criteria explained later, stacked configurations were identified.

RESULTS AND DISCUSSION

The free energy profiles corresponding to the unfolding of Trp-cage miniprotein obtained using the umbrella sampling MD simulations at six different conditions and the role of distortion of the hydrophobic core are presented first followed by the discussion of the role of hitherto unknown novel interactions between urea and aromatic residues, namely, stacking and NH–π interactions, in stabilizing aromatic residues in unfolded states. Dynamic properties such as lifetimes of these modes of binding are presented next based on 1 μs long unrestrained MD simulations followed by free energy calculations and quantum chemical calculations on model systems to further examine these interactions. Finally, experimental evidence for urea–aromatic stacking interactions and possible roles of these interactions in biological systems are discussed.

Thermal and Chemical Perturbations Facilitate Unfolding of the Trp-Cage Miniprotein. The potential of mean force profiles for the unfolding of the Trp-cage miniprotein obtained at six different conditions (in the presence and absence of urea at 300, 400, and 500 K) are given in Figure 2. Adequate overlap of the probability distributions of the reaction coordinates in the umbrella sampling simulations and convergence of the free energy surfaces were ensured (Figures S2 to S7 in the *Supporting Information*). The structure of the protein resembles the native state at about 6 Å, corresponding to the reaction coordinate except when the temperature is 500 K. The native state is characterized as a free energy minimum at 300 and 400 K in the presence or absence of urea, and the qualitative and quantitative features of the free energy surfaces seem to greatly

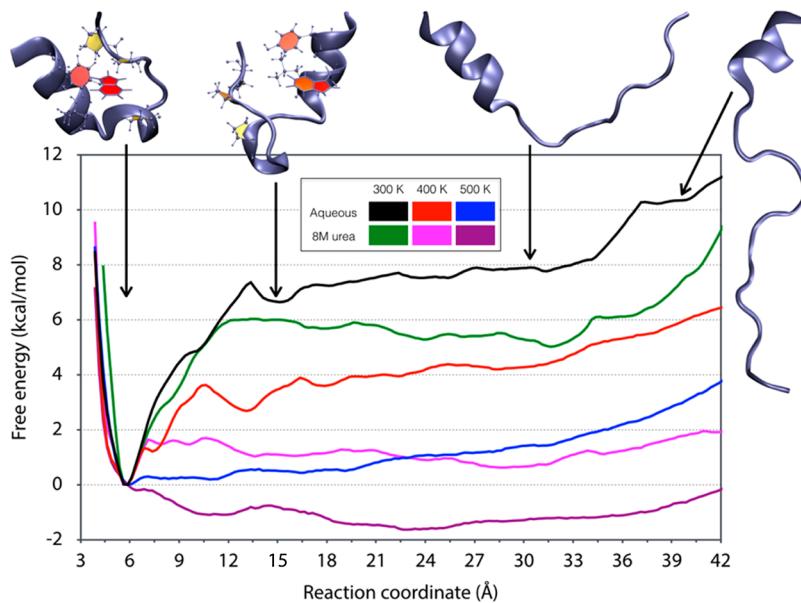


Figure 2. Potential of mean force profiles as a function of the distance between the C α atoms of the penultimate residues of the Trp-cage miniprotein obtained at the six different environments considered here. Representative structures of the protein along the reaction coordinate are also given.

depend on both the temperature and the presence/absence of the chemical denaturant. With respect to all the values of the reaction coordinate including at around 6 Å, the structures sampled at 500 K resemble unfolded states. The free energy profiles for the Trp-cage miniprotein unfolding at 300 and 400 K can be divided into three distinct regions corresponding to a hierarchical pathway (representative structures are depicted in Figure 2): (i) loss of the hydrophobic core along with the distortion of salt bridge between Arg16 and Asp9, (ii) separation of the α -helix and the PPII helix with respect to the loop, and (iii) partial unwinding of the α -helix. While the order of these events is very similar in these four conditions, the free energy corresponding to these transitions and the values of the reaction coordinate at which these transitions occur are distinctly different from each other. In aqueous solution at 300 K, the initial step involving the disruption of the hydrophobic core happens until ~13 Å along the reaction coordinate and is associated with a sharp rise in the free energy. The value of the reaction coordinate corresponding to complete opening of the Trp6 residue that is central to the hydrophobic core decreases with the increase in the severity of the conditions (~11, ~10.5, and ~8 Å in 300 K/aqueous urea, 400 K/aqueous, and 400 K/aqueous urea, respectively). Similarly, the value of the free energy at this point decreases from about 7.5 kcal/mol for the 300 K/aqueous condition to only about 1.5 kcal/mol for the 400 K/aqueous urea condition following the same trend. Such a sharp increase in the free energy due to the opening of Trp6 is followed by a rather flat surface, which corresponds to the separation of the two helices with respect to the loop. No significant change in the free energy is observed during this phase of unfolding because no significant native contacts are deformed. Finally, the α -helix in the N-terminal end of the protein tends to partially unwind, which leads to an increase in the free energy.

The reaction coordinate that is used in this study was inspired by the atomic force microscope experiments, which examine protein stability and folding using end-to-end distance as the reaction coordinate. Although it is expected that the free energy profiles and unfolding pathways are biased by the reaction coordinate used, comparison of the mechanism obtained here

with previous studies on Trp-cage miniprotein folding phenomena in the presence and absence of urea revealed good agreement (see below). Replica exchange and long unrestrained MD simulations have been used to study the dynamics of Trp-cage miniprotein in the presence and absence of urea, and free energy surfaces have been calculated by using different order parameters such as root-mean-square deviation, radius of gyration, fraction of native contacts, distance between Arg16 and Asp9, and solvation of Trp6.^{61,64,65} As discussed below, all studies suggest formation of the hydrophobic core along with the salt bridge between Arg16 and Asp9 follows the organization of the α -helix in the folding of this protein as observed in this study. Levy and co-workers observed a very similar folding mechanism, where they found that the dominant pathway of Trp-cage folding involved extended protein conformations with a preformed α -helix in the unfolded state, and folding progressed by the formation of a hydrophobic core and 3–10 helix.⁶⁷ Replica exchange simulations of Paschek et al. also suggested a similar mechanism wherein the formation of an α -helix was the initial step in the folding of the Trp-cage.⁶² Our results suggest that the unfolded states have a significant α -helical structure present in them due to only partial unwinding. This observation is consistent with the experimental results of Ahmed et al., who found an α -helical structure in the denatured states of the Trp-cage.⁵² A similar finding that the unfolded state of the Trp-cage preserves some helical content at the N-terminal region was reported by the NMR study of Trp-cage unfolding by Rovo et al.⁵⁷ This finding is in agreement with the recent computational studies of Shao et al., who suggested that the unfolded states have a high portion of α -helical structure.⁶⁶ Meuzelaar et al. suggested that formation of the α -helix preceded the formation of the hydrophobic core.⁷⁰ Interestingly, even in the presence of 8 M urea, there is substantial α -helical content present in the unfolded states. This is consistent with the results of Eberini and co-workers, who reported that urea preferentially caused loss of secondary structure in protein L by disrupting β -sheets, while the α -helical structure is largely preserved.⁹⁴ This observation is also consistent with the study by Heyda et al., who suggested that during the denaturation of the Trp-cage in urea and in guanidium

chloride the destruction of the α -helix takes place at a later stage, thus indicating that during folding the α -helix might form earlier than the hydrophobic core.⁶⁵

Preceding discussions on the free energy profiles and previous results from various groups strongly support that one of the key events for the unfolding of this protein is the distortion of the hydrophobic core. While a decrease in the free energy corresponding to unfolding due to thermal perturbations is understandable, how urea further favors this process consistently at both temperatures is examined in the following section.

Facile Distortion of the Hydrophobic Core in the Presence of Urea. Potential of mean force profiles indicate that distortion of the hydrophobic cage is associated with a sharp rise in the free energy. The solvent-accessible surface area (SASA) of the side-chain of the Trp6 residue, which is central to the hydrophobic core and buried in the folded state, was used to quantify the extent of distortion of the hydrophobic core. Mean values of the SASA value corresponding to a select range of the reaction coordinate are plotted in Figure 3, and data for the

cutoff value for the SASA of approximately 60 \AA^2 distinguishes the buried from the open states of Trp6. Notably, Guinn et al. have shown that aromatic carbons exhibit one of the strongest preferential interactions with urea.^{44–46} Thermodynamic *m*-values were calculated for the Trp6 side-chain, amide oxygen atoms, and aliphatic carbon atoms of the Trp-cage miniprotein in the folded and completely unfolded states using the ΔASA values using a procedure similar to that used before.⁴⁵ A comparison of predicted values of *m*-value/ RT per 1000 \AA^2 of surface (-0.099 , -0.029 , and -0.026 for Trp6 side-chain, amide oxygen atoms, and aliphatic carbon atoms, respectively) suggests that the exposure of the Trp6 residue may be the key step in the unfolding of the Trp-cage miniprotein. This is consistent with the above observation on the correlation between the change in the SASA value of the Trp6 side-chain along the reaction coordinate and the nature of the free energy profiles. For example, a minor change in the SASA values from 6 to 10 \AA in the presence and absence of urea at 300 K is associated with similar free energy profiles in this range. It is only after the Trp6 residue becomes exposed to the solvent that the stabilization effect of urea comes into play. This can be seen in Figure 2, where the free energy profile in the presence of urea at 300 K reaches a stable plateau after 10 \AA . The nature of the nonbonded interactions that are responsible for strong preference is investigated in detail in later sections.

van der Waals (vdW) Interactions Stabilize the Unfolded State in the Presence of Urea. Stabilization of the unfolded state of the protein due to the presence of urea was examined by calculating the interaction energies between all the residues of the protein and the solvent environment. A representative window (24 \AA) from the umbrella sampling simulations corresponding to unfolded states was chosen for this analysis. The probability distributions of the total interaction energies, the electrostatic component, and the Lennard-Jones (LJ) components along with the Gaussian fits are depicted in Figure 4. Data corresponding to the 300 K simulations in the presence and absence of urea are presented, and those obtained at 400 K are given in Figure S9 in the Supporting Information. The interaction energy between the protein and the environment becomes more favorable by about 83 kcal/mol when urea is introduced in the system. The total interaction energies were decomposed into the electrostatic and LJ contributions, which reveals that most of the stabilization of the unfolded state is contributed by the LJ-type dispersion interactions (76 kcal/mol) compared to the electrostatic contribution of only 7 kcal/mol . Although the magnitude of the total electrostatic energy is an order of magnitude higher than the total LJ energy, the differences in these energies between the two systems (in the presence and absence of urea) indicate dominance of dispersion-type interactions in stabilizing the unfolded state in the presence of urea at both 300 and 400 K . This is consistent with previous studies that showed the role of vdW interactions in stabilizing the unfolded states by aqueous urea.^{16,28} These results further substantiate the direct mechanism of the urea-assisted denaturation of proteins. Further sections examine the role of urea–aromatic interactions in the stabilization of aromatic residues in the solvent-exposed states. Note that previous studies have found that aqueous urea stabilizes all but negatively charged amino acids compared to water.³¹ Since the interactions involving aromatic groups are stronger than others and the nature of such interactions are not known, a detailed examination is undertaken here.

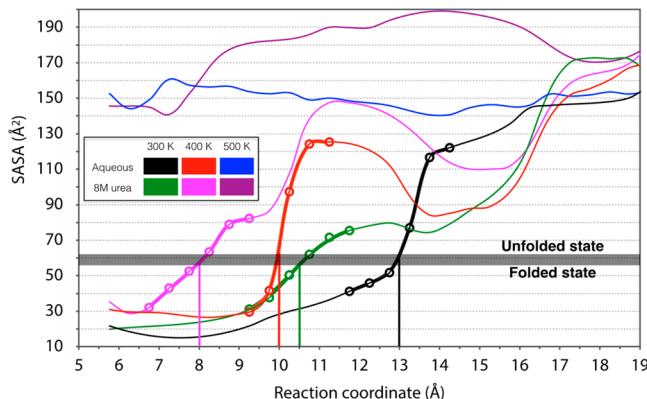


Figure 3. Mean values of the solvent-accessible surface area (SASA) of the Trp6 residue with respect to a select range of the reaction coordinate obtained at six different environments. The regions corresponding to the opening of the Trp6 residue are highlighted at 300 and 400 K .

whole range of the reaction coordinate are provided in the Supporting Information (Figure S8). A buried Trp6 side-chain is characterized by an SASA value of approximately 30 \AA^2 or below, whereas the exposed side-chain exhibits a much larger value. Consistent with discussions above, the SASA values of this residue calculated at 500 K in the presence and absence of urea indicate that it is exposed to the solvent for all values of the reaction coordinate. Transition of the residue from the buried to open state in the other four environments reveals interesting trends. In the case of the aqueous environment, this process shows a sigmoidal behavior with a sharp rise in the SASA beyond a certain value of the reaction coordinate at both 300 and 400 K (black and red thick lines in Figure 3). In the presence of urea, this transition is affected in three ways (green and magenta thick lines in Figure 3): (i) the transitions occur much earlier, (ii) the change in the SASA values are much more gradual, and (iii) the transition corresponds to a lower free energy. In the case of an aqueous solution, the Trp residue tends to be buried in the hydrophobic cage until it abruptly snaps out of the cage, effecting a sharp increase in the SASA value, but in the presence of urea, the residue slowly moves out of the hydrophobic core at both 300 and 400 K . The values of the reaction coordinates at which the Trp residue opens up coincides with those at which the free energies cease to increase sharply in all four conditions, and a

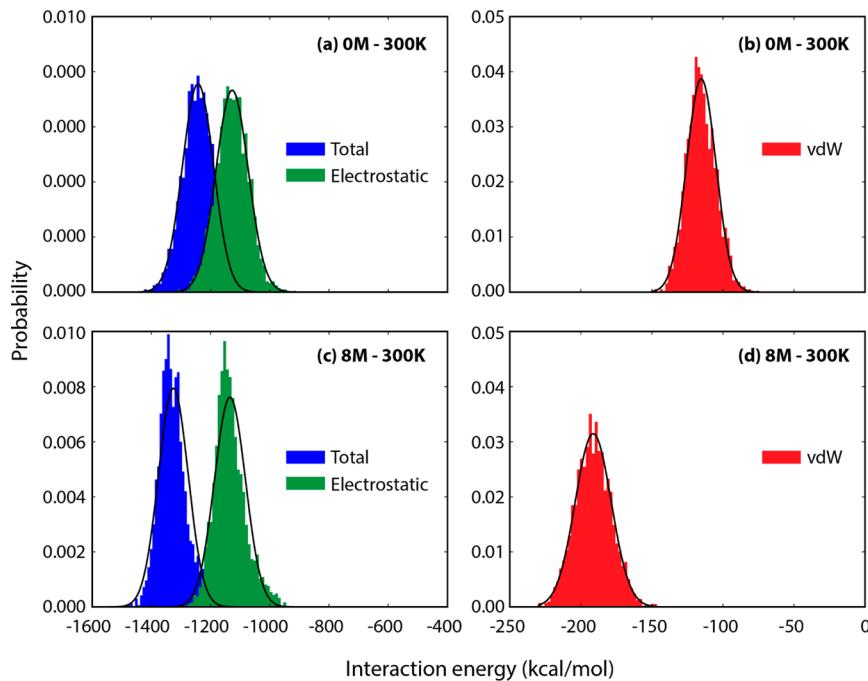


Figure 4. Probability distributions of the interaction energies between the unfolded state of the protein and the solvent environment in the presence and absence of urea at 300 K. The best fits of the Gaussian functions are depicted by black lines. The total interaction energies (blue) and the corresponding electrostatic (green) and van der Waals (red) components are given.

Dynamics of the Protein and the Trp6 Residue in Unrestrained MD Simulations. The final configurations of the 300 K umbrella sampling simulations in the presence and absence of urea corresponding to the last window (42 Å) were subjected to 1 μs long simulations without any restraints to study the dynamics of the protein and the Trp6 side-chain. The probability distributions of the root-mean-square deviation (RMSD) of the protein with respect to the experimental structure and the SASA of the Trp6 residue are given in Figure 5, and the time series data are provided in the Supporting Information (Figure S10). The initial RMSDs of these unfolded structures during the start of the simulation were close to 12 Å. Within the first 20 ns of the 1 μs simulation of the Trp-cage miniprotein in a plain aqueous environment, the protein forms a partially unfolded structure resembling the ones obtained in the umbrella sampling MD simulation, corresponding to the reaction coordinate value of 14 Å. However, in the presence of urea, the protein samples a much wider conformational space, with the RMSD fluctuating between 4 and 12 Å, and these structures resemble the ones observed in the umbrella sampling simulations (windows from 18 to 35 Å). This indicates that the free energy profiles obtained based on the reaction coordinate used here are qualitatively similar to the actual free energy surface. Differences in the extent of sampling of the conformations in the presence and absence of urea further reiterate the stabilization of the unfolded state of the protein in the presence of urea. The completely exposed Trp6 side-chain in the beginning of the unrestrained MD simulation in the absence of urea characterized by a SASA value of around 175 Å² gets buried, and the value reduces to about 50 Å² in the nanosecond time scale. This is reflected by a dominant peak centered around this value. On the other hand in the presence of urea, the Trp6 side-chain prefers to be solvated, exhibiting higher values of the SASA values. In addition to the structural characteristics obtained using the umbrella sampling, the microsecond-long unrestrained MD

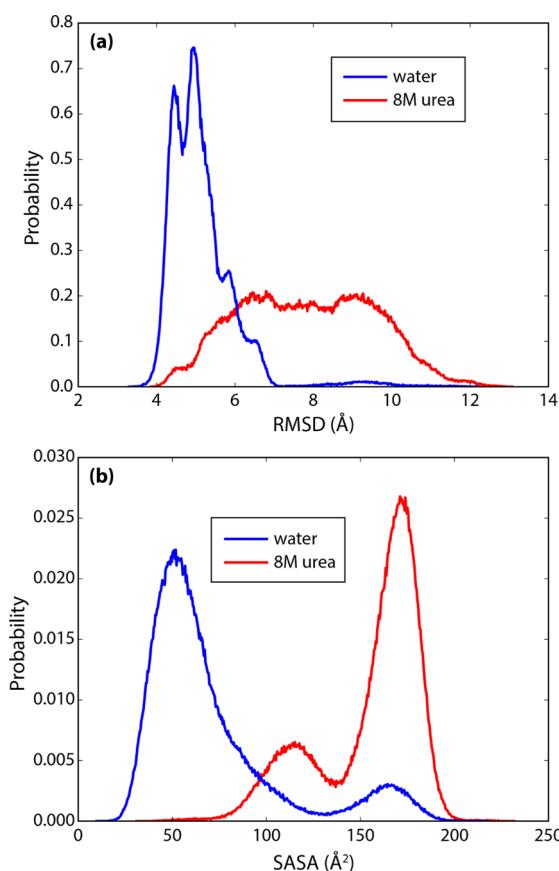


Figure 5. Probability distributions of the RMSD values of the non-hydrogen atoms of the protein (a) and solvent-accessible surface area of the side-chain of the Trp6 residue (b) obtained in the presence and absence of urea based on the 1 μs trajectories.

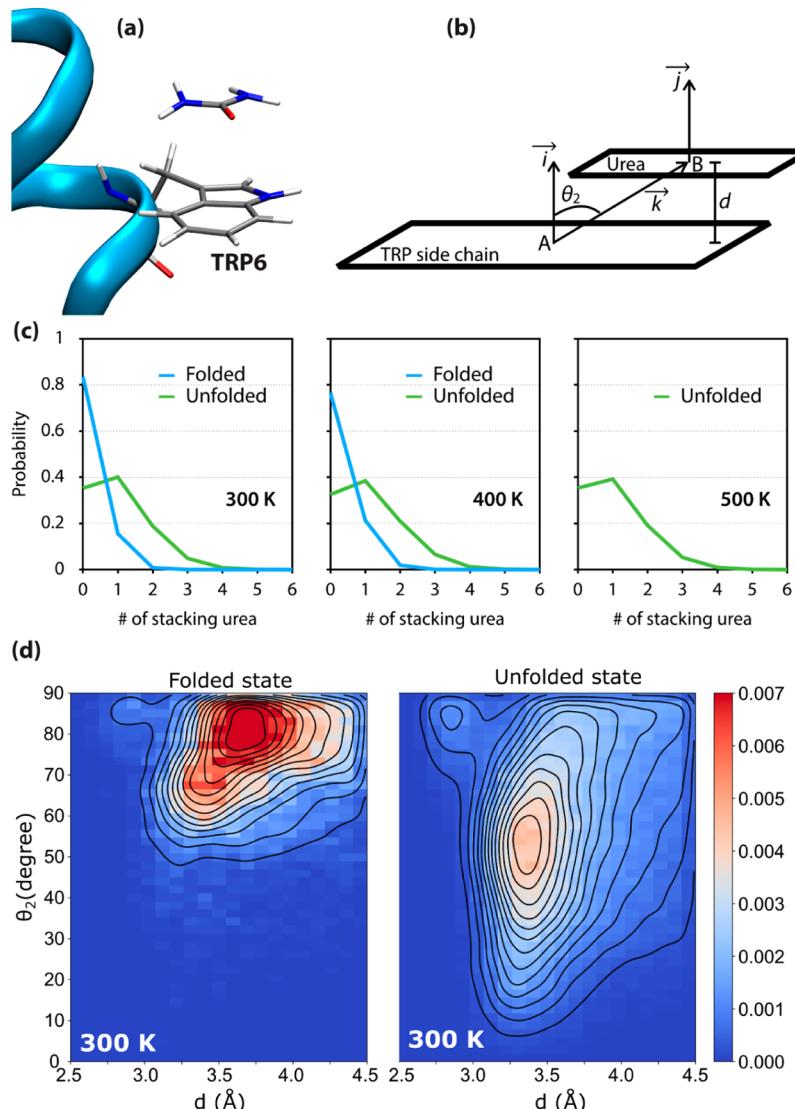


Figure 6. (a) Representative configuration of the Trp–urea stacking interaction. (b) Geometric parameters to identify stacking interactions: d is the distance between the closest atoms of the Trp side-chain and urea, θ_1 is the angle between \vec{i} and \vec{j} , and θ_2 is the angle between \vec{i} and \vec{k} . (c) Probability distributions of the number of urea molecules found to be in stacking arrangement with respect to the Trp6 side-chain in the folded and unfolded states (criteria used: $d \leq 4.5 \text{ \AA}$, $\theta_1 \leq 40^\circ$, and $\theta_2 \leq 70^\circ$). (d) Probability distribution of d and θ_2 calculated for the folded and unfolded states at 300 K with the criteria $\theta_1 \leq 40^\circ$.

simulations reveal the characteristically distinct behavior of the aromatic Trp6 residue in water vs aqueous urea. This along with experimental results points to the importance of urea–aromatic interactions in stabilizing the aromatic residues in their solvent-exposed states. Further sections of this paper focus on the nature of urea–aromatic interactions and their structural, energetic, and dynamic aspects.

Stacking and NH– π Interactions with Urea Stabilize Trp6 in the Unfolded State. Visualization of the trajectories revealed three kinds of interactions between the Trp6 side-chain and urea: (i) stacking, where the planes of the two molecules are almost parallel to each other; (ii) NH– π interaction, where one of the NH bonds of urea orients perpendicular to the aromatic ring; (iii) conventional hydrogen-bonding interaction between the hydrogen atom connected to the N atom of the Trp side-chain and the oxygen of urea. It has been shown before that the preferential interaction coefficient of urea toward aromatic groups is the maximum compared to other polar/aliphatic/ionic groups.^{44–46} However, the nature of interactions that are

responsible for such a remarkable preference of urea for aromatic groups of proteins is presented here for the first time. The prevalence of the two nonconventional interactions in the folded and unfolded states is discussed here.

Stacking Interactions. A representative structural snapshot showing the stacking interaction between the side-chain of Trp6 and a urea molecule is depicted in Figure 6a. Three vectors, first normal to the plane of Trp (\vec{i}), second, normal to the plane of urea (\vec{j}), and, third, the vector connecting the centers of masses of Trp and urea excluding the hydrogen atoms (\vec{k}), and the shortest distance (d) between the non-hydrogen atoms of the side-chain and urea were used to differentiate a stacked orientation from the others (Figure 6b). A value of 0° for the angle between \vec{i} and \vec{j} (θ_1) indicates that the two molecules are exactly parallel to each other, and a value of 0° for the angle between \vec{i} and \vec{k} (θ_2) indicates that the urea is not displaced parallel to the Trp side-chain. Cutoff values for θ_1 ensure parallel orientation of the two planar molecules, and θ_2 ensures that those urea molecules that are hydrogen bonded to Trp6 are not designated as stacked

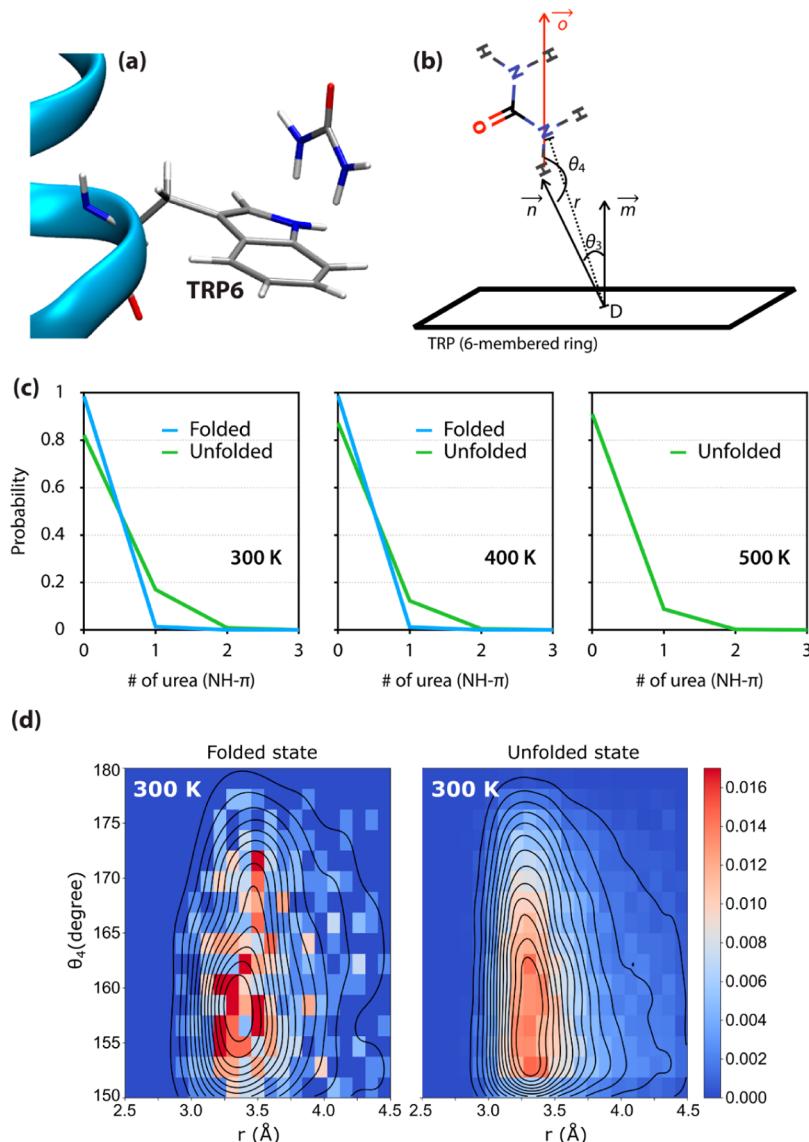


Figure 7. (a) Representative configuration of the $NH-\pi$ interaction between the Trp side-chain and urea. (b) Geometric parameters to identify $NH-\pi$ interactions: r is the distance between N of urea and the center of mass of the six-membered aromatic ring, θ_3 is the angle between \vec{m} (vector normal to the plane of the Trp side-chain) and \vec{n} (vector that connects the center of mass of the six-membered ring and hydrogen of urea connected to the nitrogen atom based on which the distance r is defined), and θ_4 is the angle between \vec{n} and \vec{o} (vector defined by the bond between N and H atoms based on which \vec{n} and \vec{o} are defined). (c) Probability distributions of the number of urea molecules found to be interacting with the Trp6 side-chain via $NH-\pi$ interaction in the folded and unfolded states (criteria used: $r \leq 4.5 \text{ \AA}$, $\theta_3 \geq 150^\circ$, and $\theta_4 \geq 150^\circ$). (d) Probability distributions of r and θ_4 calculated for the folded and unfolded states at 300 K with the criteria $\theta_3 \geq 150^\circ$.

arrangements. The probability distributions of the number of urea molecules in a stacking arrangement in the folded and unfolded states obtained from the umbrella sampling simulations at the three different temperatures calculated based on the geometric criteria defined above are given in Figure 6c. In the unfolded states, maximum probability was observed for at least one urea being stacked above or below the Trp6 residue at all three temperatures. Interestingly, the probabilities of multiple urea molecules stacking to the Trp side-chain are nonzero, suggesting the possibility of the Trp residue having stacking interactions with more than one urea molecule. In the folded state, urea is less likely to occupy the stacking orientation since the Trp6 residue is buried in the cage. Accordingly in the folded state, the maximum probability is observed for the number of stacked urea molecules to be zero. Figure 6d gives the probability distributions of the two parameters d and θ_2 within a cutoff of θ_1

$\leq 40^\circ$ calculated for the folded and unfolded states at 300 K (see Figure S11 in the Supporting Information for data from 400 K). As expected, folded states do not sample configurations exhibiting urea–aromatic stacking interactions. However, the probability distributions corresponding to the unfolded states exhibit a peak around a value of 3.3 Å for the parameter d . Note that hydrogen-bond-like interactions would exhibit a peak around 2.8 Å. High probability for the region around $\theta_2 = 50\text{--}60^\circ$ indicates that the two planes are slightly displaced during such stacking interactions. Quantum mechanical calculations confirm that urea displaced with respect to indole and other aromatic model systems is energetically preferred (see later). Notably, stacking-type interactions between the guanidinium group and aromatic side-chains have been identified earlier.^{95,96} However, this is via well-studied cation– π interactions,⁹⁷ which

are not feasible in the case of urea–aromatic stacking since urea is electrically neutral.

NH– π Interactions. In addition to stacking interactions, another kind of interaction between urea and Trp6 was identified as the NH– π interactions (Figure 7a). Such an interaction is favorable when one of the hydrogen atoms from the N–H bond of urea points toward the π -face of the aromatic group. A single molecule of urea can achieve this mode of interaction via several orientations since there are two distinguishable types of hydrogen atoms. Similar to stacking, a geometric criterion was defined based on three vectors (first normal to the plane of the six-membered ring of the Trp side-chain, \vec{m} , second, connecting the centroid of the six-membered ring of the Trp6 side-chain (D) and the closest hydrogen atom of urea, \vec{n} , and, third, the N–H bond (\vec{o}) and a distance, r (between point D and the closest nitrogen of the urea). The angle between \vec{m} and \vec{n} (θ_3) and the angle between \vec{n} and \vec{o} (θ_4) and distance r were used to differentiate orientations with NH– π interactions from others. Probability distributions for the number of urea molecules interacting with the π -surface of the aromatic side-chain via the N–H bond are given in Figure 7c. In the folded states, no significant NH– π interactions are observed with the probability for zero close to being 1. However, in the unfolded states such interactions are present in significant numbers, as reflected in nonzero probabilities for the value of 1. The probability distributions for the distance, r , and angle θ_4 (Figure 7d) for the folded and unfolded states display the prevalence of NH– π interactions in the unfolded states (see Figure S12 in the Supporting Information for data from 400 K). Values of about 3.3 to 3.4 Å for r and 155–160° for θ_4 correspond to the most preferred orientations of urea with respect to the aromatic group for this type of interaction.

The probability distributions for hydrogen-bonding interactions are given in Figure S13 of the Supporting Information. The strengths of these three interactions, namely, stacking, NH– π , and hydrogen bonding, were quantified using dissociation constants (details of the calculations are given in the Supporting Information along with the data, Table S1), which reveal the stacking to be the strongest among the three. Previous studies have shown the existence of lone pair– π interactions in proteins and nucleic acids in addition to other conventional nonbonded interactions such as hydrogen bonding.^{98–100} Analysis corresponding to interaction between the carbonyl oxygen of urea and the π -surface of Trp6 (similar criteria used for NH– π interactions) revealed negligible sampling of such interactions and hence was not considered further. A recent study by Cheng et al. showed that amide oxygen interacts with the aromatic group more preferably than the amide nitrogen, with the latter being unfavorable.⁴⁹ Notably, criteria defined for the stacking and NH– π interactions do not preclude the possibility of urea oxygen being present within the interacting distance from the Trp side-chain. The main objective of this study is to understand the modes of interaction between urea and aromatic groups, and hence atomwise contributions are not calculated. Based on the above discussions it is clear that nonconventional interactions, viz., stacking and NH– π interactions between urea and aromatic side-chains, are highly prevalent. These interactions in addition to any possible hydrogen-bonding interactions are proposed to be responsible for stabilizing Trp6 in the unfolded state. The following section examines the dynamic properties of these novel interactions by analyzing the lifetimes of these associations based on the 1 μ s unbiased MD simulation trajectory.

Urea–Aromatic Stacking Interactions Last Longer than NH– π and Hydrogen-Bonding Interactions. Urea–aromatic side-chain stacking contributions in protein denaturation are shown here for the first time. Visualization of these interactions indicates that some are long-lived (inset of Figure 8a); mean lifetimes of the urea–Trp stacking interactions were calculated using the 1 μ s long MD trajectory and were compared with those of the NH– π and hydrogen-bonding interactions. Using the criteria for stacking between urea and the Trp6 side-chain defined above, dwell time distributions for all the stacking events throughout the 1 μ s were calculated using the following equation.¹⁰¹

$$P(t; t^*) = \frac{1}{N} \sum_{i=1}^N \delta[\tau_i(t^*) - t] \quad (1)$$

Here, t^* , the transient disruption time or the relaxation time, allows for considering events that have temporary disruptions, which is common in liquids. For example, if a stacking interaction is formed between a given urea molecule and the Trp6 side-chain and if the interaction is disrupted for a time less than t^* followed by reassociation, then such a disruption is ignored and the interaction is considered to be present continuously. N is the number of stacking events (dwells) during the whole simulation time, $\tau_i(t^*)$ is the time interval of the i^{th} dwell, and $\delta[\tau_i(t^*) - t] = 1$ if $t < \tau_i(t^*) < (t + dt)$ and 0 otherwise. Dwell time distributions were calculated for a wide range of t^* from 5 ps to 5 ns, and the one obtained with a t^* value of 1 ns is given in Figure 8a. The dwell time distribution reveals that the probability exponentially decreases with respect to the increase in the total duration of the event. Interestingly, the longest stacking event was found to be for a duration of 49.2 ns with no other event after about 12 ns. The time series of the distance, d , during the part of the simulation where stacking between this urea and the Trp6 side-chain occurs, is given in Figure 8b, which shows that a distant urea molecule diffuses close to the Trp6 residue, gets involved in a stacking interaction with its side-chain for about 49 ns, and then diffuses away. Visualization of this event revealed that this urea molecule is constrained to be stacked because of cooperativity from other conventional interactions with the protein. For a given t^* , survival probabilities of the stacking interactions using the dwell time distributions were calculated using the following expression (eq 2). The survival probability thus calculated was fit to a triexponential function, using which the average lifetimes of stacking interactions were calculated using eq 3. Figure 8c shows the calculated survival probability for $t^* = 1$ ns and the best triexponential fit.

$$S(t; t^*) = 1 - \int_0^t P(\tau; t^*) d\tau \quad (2)$$

$$\langle \tau_i(t^*) \rangle = \int_0^\infty S(t; t^*) dt \quad (3)$$

As mentioned in the previous section, there are instances where the Trp6 side-chain interacts with a multiple number of urea molecules via stacking interactions. Calculations were also performed for two and three molecules of urea simultaneously interacting with Trp6 using the same procedure explained above except that the definition of the events changes. For two and three ureas simultaneously interacting with the Trp side-chain, the geometric criteria for every possible pair and triple of urea molecules, respectively, were analyzed. The average lifetimes calculated using eq 3 for multiple values of t^* for single, double,

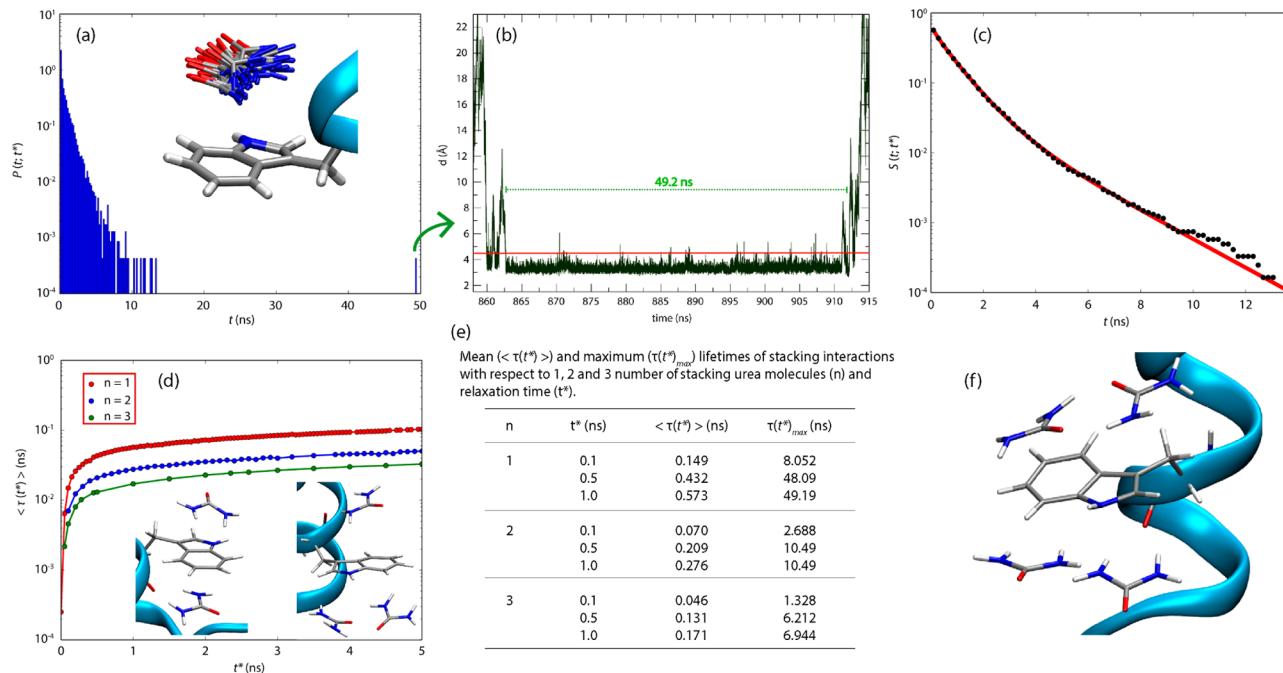


Figure 8. Trp–urea stacking lifetime calculations. (a) Stacking dwell time distributions ($P(t; t^*)$) obtained for the 1 μ s unbiased simulation of the protein in 8 M aqueous urea with $t^* = 1$ ns. (b) Time series of distance between the closest atoms of the atoms of the Trp side-chain and urea (r) for the longest event ($t = 49.2$ ns). (c) Survival probability distribution calculated using the stacking dwell time distribution (a) given in black dots for $t^* = 1$ ns. The best fit triexponential function for the survival probability is depicted by the red line. (d) Mean lifetimes (red) calculated based on the best fit triexponential functions for different transient disruption times (t^*). The mean lifetimes for two (blue) or three (green) urea molecules simultaneously interacting with the aromatic side-chain (insets show representative structures of such stacking arrangements) are also given. (e) Mean and maximum lifetimes of stacking interactions between one, two, and three urea molecules with the Trp side-chain calculated for select values of t^* . (f) Structural representation of one of the rare events of stacking interaction involving four different urea molecules.

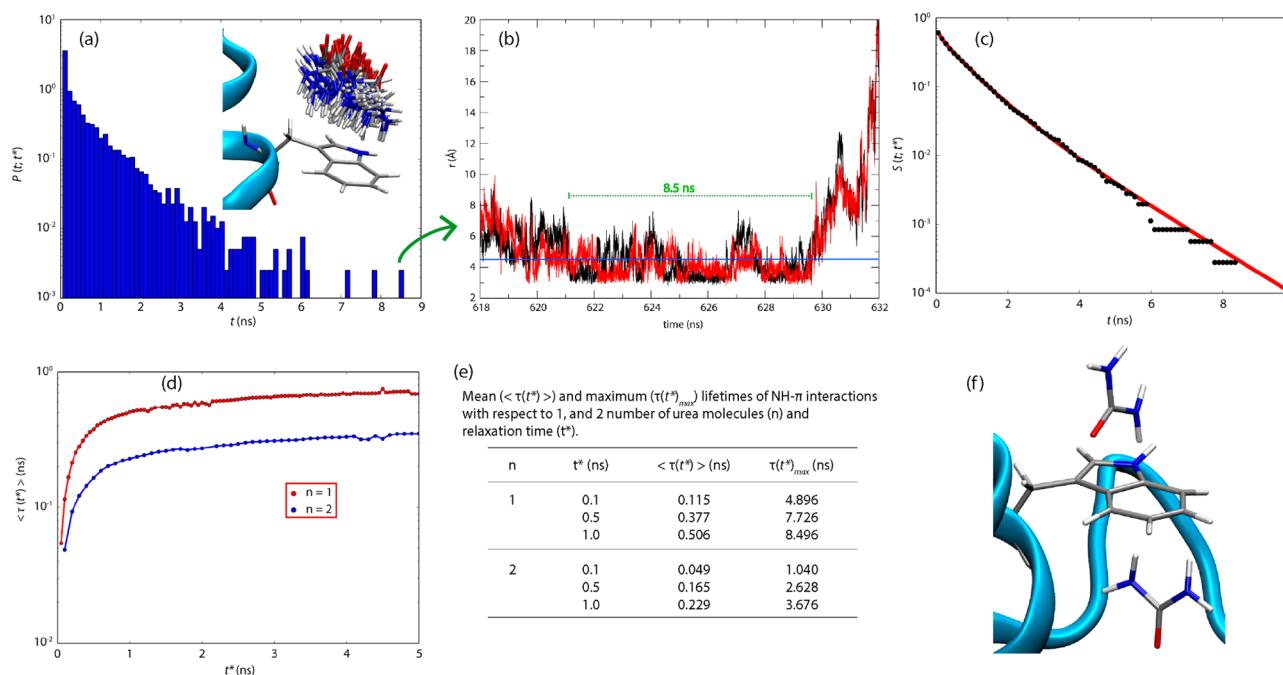


Figure 9. Trp–urea NH- π lifetime calculations. NH- π stacking dwell time distributions (a), time series of the distance d corresponding to the longest NH- π interaction (b), survival probability and the triexponential fit (c), and the mean lifetimes with respect to one and two ureas interacting with the Trp6 residue via NH- π interaction (d). The mean and maximum lifetimes of NH- π interactions involving one and two urea molecules (e) and a structural snapshot showing Trp6 interacting with two urea molecules simultaneously (f) are given.

and triple urea molecule stacking with the Trp6 side-chain are given in Figure 8d. The insets in this figure depict the illustrations of representative structural snapshots where two and three urea

molecules are simultaneously involved in stacking. The mean and maximum lifetimes for the stacking events for select values of t^* are given in Figure 8e. This reveals that the mean lifetimes of

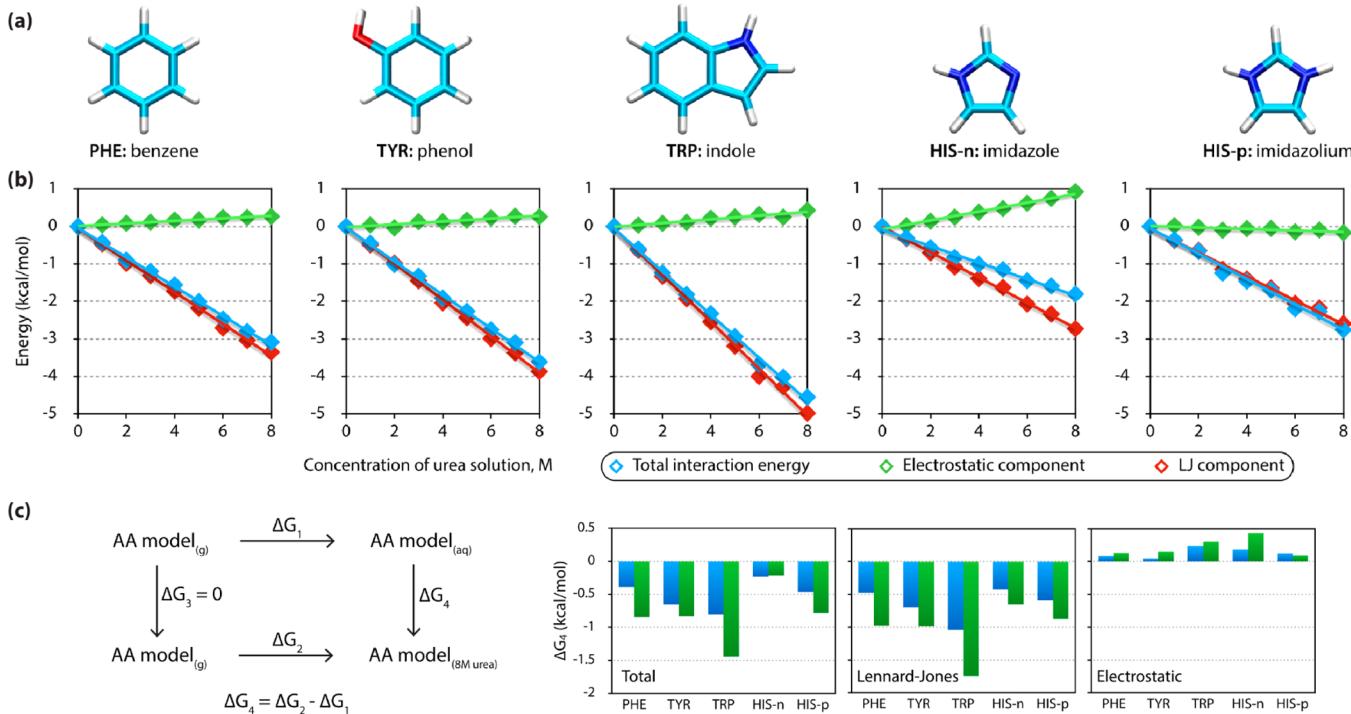


Figure 10. (a) Model systems conceived representing five different aromatic side-chains. (b) Interaction energies between the aromatic model systems and the solvent environment (blue) with respect to the concentration of the urea solution—energy values relative to those obtained in an aqueous solution (0 M urea)—are presented. The electrostatic and van der Waals contributions to the total interaction energies are also given (green and red, respectively). (c) Thermodynamic cycle considered for calculating the transfer free energies using the thermodynamic integration method. The total free energies and the electrostatic and Lennard-Jones contributions for 0 to 4 M (blue) and 0 to 8 M (green) are plotted.

these interactions are on the order of hundreds of picoseconds, indicating that the stacking interactions are reasonably long-lived. Very few instances of four urea molecules interacting with the aromatic side-chain were observed (Figure 8f), and statistically meaningful mean lifetimes could not be calculated. The mean lifetimes for NH–π interactions (Figure 9) are found to be marginally less than those calculated for the stacking interactions, and this trend is consistent for all values of t^* . Lifetimes for hydrogen bonding between the N–H of Trp and the O atom of urea (Figure S14 in the Supporting Information) reveal that these interactions are short-lived compared to both stacking and NH–π interactions.

Similar to urea, hydrogen-bonding interactions and OH–π interactions are possible between aromatic groups and water. Hence, stacking interactions between urea and aromatic groups is proposed to be responsible for stabilizing the solvent-exposed aromatic residues, which is otherwise not possible in the water-alone system. The next section investigates the ability of other aromatic amino acid residues to form favorable stacking interactions with urea.

All Four Aromatic Residues Can Form Stacking and NH–π Interactions with Urea. It is evident from the above discussions that stacking is ubiquitous between the Trp side-chain and urea. Following this, a straightforward question, “are other aromatic amino acids capable of forming stacking interactions with urea?”, arises. Five model systems (benzene, phenol, indole, imidazole, and imidazolium ion, Figure 10a) were conceived to examine urea stacking and NH–π interactions with all the aromatic residues. Each of these molecules was simulated in the presence of water and different concentrations of aqueous urea (1 M, 2 M, 3 M, 4 M, ..., 8 M). Similar to the Trp side-chain in the Trp-cage miniprotein, stacking and NH–π interactions are

evident in all of them (Figures S15–S25 in the Supporting Information). The interaction energies between the aromatic model system and the environment and the LJ and electrostatic components were calculated to examine the stabilization of the aromatic groups with respect to different concentrations of aqueous urea. The interaction energies thus calculated relative to the interaction energies obtained for the 0 M system are given in Figure 10b. In all the systems, the favorability of the interaction between the aromatic model compounds and the surroundings increases with respect to an increase in the concentration of urea, showing a linear dependence. Interestingly, the electrostatic contribution to the change in the interaction energies is slightly repulsive except in the imidazolium ion, where it is marginally attractive. Stabilization due to the presence of urea compared to the plain water system is primarily due to vdW-type dispersion interactions in all the systems. Since these interaction energies do not include entropy contributions, transfer free energies for water to 4 and 8 M aqueous urea solutions were calculated using the thermodynamic integration method (Figure 10c). The transfer free energies were calculated for both directions (solvation and desolvation), and convergence of these two values was ensured. Transfer and hydration free energies obtained here are in good agreement with experimental data (Tables S2 and S3). Similar to the interaction energies, the solvation free energies of all the aromatic model systems become more favorable in the presence of urea due to stabilizing LJ contributions where the electrostatic contributions are marginally repulsive. Ganguly and van der Vegt have studied a benzene–aqueous system using a coarse-grained force field based on KBFF and found that benzene has a significantly reduced tendency to aggregate and form clusters in the presence of urea.^{102,103} The

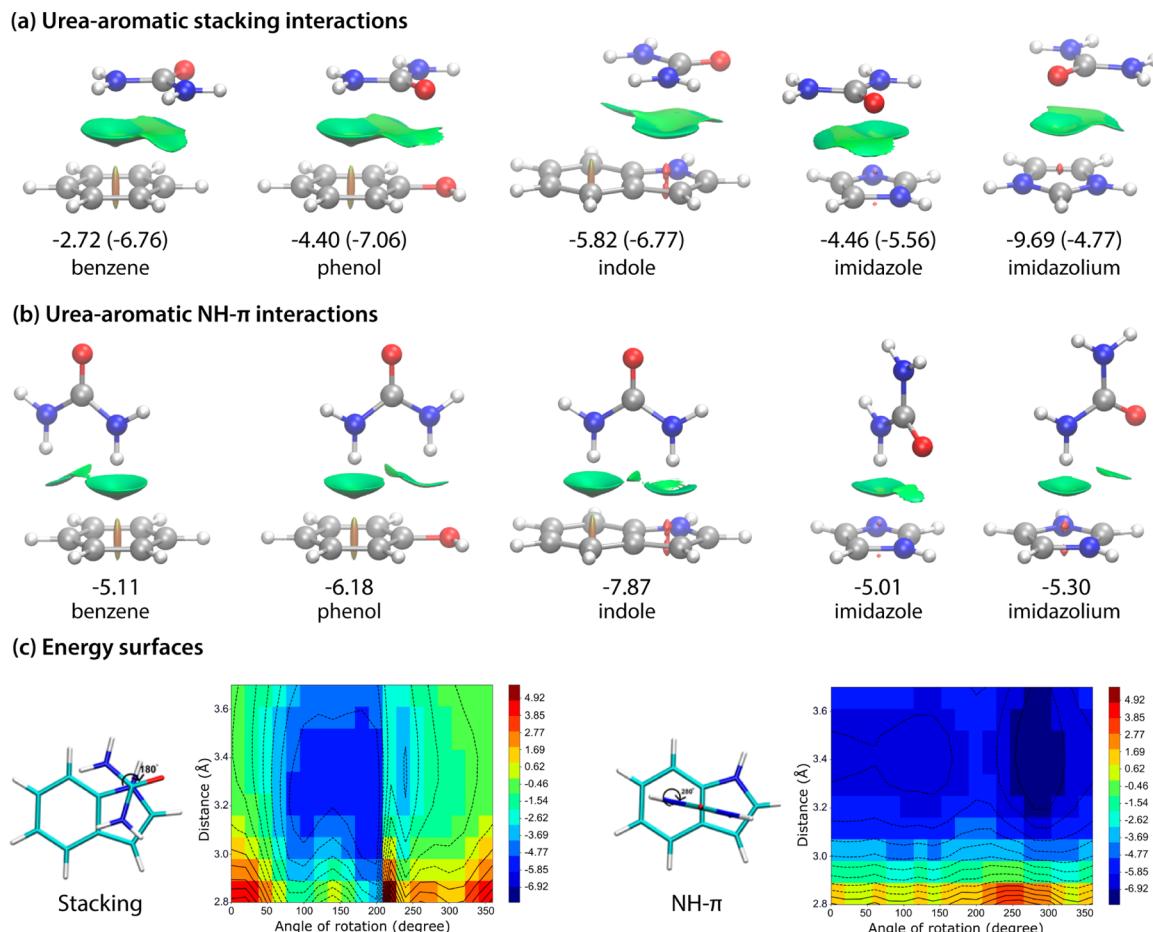


Figure 11. (a) Most stable stacking arrangement between the five different aromatic model systems with a urea molecule. The green-colored isosurfaces obtained from NCI plot show a weak noncovalent interaction between the two. The interaction energy values along with the dispersion contribution (in parentheses) obtained at the RI-MP2/cc-pVQZ level are given below the structures. (b) Structural representations of the most favorable orientations of urea with respect to each of the five aromatic model systems due to NH- π interactions. The green isosurfaces showing favorable interaction and the interaction energies are given. (c) Potential energy contour surface corresponding to the most stable stacking and NH- π interactions obtained at the RI-MP2/aug-cc-pVDZ level. The interaction energies were calculated with respect to the distance between urea and indole, and the orientation of urea with respect to indole (see *Results and Discussion* for more details). All values are given in kcal/mol.

urea–aromatic interactions proposed here explain this phenomenon well.

Quantum Mechanical Calculations Support Urea–Aromatic Novel Interactions Dominated by Dispersion. The reliability of the results from MD simulations and free energy calculations depends on the quality of the force field parameters used. Interactions such as stacking and NH- π interactions identified here are usually not explicitly considered during the force field parametrization process, although the factors responsible for such interactions are indirectly considered during calculations of solvation free energy and other condensed phase properties. Quantum mechanical calculations at the RI-MP2/aug-cc-pVDZ level of theory were performed on the five model systems (Figure 10a) to validate the existence of stacking and NH- π interactions between urea and aromatic groups. Initially, 216 different orientations of urea with respect to indole in their individually optimized geometries were modeled so that the two molecular planes are parallel to each other and the distance between the planes is 3.4 Å (since the probability distribution had a maximum at this distance, Figures 6). These 216 orientations differ from each other based on the orientation of the urea molecule and the position of the central carbon of urea with respect to the indole molecule (Figure S26 in the

Supporting Information). Interaction energies were calculated after accounting for the basis set superposition error and were subjected to energy decomposition analysis. Interaction energies were calculated for the most stable system using the larger cc-pVQZ basis set. The most favorable interaction (-6.3 kcal/mol) was observed for the arrangement where the vector joining the nitrogen atom of the indole ring and the central carbon of urea is perpendicular to both molecular planes (Figure 11a). All the other interaction energies are presented in Figures S27 to S32 of the *Supporting Information*. Energy decomposition analysis shows that the dispersion contribution to this interaction energy is -7.3 kcal/mol and that the electrostatic contribution is slightly repulsive. Similarly, hundreds of model systems were conceived for the other four model systems, and the most stable arrangement for each of them along with their interaction energies are given in Figure 11. The figure also depicts the noncovalent interaction index calculated from the electron density and its derivatives showing weak interaction between the two molecules. Box plots corresponding to the interaction energies and electrostatic and dispersion contributions with respect to the change in the interplanar distances show that the dispersion interactions dominate the total interaction energies (Figure S33 in the *Supporting Information*). In all five model

systems, arrangements with the most stable stacking interaction have urea parallelly displaced with reference to the aromatic ring. This is consistent with the stacking probability distributions where maximum probability for stacking was found for angles close to 50° (Figure 6). Incidentally, the most stable π -stacking benzene dimer has a parallel displaced structure. The most stable urea–indole stacked complex was optimized at the MP2 level, and a stationary point on the potential energy surface was obtained. Similarly, a model system with two urea molecules interacting via either side of the indole molecular plane was optimized, and a stationary point was obtained (Figure S34 in the Supporting Information). NH– π interactions were investigated, and the most stable structures along with interaction energies are given in Figure 11b. For NH– π interactions, both possible orientations of urea where the carbonyl bond is pointed toward and away from the aromatic ring, were considered. Starting from the two most stable complexes, one each from stacking and NH– π interactions, potential energy surface calculations at the MP2/aug-cc-pVDZ level were carried out. In the case of a stacking interaction, energies were calculated with respect to the change in the angle of rotation of urea with respect to indole, and the N(indole)–C(urea) distance. For NH– π interaction, energies were calculated with respect to the change in the distance between N of urea and the centroid of the six-membered ring of indole, and the angle of rotation with respect to the vector connecting these two points. Figure 11c depicts the energy surfaces for both of these, which shows well-defined minima corresponding to both stacking and NH– π interactions (see Figures S35 and S36 in the Supporting Information for the potential energy surfaces for all the models). Thus, the results obtained using the MD simulations and the quantum mechanical calculations are consistent with each other, substantiating the role of stacking and NH– π interactions in favoring the exposure of aromatic residues in the presence of urea. The interaction energies for the most stable systems were also calculated using the CHARMM force field and are in good agreement with the QM results (Table S4 in the Supporting Information). The next section presents experimental evidence for a urea–aromatic stacking interaction and its possible role in biological systems.

Experimental Evidence for Urea–Aromatic Stacking Interactions. NH– π -type interactions between amine/amide hydrogens and aromatic groups have been observed previously.^{104–106} However, stacking between urea and aromatic groups is rather unintuitive. It has previously been shown to play a crucial role in the chemical denaturation of RNA.^{107–110} In this section, we explore possible roles of urea–aromatic stacking in biological systems and in existing bimolecular structures.

Urea Transporters. These are transmembrane proteins that selectively allow for permeation of urea molecules across a lipid bilayer of cells. Three-dimensional structures of three of these proteins have been characterized.^{111–113} In all three structures, aromatic residues have been found to line the selectivity filter of the pore through which urea permeates. Structural representation of the urea transporter and the selectivity filter from *Desulfovibrio vulgaris* are depicted in Figure 12a.¹¹³ The two pairs of Phe residues form the beginning and the end of the selectivity filter. In each of the two pairs, the six-membered rings are separated by a distance of about 7.5 Å. When a urea molecule passes through the aromatic groups, it is structurally constrained to be parallel to the phenyl rings. The mechanism of urea transport across this channel was studied using MD simulations, and it was found that stacking interactions between urea and

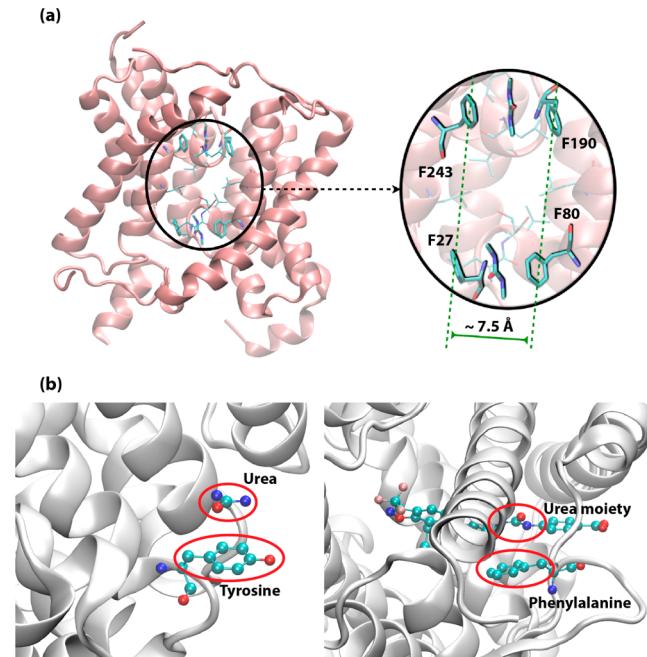


Figure 12. Experimental evidence for stacking interactions between aromatic groups and urea. (a) Crystal structure of the urea transporter from *Desulfovibrio vulgaris*. The selectivity filter of the transporter is shown, highlighting the arrangement of the two pairs of phenylalanine residues interacting with dimethylurea molecules via stacking. (b) Structural representations of two crystal structures where a urea molecule is stacking with a tyrosine residue and in the other where a urea derivative is in stacking arrangement with a phenylalanine residue.

phenylalanine side-chains are crucial for the selectivity of the channel.¹¹⁴

Crystal Structures. Several protein structures in the Protein Data Bank contain urea or derivatives of urea as ligand or as cosolvent molecules. A total of 420 such structures, 61 with urea and 359 with urea derivatives bound to protein structures, were obtained from the PDB. All instances of the existence of urea–aromatic stacking interactions were identified using the geometric criteria defined above (minimum distance between aromatic group and urea atoms \leq 4.5 Å; angle between the two planes \leq 40°; angle between the normal of the aromatic group and the center of mass of urea \leq 70°). From these structures, 38% of the structures with urea and 25% of the ones with urea derivatives were found to have a stacking arrangement with one of the four aromatic residues. Structural representations of one of the structures from each of these two sets are given in Figure 12b (PDB ID: 3IPU and 4EV9).¹¹⁵ The probability distributions of the distance and angle corresponding to the urea derivatives presented in Figure S37 of the Supporting Information are similar to the distributions obtained in this study (Figure 6d). A similar plot could not be obtained for unsubstituted urea due to the scarcity of data.

CONCLUSIONS

This study reports a series of MD simulations and quantum mechanical calculations to investigate the nature of interactions that are responsible for maximum affinity between urea and aromatic groups compared to other functional groups in proteins. Initially, the effect of temperature and the presence of urea on the unfolding free energy profiles of Trp-cage miniprotein was studied using umbrella sampling simulations.

Free energy calculations and analysis of geometric features demonstrate the importance of distortion of the hydrophobic core along the unfolding pathway of Trp-cage miniprotein. Stabilization of the Trp6 residue of unfolded protein in the presence of urea is due to the prevalence of hitherto unknown stacking and NH- π interactions between urea and the aromatic group. This was further confirmed by microsecond-long simulations in water and in aqueous urea. The unfolded states in aqueous solution are stabilized in the presence of urea mainly due to vdW and not electrostatic interactions. Mean lifetimes of the different kinds of interactions between urea and the Trp6 side-chain based on the long MD trajectory revealed that the stacking interactions are long-lived compared to both NH- π and hydrogen-bonded interactions. The ability of urea to form these unconventional associations was confirmed by analyzing geometric parameters and transfer free energies of different aromatic model systems. Further quantum mechanical calculations validated the stabilizing nature of both stacking and NH- π interactions. Among the three types of interactions, N(O)H- π and hydrogen bonding are possible for both urea and water. However, the ability of urea to form stacking interactions is key and is proposed to be responsible for the difference in the behavior of aromatic groups in water and in aqueous urea. Urea-aromatic stacking interactions seem to be manifested in certain biological systems such as urea transporters. Urea transporters use stacking interactions to discriminate and regulate urea permeation across cell membranes. Experimental observations of such novel interactions further support our findings that urea-aromatic stacking interactions are crucial in the shifting of the protein equilibrium process when aromatic amino acids are involved. We propose that such interactions can possibly be exploited to increase the specificity and affinity of inhibitors during the lead optimization stage in the drug design and development process.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/jacs.7b05463](https://doi.org/10.1021/jacs.7b05463).

Figures and tables showing convergence of PMF profiles, adequate overlap between successive windows in umbrella sampling simulations, SASA values for the whole range of the reaction coordinate, probability distributions of the interaction energies obtained at 400 K, time series of the RMSD and SASA values obtained using 1 μ s simulations, probability distributions showing stacking and NH- π interactions, dissociation constants, MM interaction energies, comparison of computed and experimental hydration/transfer free energies, and quantum mechanically obtained potential energy surfaces ([PDF](#))

AUTHOR INFORMATION

Corresponding Author

*deva@iiit.ac.in

ORCID

U. Deva Priyakumar: [0000-0001-7114-3955](https://orcid.org/0000-0001-7114-3955)

Author Contributions

[†]S. Goyal and A. Chattopadhyay contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank the Indian National Science Academy for funding via the research grant for INSA medal awardees (No. SP/YSP/116/2015/103, dt.third Dec, 2015). We would like to thank Dr. Prabhakar Bhimalapuram for fruitful discussions.

REFERENCES

- (1) Dill, K. A. *Biochemistry* **1990**, *29*, 7133–7155.
- (2) Best, R. B. *Curr. Opin. Struct. Biol.* **2012**, *22*, 52–61.
- (3) Englander, S. W.; Mayne, L. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 15873–15880.
- (4) Malhotra, P.; Udgaonkar, J. B. *Protein Sci.* **2016**, *25*, 1924–1941.
- (5) Rao, V. H. G.; Gosavi, S. *Curr. Opin. Struct. Biol.* **2016**, *36*, 67–74.
- (6) Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 8069–8072.
- (7) Anderson, D. E.; Becktel, W. J.; Dahlquist, F. W. *Biochemistry* **1990**, *29*, 2403–2408.
- (8) Toba, S.; Hartsough, D. S.; Merz, K. M. *J. Am. Chem. Soc.* **1996**, *118*, 6490–6498.
- (9) Colombo, G.; Toba, S.; Merz, K. M. *J. Am. Chem. Soc.* **1999**, *121*, 3486–3493.
- (10) Lin, L.-N.; Brandts, J. F.; Brandts, J. M.; Plotnikov, V. *Anal. Biochem.* **2002**, *302*, 144–160.
- (11) Mitra, L.; Smolin, N.; Ravindra, R.; Royer, C.; Winter, R. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1249–1265.
- (12) Cafisch, A.; Karplus, M. *Structure* **1999**, *7*, 477–S2.
- (13) Alonso, D. O.; Daggett, V. *J. Mol. Biol.* **1995**, *247*, 501–520.
- (14) Pace, C. *Methods Enzymol.* **1986**, *131*, 266–280.
- (15) England, J. L.; Haran, G. *Annu. Rev. Phys. Chem.* **2011**, *62*, 257–277.
- (16) Canchi, D. R.; García, A. E. *Annu. Rev. Phys. Chem.* **2013**, *64*, 273–293.
- (17) Bandyopadhyay, D.; Mohan, S.; Ghosh, S. K.; Choudhury, N. *J. Phys. Chem. B* **2014**, *118*, 11757–11768.
- (18) Watlaufer, D. B.; Malik, S. K.; Stoller, L.; Coffin, R. L. *J. Am. Chem. Soc.* **1964**, *86*, 508–514.
- (19) Rupley, J. *J. Phys. Chem.* **1964**, *68*, 2002–2003.
- (20) Frank, H. S.; Franks, F. *J. Chem. Phys.* **1968**, *48*, 4746–4757.
- (21) Finer, E.; Franks, F.; Tait, M. *J. Am. Chem. Soc.* **1972**, *94*, 4424–4429.
- (22) Vanzi, F.; Madan, B.; Sharp, K. *J. Am. Chem. Soc.* **1998**, *120*, 10748–10753.
- (23) Klimov, D.; Straub, J. E.; Thirumalai, D. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 14760–14765.
- (24) Canchi, D. R.; García, A. E. *Biophys. J.* **2011**, *100*, 1526–1533.
- (25) O'Brien, E. P.; Dima, R. I.; Brooks, B.; Thirumalai, D. *J. Am. Chem. Soc.* **2007**, *129*, 7346–7353.
- (26) Lim, W. K.; Rösgen, J.; Englander, S. W. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 2595–2600.
- (27) Berteotti, A.; Barducci, A.; Parrinello, M. *J. Am. Chem. Soc.* **2011**, *133*, 17200–17206.
- (28) Hua, L.; Zhou, R.; Thirumalai, D.; Berne, B. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16928–16933.
- (29) Candotti, M.; Esteban-Martín, S.; Salvatella, X.; Orozco, M. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 5933–5938.
- (30) Lindgren, M.; Westlund, P.-O. *Phys. Chem. Chem. Phys.* **2010**, *12*, 9358–9366.
- (31) Stumpe, M. C.; Grubmüller, H. *J. Am. Chem. Soc.* **2007**, *129*, 16126–16131.
- (32) Auton, M.; Holthauzen, L. M. F.; Bolen, D. W. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 15317–15322.
- (33) Stumpe, M. C.; Grubmüller, H. *PLoS Comput. Biol.* **2008**, *4*, e1000221.
- (34) Stumpe, M. C.; Grubmüller, H. *Biophys. J.* **2009**, *96*, 3744–3752.
- (35) Tirado-Rives, J.; Orozco, M.; Jorgensen, W. L. *Biochemistry* **1997**, *36*, 7313–7329.
- (36) Mountain, R. D.; Thirumalai, D. *J. Am. Chem. Soc.* **2003**, *125*, 1950–1957.

- (37) Candotti, M.; Perez, A.; Ferrer-Costa, C.; Rueda, M.; Meyer, T.; Gelpí, J. L.; Orozco, M. *PLoS Comput. Biol.* **2013**, *9*, e1003393.
- (38) Bennion, B. J.; Daggett, V. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 5142–5147.
- (39) Wallqvist, A.; Covell, D.; Thirumalai, D. *J. Am. Chem. Soc.* **1998**, *120*, 427–428.
- (40) Zou, Q.; Bennion, B. J.; Daggett, V.; Murphy, K. P. *J. Am. Chem. Soc.* **2002**, *124*, 1192–1202.
- (41) Wei, H.; Fan, Y.; Gao, Y. Q. *J. Phys. Chem. B* **2009**, *114*, 557–568.
- (42) Caballero-Herrera, A.; Nordstrand, K.; Berndt, K. D.; Nilsson, L. *Biophys. J.* **2005**, *89*, 842–857.
- (43) Moeser, B.; Horinek, D. *J. Phys. Chem. B* **2013**, *118*, 107–114.
- (44) Guinn, E. J.; Schwinefus, J. J.; Cha, H. K.; McDevitt, J. L.; Merker, W. E.; Ritzer, R.; Muth, G. W.; Engelsgjerd, S. W.; Mangold, K. E.; Thompson, P. J. *J. Am. Chem. Soc.* **2013**, *135*, 5828–5838.
- (45) Guinn, E. J.; Pegram, L. M.; Capp, M. W.; Pollock, M. N.; Record, M. T. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 16932–16937.
- (46) Diehl, R. C.; Guinn, E. J.; Capp, M. W.; Tsodikov, O. V.; Record, M. T., Jr. *Biochemistry* **2013**, *52*, 5997–6010.
- (47) Guinn, E. J.; Kontur, W. S.; Tsodikov, O. V.; Shkel, I.; Record, M. T. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 16784–16789.
- (48) Guinn, E. J.; Jagannathan, B.; Marqsee, S. *Nat. Commun.* **2015**, *6*, 6861.
- (49) Cheng, X.; Shkel, I. A.; O'Connor, K.; Henrich, J.; Molzahn, C.; Lambert, D.; Record, M. T., Jr. *J. Am. Chem. Soc.* **2017**, *139*, 9885–9894.
- (50) Qiu, L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. *J. Am. Chem. Soc.* **2002**, *124*, 12952–12953.
- (51) Streicher, W. W.; Makhatadze, G. I. *Biochemistry* **2007**, *46*, 2876–2880.
- (52) Ahmed, Z.; Beta, I. A.; Mikhonin, A. V.; Asher, S. A. *J. Am. Chem. Soc.* **2005**, *127*, 10943–10950.
- (53) Neuweiler, H.; Doose, S.; Sauer, M. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 16650–16655.
- (54) Mok, K. H.; Kuhn, L. T.; Goez, M.; Day, I. J.; Lin, J. C.; Andersen, N. H.; Hore, P. *Nature* **2007**, *447*, 106–109.
- (55) Culik, R. M.; Serrano, A. L.; Bunagan, M. R.; Gai, F. *Angew. Chem., Int. Ed.* **2011**, *50*, 10884–10887.
- (56) Halabis, A.; Źmudzińska, W.; Liwo, A.; Ołdziej, S. *J. Phys. Chem. B* **2012**, *116*, 6898–6907.
- (57) Rová, P.; Stráner, P.; Láng, A.; Bartha, I.; Huszár, K.; Nyitrai, L.; Perczel, A. *Chem. - Eur. J.* **2013**, *19*, 2628–2640.
- (58) Chowdhury, S.; Lee, M. C.; Xiong, G.; Duan, Y. *J. Mol. Biol.* **2003**, *327*, 711–717.
- (59) Pitera, J. W.; Swope, W. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 7587–7592.
- (60) Zhou, R. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 13280–13285.
- (61) Juraszek, J.; Bolhuis, P. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 15859–15864.
- (62) Paschek, D.; Hempel, S.; García, A. E. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17754–17759.
- (63) Gattin, Z.; Riniker, S.; Hore, P.; Mok, K.; Van Gunsteren, W. *Protein Sci.* **2009**, *18*, 2090–2099.
- (64) Canchi, D. R.; Paschek, D.; García, A. E. *J. Am. Chem. Soc.* **2010**, *132*, 2338–2344.
- (65) Heyda, J.; Kožíšek, M.; Bednárová, L.; Thompson, G.; Konvalinka, J.; Vondrášek, J. i.; Jungwirth, P. *J. Phys. Chem. B* **2011**, *115*, 8910–8924.
- (66) Shao, Q.; Shi, J.; Zhu, W. *J. Chem. Phys.* **2012**, *137*, 125103.
- (67) Deng, N.-J.; Dai, W.; Levy, R. M. *J. Phys. Chem. B* **2013**, *117*, 12787–12799.
- (68) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y. *Science* **2010**, *330*, 341–346.
- (69) Han, W.; Schulten, K. *J. Phys. Chem. B* **2013**, *117*, 13367–13377.
- (70) Meuzelaar, H.; Marino, K. A.; Huerta-Viga, A.; Panman, M. R.; Smeenk, L. E.; Ketelarij, A. J.; van Maarseveen, J. H.; Timmerman, P.; Bolhuis, P. G.; Woutersen, S. *J. Phys. Chem. B* **2013**, *117*, 11490–11501.
- (71) Kannan, S.; Zacharias, M. *PLoS One* **2014**, *9*, e88383.
- (72) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. *Nat. Struct. Biol.* **2002**, *9*, 425–430.
- (73) Byrne, A.; Williams, D. V.; Barua, B.; Hagen, S. J.; Kier, B. L.; Andersen, N. H. *Biochemistry* **2014**, *53*, 6011–6021.
- (74) Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545–614.
- (75) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–802.
- (76) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; Mackerell, A. D., Jr. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
- (77) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., Jr. *J. Comput. Chem.* **2010**, *31*, 671–90.
- (78) Weerasinghe, S.; Smith, P. E. *A. J. Phys. Chem. B* **2003**, *107*, 3891–3898.
- (79) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14* (33–8), 27–8.
- (80) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. *J. Comput. Chem.* **1993**, *14*, 1347–1363.
- (81) Kästner, J. *WIRES Comput. Mol. Sci.* **2011**, *1*, 932–942.
- (82) Guinn, E. J.; Marqsee, S. *J. Mol. Biol.* **2017**, DOI: 10.1016/j.jmb.2017.07.022.
- (83) Bung, N.; Priyakumar, U. D. *J. Mol. Model.* **2012**, *18*, 2823–2829.
- (84) Carrion-Vazquez, M.; Oberhauser, A. F.; Fowler, S. B.; Marszalek, P. E.; Broedel, S. E.; Clarke, J.; Fernandez, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 3694–3699.
- (85) Botello, E.; Harris, N. C.; Sargent, J.; Chen, W.-H.; Lin, K.-J.; Kiang, C.-H. *J. Phys. Chem. B* **2009**, *113*, 10845–10848.
- (86) Cao, Y.; Li, H. *J. Mol. Biol.* **2008**, *375*, 316–324.
- (87) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (88) Grossfield, A. *WHAM: The weighted histogram analysis method*, version 2.0.9, <http://membrane.urmc.rochester.edu/content/wham>.
- (89) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
- (90) Souaille, M.; Roux, B. *t. Comput. Phys. Commun.* **2001**, *135*, 40–57.
- (91) Su, P.; Li, H. *J. Chem. Phys.* **2009**, *131*, 014102.
- (92) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506.
- (93) Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. *J. Chem. Theory Comput.* **2011**, *7*, 625.
- (94) Rocco, A. G.; Mollica, L.; Ricchiuto, P.; Baptista, A. M.; Gianazza, E.; Eberini, I. *Biophys. J.* **2008**, *94*, 2241–2251.
- (95) Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Saboungi, M.-L.; Dempsey, C. E.; MacKerell, A. D.; Brady, J. W. *J. Am. Chem. Soc.* **2004**, *126*, 11462–11470.
- (96) Mason, P. E.; Dempsey, C. E.; Vrbka, L.; Heyda, J.; Brady, J. W.; Jungwirth, P. *J. Phys. Chem. B* **2009**, *113*, 3227–3234.
- (97) Mahadevi, A. S.; Sastry, G. N. *Chem. Rev.* **2013**, *113*, 2100–2138.
- (98) Ran, J.; Hobza, P. *J. Chem. Theory Comput.* **2009**, *5*, 1180–1185.
- (99) Egli, M.; Sarkhel, S. *Acc. Chem. Res.* **2007**, *40*, 197–205.
- (100) Jain, A.; Ramanathan, V.; Sankararamakrishnan, R. *Protein Sci.* **2009**, *18*, 595–605.
- (101) Yoon, J.; Thirumalai, D.; Hyeon, C. *J. Am. Chem. Soc.* **2013**, *135*, 12112–12121.
- (102) Ganguly, P.; Mukherji, D.; Junghans, C.; van der Vegt, N. F. J. *Chem. Theory Comput.* **2012**, *8*, 1802–1807.
- (103) Ganguly, P.; van der Vegt, N. F. J. *Chem. Theory Comput.* **2013**, *9*, 5247–5256.

- (104) Krishna, R.; Velmurugan, D.; Babu, G.; Perumal, P. *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.* **1999**, *55*, 75–78.
- (105) Mohan, N.; Vijayalakshmi, K. P.; Koga, N.; Suresh, C. H. *J. Comput. Chem.* **2010**, *31*, 2874–2882.
- (106) Faggi, E.; Luis, S. V.; Alfonso, I. *RSC Adv.* **2013**, *3*, 11556–11565.
- (107) Priyakumar, U. D.; Hyeon, C.; Thirumalai, D.; MacKerell, A. D., Jr *J. Am. Chem. Soc.* **2009**, *131*, 17759–17761.
- (108) Kasavajhala, K.; Bikkina, S.; Patil, I.; MacKerell, A. D., Jr; Priyakumar, U. D. *J. Phys. Chem. B* **2015**, *119*, 3755–3761.
- (109) Miner, J. C.; Garcia, A. E. *J. Phys. Chem. B* **2017**, *121*, 3734–3746.
- (110) Suresh, G.; Padhi, S.; Patil, I.; Priyakumar, U. D. *Biochemistry* **2016**, *55*, 5653–5664.
- (111) Levin, E. J.; Cao, Y.; Enkavi, G.; Quick, M.; Pan, Y.; Tajkhorshid, E.; Zhou, M. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 11194–11199.
- (112) Strugatsky, D.; McNulty, R.; Munson, K.; Chen, C.-K.; Soltis, S. M.; Sachs, G.; Luecke, H. *Nature* **2013**, *493*, 255–258.
- (113) Levin, E. J.; Quick, M.; Zhou, M. *Nature* **2009**, *462*, 757–761.
- (114) Padhi, S.; Priyakumar, U. D. *J. Chem. Theory Comput.* **2016**, *12*, 5190–5200.
- (115) Fradera, X.; Vu, D.; Nimz, O.; Skene, R.; Hosfield, D.; Wynands, R.; Cooke, A. J.; Haunø, A.; King, A.; Bennett, D. J. *J. Mol. Biol.* **2010**, *399*, 120–132.