THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

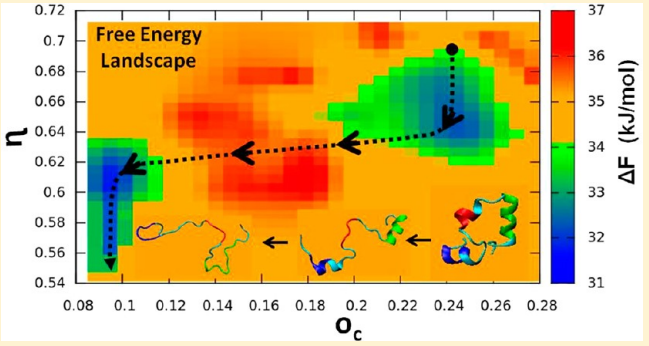
pubs.acs.org/JPCB

Chemical Unfolding of Chicken Villin Headpiece in Aqueous Dimethyl Sulfoxide Solution: Cosolvent Concentration Dependence, Pathway, and Microscopic Mechanism

Susmita Roy and Biman Bagchi*

Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore 560012, India

ABSTRACT: Unfolding of a protein often proceeds through partial unfolded intermediate states (PUIS). PUIS have been detected in several experimental and simulation studies. However, complete analyses of transitions between different PUIS and the unfolding trajectory are sparse. To understand such dynamical processes, we study chemical unfolding of a small protein, chicken villin head piece (HP-36), in aqueous dimethyl sulfoxide (DMSO) solution. We carry out molecular dynamics simulations at various solution compositions under ambient conditions. In each concentration, the initial step of unfolding involves separation of two adjacent native contacts, between phenyl alanine residues (11−18 and 7−18). This first step induces, under appropriate conditions, subsequent separation among other hydrophobic contacts, signifying a high degree of cooperativity in the unfolding process. The observed sequence of structural changes in HP-36 on increasing DMSO concentration and the observed sequence of PUIS, are in approximate agreement with earlier simulation results (in pure water) and experimental observations on unfolding of HP-36. Peculiar to water-DMSO mixture, an intervening structural transformation (around 15% of DMSO) in the binary mixture solvent retards the progression of unfolding as composition is increased. This is reflected in a remarkable nonmonotonic composition dependence of RMSD, radius of gyration and the fraction of native contacts. At 30% mole fraction of DMSO, we find the extended randomly coiled structure of the unfolded protein. The molecular mechanism of DMSO induced unfolding process is attributed to the initial preferential solvation of the hydrophobic side chain atoms through the methyl groups of DMSO, followed by the hydrogen bonding of the oxygen atom of DMSO to the exposed backbone NH groups of HP-36.



1. INTRODUCTION

Folding and unfolding of polymers[1] and biopolymers[2−7] pose highly interesting problems that have been studied intensely by different experimental and theoretical techniques, including single molecule spectroscopy, usual kinetic methods and computer simulations. Interacting homopolymers, such as poly[2-methoxy-5-(2′-ethylhexyloxy)-p-phenylene-vinylene] (MEH-PPV), fold into an ordered defect structure that has many characteristics of the native state of a protein.[1] Barbara and co-workers made seminal contributions to this area of research that initiated a paradigm shift in our understanding of polymer folding.[1]

Characterization of protein folding/unfolding pathways has made great strides in the last one decade but actually required many decades of studies to arrive at comprehensive level of understanding owing to the complexity of the problem. Even now many aspects have remained ill-understood. Soon after the unraveling of Levinthal paradox that pointed out conceptual difficulty in searching a native folded state at biologically significant time, efforts were made to understand the folding pathway through constructing mathematical models that could significantly reduce Levinthal's time.[2−4] While experimental, theoretical, and simulation studies have been exploited in elucidating the mechanisms underlying the protein folding process,[8] it is found to be a daunting task to explore and map out the full folding pathway at an atomic level. Therefore, the statistical energy landscape approach recently employed has been quite successful in explaining a limited number of sequence dependent specific folding pathways in different thermodynamic conditions and in understanding the energetics of different protein conformations.[6]

In nature a protein is frequently found to convert its folded native state to the unfolded states through a number of partially unfolded intermediate states (PUIS) separated by relatively low energy barriers. These intermediates are rather difficult to trap, as they are short-lived, with lifetimes of the order of a few ns only. To study the essential biological function of a protein majority of research attention has naturally been focused on the native conformation of a protein than on its corresponding unfolded state. Further, the extraordinary diversity of conforma-

Special Issue: Paul F. Barbara Memorial Issue

Received: August 29, 2012
Revised: November 18, 2012

 © XXXX American Chemical Society dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

tional states of an unfolded protein makes it extremely difficult to trace.[9,10] Nevertheless, the unfolded state of a protein is also important because they are the main precursor of folding process, play significant role in conformational stability, protein degradation, protein turnover and biological transport phenomena through membranes.

Solvent conditions (pH, composition) as well as temperature and pressure play important role in determining the energy of the protein as a function of its configuration. In the protein folding process, water is not only crucial for correct folding of proteins, but also essential in stabilizing the folded state.[6] It is worth mentioning here that both thermal and chemical denaturation processes are associated with interruption of the spanning water network around the protein. Structural transformation and also hydration change have a combined effect on the free energy associated with protein folding/unfolding.[11] The associated free energy receives both enthalpic and entropic contributions. Thermal denaturation is clearly driven by the larger entropic contribution at elevated temperatures. However, although in real experiments even a few degrees of temperature change cause denaturation, in simulation, one often has to raise the temperature by unphysically large amount (even by 350 °C) to induce thermal denaturation. In chemical denaturation, unfolding can be induced in milder conditions by using appropriate cosolvents, as enthalpic contribution that plays an important role here[12] can be systematically modified at room temperatures.

A verity of solvent additives or cosolvents has been exploited to modulate various structural and functional properties of a protein. These cosolvents not only stabilize native-like secondary structures, but also transform some structural moieties to their nonnative states. Some denaturants can destabilize, unfold, or denaturate the protein chain. Among them, glycerol or trimethylamine N-oxide can stabilize the folded state. On the contrary, guanidinium chloride (GdmCl) and urea act as a protein denaturants.

Interestingly, among such cosolvents, dimethyl sulfoxide (DMSO) is unique. It can play a role as a stabilizer, an activator, a denaturant, an inhibitor, and also as a cryoprotector. In drug discovery processes, DMSO is commonly utilized to prepare the standard solvent for preparing stock solutions of compounds. Furthermore, moderate concentrations of DMSO have been found to enhance enzyme activity by accelerating the conformational flexibility of the protein. There are only a few reports on how low concentrations of DMSO affect a protein. In a recent article, Yang et al. show a distinct change in the unfolding pathway of dimeric bacterial NAD+ synthetase upon adding DMSO to a concentration of 2.5% (V/V).[13] By UV circular dichroism spectral study, Balaram and co-worker detected a partially folded state of hen egg-white lysozyme in 50% DMSO (~18% mole fraction) during the unfolding process.[14] In a previous study, we also identified the experimentally observed DMSO concentration dependent intermediates and explained the microscopic origin of emergence of such intermediates in terms of protein−cosolvent interaction.[15] The tertiary fold of lysozyme is found to become unfolded at concentration around 30−40% mole fraction of DMSO.

Solvation dynamics has also been proven to be a useful tool to characterize both the native and the unfolded states of a protein. Very recently, dynamics of aqueous folded and guanidine induced unfolded states have been studied by Fayer and co-workers by a series of two-dimensional infrared vibrational echo experiments performed on a nitrile-labeled villin headpiece.[16]

For the past decade, attempts have been made to understand the microscopic basis of protein denaturation mechanism and the associated effects of GdmCl and urea on the protein structure and stability. Recently it has been reported that urea is preferred over water for binding to the protein backbone and to the charged side chains through hydrogen bonds, and it drags the protein to unfold or denaturate.[17] On the contrary, in the case of Gdmcl, the interaction between positively charged guanidinium ions with hydrophobic surfaces is the principal driving force for denaturation.[18] Thus, all these well-known denaturants denature proteins by different mechanisms.[19] Despite the existence of such diverse mechanisms, in recent years numerous spectroscopic studies have observed that a number of proteins appear to retain significant residual contacts, even in strongly denaturing conditions.[20] For a globular protein the stabilization of a native folded structure is primarily governed by the favorable interaction between hydrophobic side chain residues, specific hydrogen bonds formation, salt bridge construction, and so on. Nevertheless, a number of nonspecific interactions such as, nonpolar backbone contacts have immense importance in formation of a collapse state during the process of protein folding. In fact, the interaction of chemical denaturants with protein seems to utilize this nonspecific interaction.[21]

The possibility of diverse mechanisms and different unfolding pathways (characterized by appearance of different sequence of intermediate states) in different cosolvents motivated us to study the DMSO-induced chemical unfolding of proteins. To approach the detailed investigation at a molecular level, we have carried out molecular dynamics simulation studies of model systems such as the 36-residue villin headpiece protein (HP-36). The detailed information about this protein has been described in our previous article.[22] The well packed hydrophobic core makes HP-36 an ideal model protein bearing the characteristics of a large protein.

The present simulations reveal several interesting results. Here we report potentially significant correlations between the DMSO concentration dependent dynamical evolution of the trajectories from the native to unfolded state of HP-36 and analyze the subsequent changes with changing solvent configuration as well. At the critical DMSO concentration for unfolding, we have studied the time evolution of native contact pair formation, sequence dependent contact distance, and the other important order parameters. The time evolution of each order parameter of the protein has been found to follow consistent multistage dynamics. Most importantly, we found that the dynamics of cosolvent molecules is strongly coupled with the time evolution of the unfolding process. A number of metastable states has been detected that have quite resemblance with the intermediates found in thermal denaturation process of the same protein studied by Bandyopadhyay et al.[23] Moreover, we have reported a molecular mechanism of DMSO-induced unfolding process based on protein−cosolvent interaction.

The first step of unfolding (separation of Phe18 from Phe11 and Phe7) appears to be common to other ways of unfolding, such as thermal denaturation.[23] In the present case of chemical denaturation, the hydrophobic interactions between clusters formed by Phe18, Phe11 and Phe7 are weakened due to solvation by the hydrophobic methyl groups of DMSO and entropic driving force helps them to separate in space. The

dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

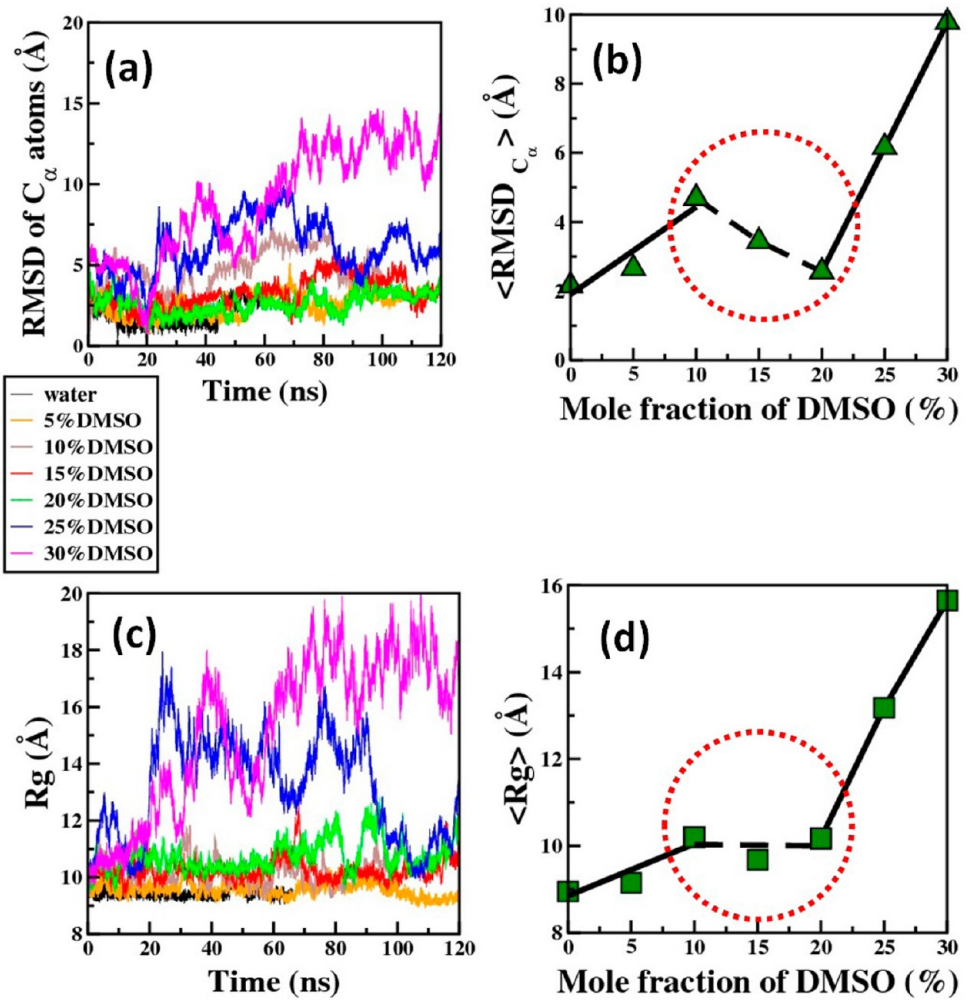


Figure 1. Conformational fluctuation and unfolding of HP-36. (a) The time evolution of backbone (C_α) RMSD and (b) the equilibrium average values of C_α RMSD at several compositions of aqueous DMSO solution. Note the partial unfolding region (0−10%) and the collapsed region (10−20%) together bear the signature of nonmonotonic concentration dependent conformational fluctuation of HP-36. (c) Time variation in radius of gyration (R_g) and (d) the equilibrium average values of R_g at different DMSO concentrations. Due to the sudden increase of C_α RMSD and R_g at 25−30% DMSO mole fraction, this concentration range has been identified as the critical unfolding concentration region for HP-36. The color code for different DMSO mole fraction is indicated in a legend box.

effect of the methyl groups can be understood at least partly by considering the potential of mean force between two phenyl alanine groups quantified in several earlier studies.[24] Subsequent steps of unfolding are clearly coupled to this initial step where the second step is the separation between Ala17 and Leu29 and Ala9 and Leu35. Separation of these long distance tertiary contacts signifies a significant expansion of the protein. The last stages are the melting of the helix-1 and helix-3, respectively.

The rest of the article is planned as follows. In section 2, system setup and simulation details are discussed. Composition dependent dramatic variation in protein properties in aqueous DMSO mixture are discussed in section 3. Section 4 includes unfolding pathway at critical DMSO concentration required for unfolding. In section 5, we develop theoretical analyses of unfolding phenomenon. Section 6 explains the mechanism of DMSO induced unfolding of HP-36. In section 7, we discuss the unfolding aspects in correlation with solvation dynamics studies. Last of all, section 8 concludes with a brief summary of results.

2. SYSTEM SETUP AND SIMULATION DETAILS

We performed molecular dynamics (MD) simulations of a protein in water−DMSO binary mixture at several DMSO compositions by using the GROMACS Package (version 4.0.5). The simulation began with the crystal structure of the HP-36 obtained from the NMR structure of the villin headpiece subdomain as reported by McKnight et al.[25] The initial coordinates were collected from the Protein Data Bank (PDB ID: 1VII). The two ends (Met1 and Phe36) of the protein were capped properly. We used SPC/E model for water molecules.[26] Methyl groups of DMSO were modeled as united atom within GROMOS96 53A6 force field.[27] Numerous evaluations suggest that GROMOS force field can be successfully applied for simulating several biomolecular systems including a number of solvents such as water, chloroform, methanol, dimethyl sulfoxide, carbon tetrachloride, and so on. Oostenbrink reported that the united atom model of DMSO combined with the GROMOS force field is reliable to produce any physical properties of liquid DMSO, including rotational correlation time, thermal expansion coefficient, isothermal compressibility, specific heat, excess Helmholtz free energy,

dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

static dielectric permittivity, and shear viscosity, to name a few. All results were in good agreement with experiments.[27]

In the present work at first we prepared water−DMSO binary mixture at various concentrations in cubic boxes, with sides 3.0 nm. After steepest descent energy minimization, each trajectory was propagated in a NVT ensemble and equilibrated for 2 ns. All the simulations in this study were done at 300 K and 1 bar pressure. The temperature was kept constant using the Nose-Hoover thermostat.[28] It was followed by a NPT equilibration for 20 ns using the Parinello−Rahman barostat.[29] After preparing the binary solvent at various concentrations, protein was dissolved to each of them and again followed the same procedure of energy minimization. A total of around 6000 solvent molecules were taken. To further equilibrate the solvent before starting a full molecular dynamics simulation, we hold the protein fixed while allowing the solvent to move around at constant temperature by performing position restrained molecular dynamics for 5 ns. This allows the solvent to relax to a state which is natural for the current (native) conformation of the protein. Finally, production runs were performed for each system in NPT ensemble. All the results were extracted from the 120 ns trajectory. The box size is enlarged to 6.7 nm to accommodate all the molecules. Periodic boundary conditions were applied and nonbonded force calculations were employed in a grid system for neighbor searching.[30] Neighbor list generation was performed after every 10 steps using a cutoff 0.9 nm. A cutoff radius of 1.2 nm was used for van der Waal's interaction. To calculate the electrostatic interactions, we used PME37 with a grid spacing of 0.136 nm and an interpolation order of 4.

3. COMPOSITION-DEPENDENT VARIATION OF PROTEIN STRUCTURE

As mentioned earlier, many properties of the water−DMSO binary mixture exhibit anomalies in the 10−20% concentration range.[31] These anomalies have been attributed to an aggregation phenomenon that is driven by both the hydrophobic interactions between the methyl groups and the strong hydrogen bonding ability of the oxygen atom of >S═O group. The aggregation of methyl groups influences the unfolding scenario because near the aggregation driven structural transformation, the methyl groups remain engaged in the self-association and are not readily available for solvation and hydrophobic interaction with other groups of protein. However, at higher concentrations, the number of methyl groups becomes sufficient to participate in such interactions. The presence of this structural transformation gives the composition dependence of unfolding in water−DMSO mixture, novel characteristics that are investigated in this article.

3.1. Backbone Structural Changes via Cα-RMSD. A routine checkup to analyze globular protein conformations often begins with the measurement of Cα RMS deviation from its corresponding crystal structure. Figure 1a shows the time progression of Cα RMSD of HP-36 at different mole fraction of DMSO. Concentration dependence of conformational fluctuation becomes more interesting when we compare the equilibrium average value of the RMS deviation at various DMSO compositions (see Figure 1b). Note that starting from the initial addition of DMSO to 10%, the RMSD value gradually increases suggesting the continuous partial expansion of the protein structure. Such continuation of the unfolding process is interrupted in the concentration range of 10−15% mole fraction of DMSO. In this range, protein goes to a more compact structure. Further addition of DMSO again forward protein's sojourn toward unfolding. Finally, at 30% mole fraction of DMSO protein structure attains its equilibrium unfolded conformation.

3.2. Radius of Gyration. The most useful measurement is radius of gyration (R_g) of a protein to monitor the large structural variation which indeed offers a straightforward indication for protein unfolding. Therefore, we have calculated the radius of gyration of HP-36 in different solvent compositions and presented in Figure 1c. The equilibrium average value of R_g is also plotted to check the concentration-dependent unfolding pattern (see Figure 1d). The variation found here is consistent with the same nonmonotonic variation as found in C_α RMSD. At 30% mole fraction of DMSO, the value of R_g raises with a large fluctuation, which provides an elegant evidence for unfolding. The evaluated RMSD as well as R_g of unfolded protein is well consistent with the early simulation results.[32]

3.3. Unfolding Characteristics and Partially Unfolded Intermediate States (PUIS). The progression of DMSO induced unfolding of HP-36 triggers with the deformation of Helix-2 region where the congregated hydrophobic patch involving Phe7, Phe11, and Phe18 starts to become disconnected even at 5% mole fraction of DMSO. This initial event is found to be similar as one observes in thermal denaturation or Gdmcl-mediated unfolding processes.[23,16] The structural transition from 10 to 20% mole fraction of DMSO is responsible for generating a partially unfolded intermediate as is observed earlier in many experiments and simulation processes.[14,15] Study of the dynamics of the folded and unfolded villin headpiece with ultrafast 2D IR vibrational echo spectroscopy also suggests that unfolding moves the nitrile-functionalized phenylalanine group from the interior of the folded peptide to the solvent-exposed exterior.[16] It is evidently observed from the snapshots that there is no significant structural changes occurs from 10−20% mole fraction of DMSO (see Figure 2). It seems that in this concentration ranges protein structure becomes highly restricted. However, at 30% DMSO concentration, the emergence of unfolding conformation is clearly distinguished from the other prevailing conformations.

3.4. Native Contact Dynamics. Formation or breakage of a specific contact between two residues of a polypeptide chain is an essential elementary process in protein folding/unfolding study. Energetic stabilization of the folded native state of a protein is driven by the making of appropriate, often hydrophobic contacts.[33] Here we study the time evolution of the native contact exclusion/inclusion dynamics at different DMSO concentrations. A considerable amount of correct native contact formation is diminished even in aqueous media. But further deviation from the correct native contact arrangements up to 10% DMSO concentration strongly supports the proposition of partial unfolding in the protein conformation and also the destabilization phenomenon. The loss of native contacts beyond 10% of DMSO concentration shows that in this concentration region protein attains a collapsed conformation. It is better to say that in these concentration ranges protein structure is dramatically inhibited to unfold. In fact the hydrophobic core packing does not deviate much from the native structure until the DMSO concentration reaches beyond 20%. At 30% mole fraction of DMSO, the unfolding turns out to be more gradual and to this end we reach to the finer structure of unfolded protein.

dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

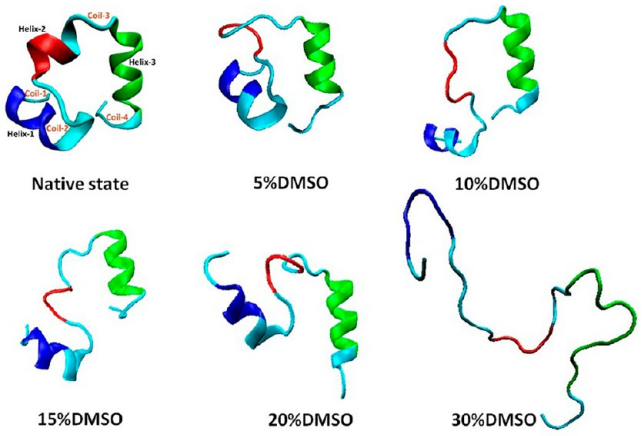


Figure 2. Concentration-dependent equilibrium structures of HP-36 at various DMSO mole fraction. Progression of unfolding from native structure of HP-36 to the extracted equilibrium conformation at different mole fraction of DMSO has been demonstrated. Different structural regions of HP-36 are indicated in the native structure are as follows: Helix-1 (4−8 amino acid residues colored as blue), Helix-2 (15−18 amino acid residues colored as red), Helix-3 (23−32 amino acid residues colored as green), and others are Coil-1, Coil-2, Coil-3, and Coil-4 (colored as cyan). The 1st event of unfolding is the unfastening of Helix-2 initiated even at 5% mole fraction of DMSO. The concurrent separation of hydrophobic core patches results in different conformations of partially unfolded intermediate states (PUIS) from 10%. At around 10−20% the structure becomes inflexible, rather restricted to unfold. The randomly coiled unfolded structure of HP-36 is obtained at 30% DMSO mole fraction.

3.5. Protein−Solvent Interaction Energies. To understand the role of cosolvent concentration dependence in driving the unfolding transition, we monitor the average protein−protein, protein−solvent, and the overall combined interaction energy at each concentration of DMSO. The concentration-dependent nonmonotonic behavior correlates well with the other protein properties. This strong correlation between the protein and solvent dynamics strongly signifies that the solvent organization plays an important role in stabilizing structures and assists the protein to follow a minimum energy unfolding pathway. In Figure 4a, we observe, until 10% DMSO concentration, the protein−protein interaction energy increases, and the protein−solvent interaction energy decreases (see Figure 4b) that suggest the structure changes to acquire a partially unfolded state assisted by the lowering of protein−solvent interaction energy. From 10 to 15%, the diminished protein−solvent interaction energy restricts the structure toward unfolding. Beyond 15%, protein−solvent interaction energy really drags the system to unfold. It is important to note here that the unfolding process is largely driven by the protein−solvent interaction energy, which ultimately stabilizes the unfolded state. Figure 4c clearly shows that protein−solvent interaction energy wins as DMSO concentration reaches to its critical concentration (30% mole fraction of DMSO) for unfolding.

3.6. Analysis of Nonmonotonic Behavior in Protein Structural Properties. In an earlier simulation-based theoretical study, the origin of the observed significant changes in both structure and function of lysozyme on addition of a small amount of DMSO in aqueous solution was extensively analyzed. The observed changes were attributed to the preferential solvation of exposed hydrophobic residues of the protein by methyl groups of DMSO.[15]

However, this preferential solvation sustains only up to 10% mole fraction of DMSO concentration. As a consequence, in this range, one finds, with increase in DMSO concentration, an enhanced RMSD, enlarged radius of gyration, loss in native contacts, and increased protein−solvent interaction energy.

Water−DMSO binary mixture is found to exhibit rich dynamics due to a weak, percolation driven structural phase transition at around 10−20% DMSO concentration region. DMSO molecules form clusters through the aggregation of methyl groups due to an intervening phase transition at around 15% DMSO mole fraction.[31] This structural phase transition plays an inevitable role in the solvation of the hydrophobic amino acid residues of the protein. At this range, DMSO concentration close to the hydrophobic surface decreases. As a result, a large number of hydrophobic residues are sequestered

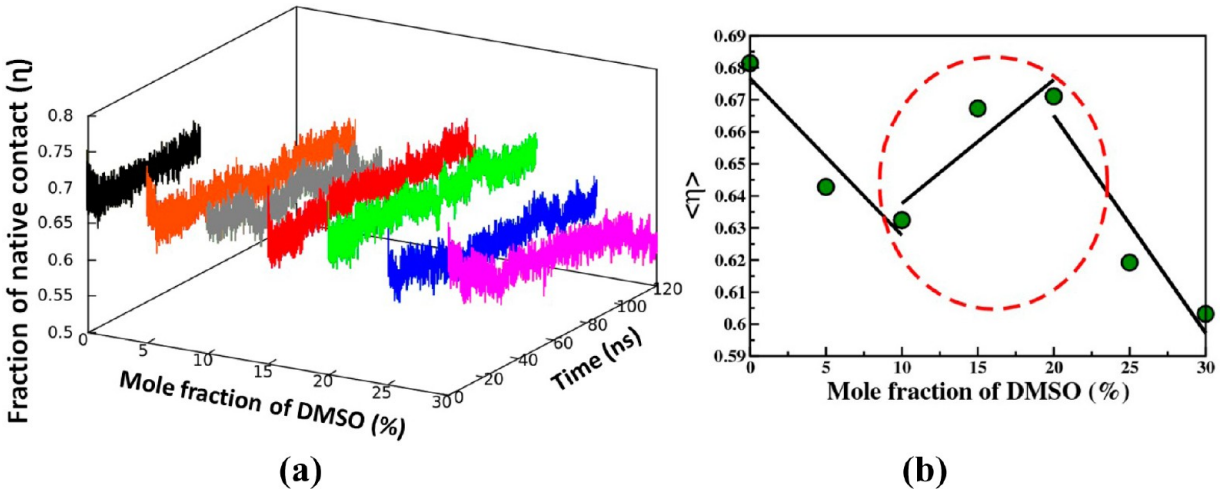


Figure 3. Fraction of native contact (η) dynamics and DMSO concentration dependence. (a) Time trajectories of fraction of native contact of HP-36 in the aqueous environment and with change in mole fraction of DMSO are plotted. Contact cutoff is taken as 4 Å. (b) Equilibrium average of fraction of native contact. Loss of fraction of native contact increases from 0 to 10% mole fraction of DMSO indicates the expenses of more stability and inducing the partial unfolding than its equilibrium conformation in aqueous media. Note, the same trend preserves as we observe in C_α-RMSD and R_g.

dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

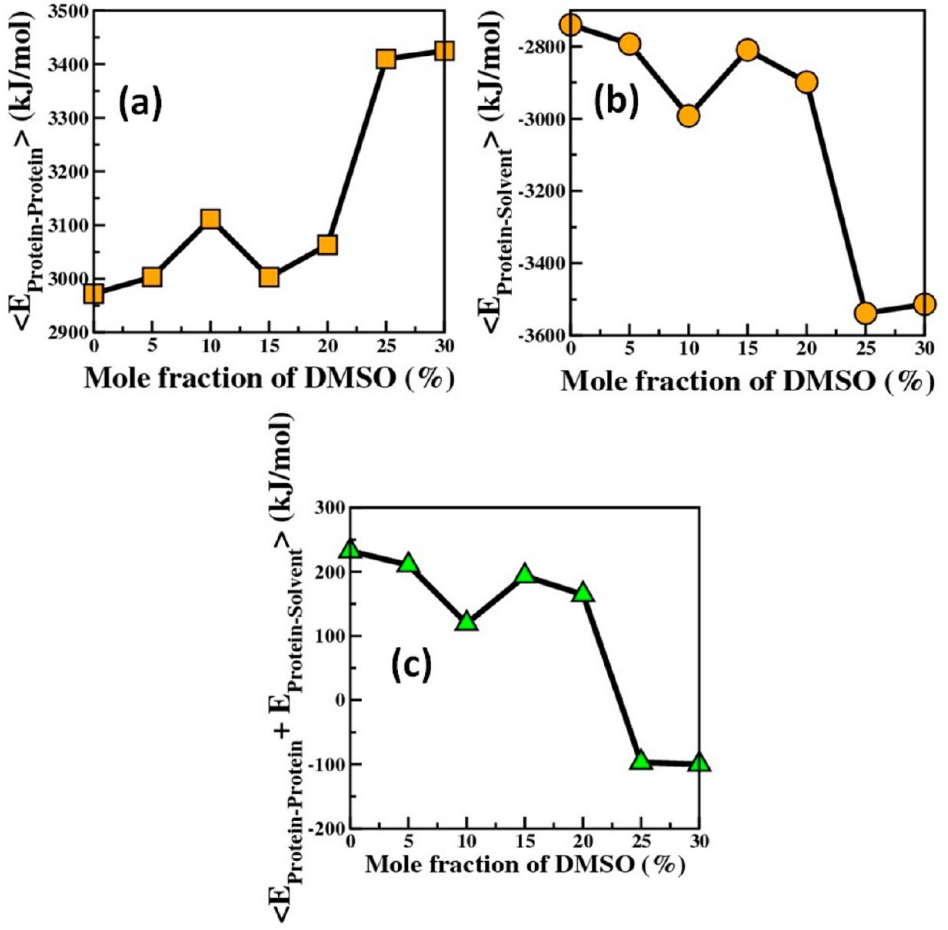


Figure 4. Nonmonotonicity in protein−protein and protein−solvent interaction energy with variation in DMSO concentrations. Average interaction energy between (a) protein−protein, (b) protein−solvent, and (c) their overall sum at different DMSO concentrations. The competition between protein−protein and protein−solvent interactions largely affect the organization of the amino acid chain. Ultimately protein−solvent interaction energy wins to favor the unfolding process.

in that core structure and the event is named as "hydrophobic zipping".[14] The fluctuation of some residues (such as, Phe11, Phe18) gets quenched. Hence, RMSD falls, radius of gyration decreases, some native contacts regenerate, and protein−solvent interaction energy diminishes. Around 25−30% or above it follows its way to denaturation by continuous unfolding. As a result increase in C_α RMSD, gradually enlarged R_g, loss in native contacts, and stronger protein−solvent interaction energy are observed along with the increase in DMSO content in the hydration layer.

4. UNFOLDING PATHWAY AT 30% DMSO CONCENTRATION

4.1. Multistage Dynamics Captured in Fraction of Native Contacts. The partial unfolding intermediates in unfolding process have been identified in several experiments and simulation studies. Current simulation has not only detected those intermediates but also analyzed the complete unfolding pathway in terms of solute−solvent interaction. The time trajectory of the fraction of native contact at 30% mole fraction of DMSO reveals several interesting results. We find that the unfolding trajectories of HP-36 exhibit multistage dynamics (see Figure 5). Despite the initial fast decay of the fraction of native contacts, the hydrophobic core contacts remains largely unaltered for a very long time. Such "hydrophobic locked structure" protects the protein from achieving a globally unfolded state and emerge a "native-like compact structure" (see snap2 in Figure 5). At the end of this initial stage of rapid native contact loss, second helix deforms as consequence of separation between Phe11 and Phe18 residues that serve as a hot contact to initiate the unfolding process (see snap3). This step is followed by long plateau regions starting from 30 to 90 ns where interestingly native contacts fluctuate without any significant hop. But at the same time, some hydrophobic important contacts separately form hydrophobic patches (see snap4). These events can be termed as hydrophobic flocculation. As a consequence, native-like partly extended states are observed and hence the long plateau region is ended with a partially unfolded intermediate. The breaking of such hydrophobic patches finally approaches to the fully unfolded state.

4.2. Average Sequence Distance between the Nonlocal Hydrophobic Contacts. The hydrophobic flocculation is driven by the breaking of the hydrophobic core contacts. In protein folding study the often used average sequence distance between the nonlocal hydrophobic contacts is the appropriate one to check the hydrophobic compact state. In the same way to understand the unfolding phenomena we need to study the dynamics of formation of long-range hydrophobic contacts.[34] On such basis we have used relative contact order parameter, O_c, which reflects the relative importance of local and nonlocal

dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

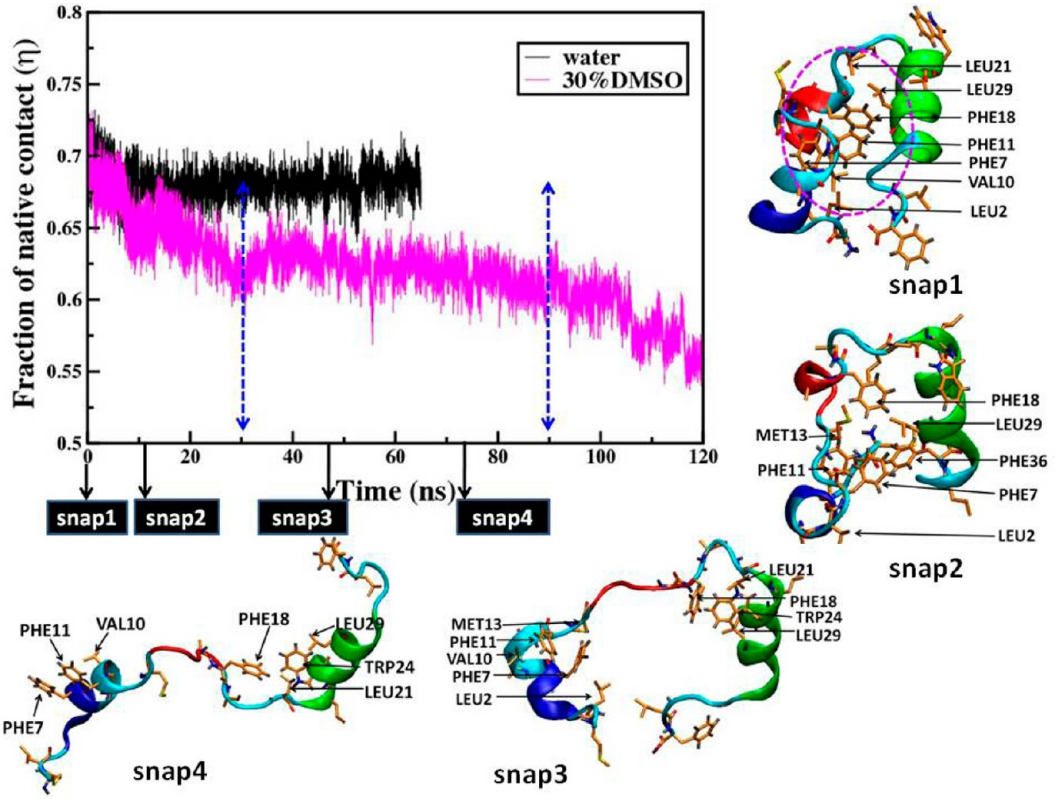


Figure 5. Multistage dynamics of HP-36 at 30% mole fraction of DMSO and comparison with the same in water. Time evolution of the fraction of native contact (η) and multistage dynamics during unfolding of HP-36 is shown at 30% aqueous DMSO solution. A clear comparison with the dynamics of water has been characterized. The multistage dynamics evolves with a number of intermediate structures. The corresponding intermediates are demonstrated accordingly they appear in individual stages. A series of hydrophobic flocculation events are observed from snap 2 to snap 3−4 to trigger the overall unfolding process.

hydrophobic contacts. Relative contact order is the average sequence distance between all pairs of hydrophobic contacting residues normalized by the total sequence length:

$$O_c = \frac{\sum_{i,j}(s_j - s_i)}{LN_c} \quad (1)$$

where (i,j) denotes the specific hydrophobic pair contacts, N_c is the number of hydrophobic contacts, and L represents the total number of hydrophobic amino acids presents in the system; s_i and s_j are the sequence number of the corresponding hydrophobic amino acids along the protein chain.[35]

The dynamics of relative contact order parameter (O_c) plotted in Figure 6 gives further insight into the dynamics of unfolding of HP-36. The multistage unfolding dynamics observed in the previous figure for the time evolution of energy seems also to sustain here. Note that, O_c is the average sequence separation. So, an increase in O_c signifies the formation of long-range contacts. The plot provides three kinds of intermediate structures driven by such long-range contact dynamics: (1) the initial preservation of the hydrophobic topological contact results hydrophobic compact structures (indicated by "1" in Figure 6); (2) the subsequent hydrophobic contact loss results in hydrophobic partly extended states (indicated by "2" in Figure 6); (3) the late stage transition provides hydrophobic extended states (indicated by "3" in Figure 6).

4.3. Pathway Search. To investigate DMSO-induced exact unfolding pathway of HP-36, we have plotted fraction of native

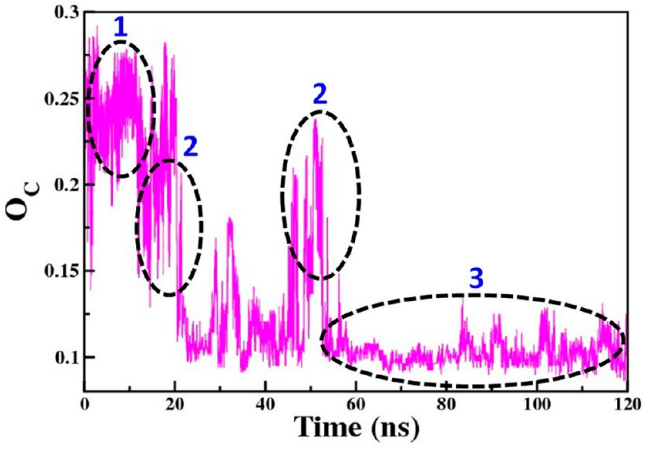


Figure 6. Average sequence distance between hydrophobic contacts (O_c) at 30% mole fraction of DMSO. The different regions in the time evolution of this relative contact order parameter are allocated as (1), (2), and (3), corresponding to the different intermediates shaped and tailored by the long-range hydrophobic contacts. The three intermediates are as follows: (i) hydrophobic compact structures (indicated by "1"), (ii) hydrophobic partly extended ("2"), and (iii) hydrophobic extended states ("3").

contact (η) vs O_c to check their simultaneous variation at explicit time (see Figure 7). The plot provides a deep insight for better understanding of the mechanism. This permits us to present a five state model that includes (1) native (N), (2) near-native hydrophobic compact (NNHC), (3) near-native

dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

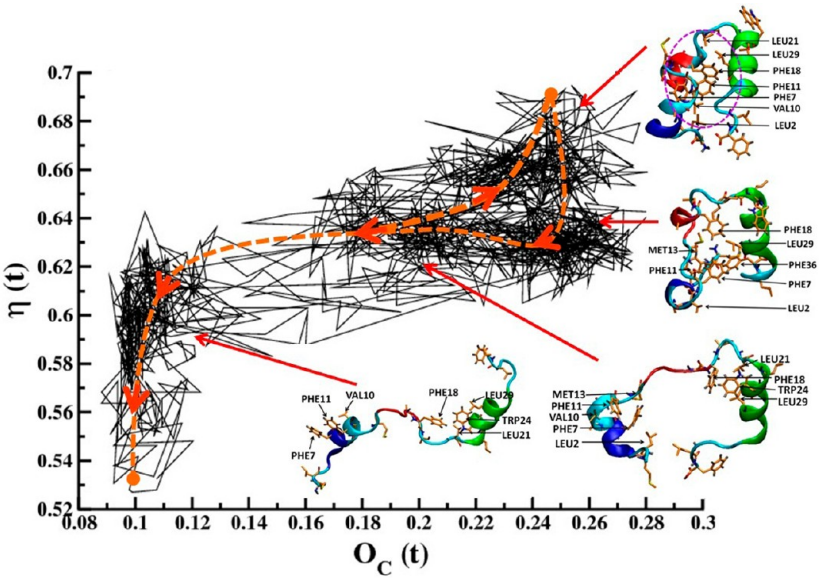


Figure 7. Protein's conformational degrees of freedom visited along the unfolding pathway. The *X*-axis and *Y*-axis are depicted as an average sequence distance between hydrophobic contacts ($O_c(t)$) and the fraction of native contact ($\eta(t)$), respectively, at their explicit time. The most visited conformational spaces allow us to derive a probable pathway that might be the dominant lower energy path followed by HP-36 in the denaturating condition. The orange arrowed dashed line is the schematic guide of the most probable pathway involving the following five states that determine the dominant path of HP-36 unfolding: (i) native (N), (ii) near-native hydrophobic compact (NNHC), (iii) near-native hydrophobic partially extended (NNHPE), (iv) near-native hydrophobic extended (NNHE), and (v) nonnative fully unfolded states (U).

hydrophobic partially extended (NNHPE), (4) near-native hydrophobic extended (NNHE), and (5) nonnative fully unfolded states (U). The fully unfolded snap is not shown in Figure 7. The first three intermediate states account for near-native fluctuations but solely driven by the long-range hydrophobic contacts. Emergence of these intermediates allows us to sketch a qualitative property based free energy path (see Figure 8). The free energy path discloses more important information.[36] It also provides the low free energy separation between the folding and unfolding minima. The below contour plot indicates the transition from native folded state to unfolded structure through minimum energy pathway. The highly populated native circle suggests that the initial structures are rattling around the stable native state. We identify that the ensemble of near native configuration are often visited in any denaturating condition. Temperature unfolding and also studies by using GdmCl denaturant have observed such visits. Hence, one can conjecture that the initial protein unfolding steps in any denaturating condition often proceeds by following a minimum energy path, which appears to be dominant among the multiple pathways of unfolding.

4.4. Correlated Protein−Solvent Dynamics. The specific unfolding pathway and the corresponding conformational changes are controlled by the denaturant condition, whether it is temperature or any suitable cosolvent/denaturants. To understand the role of such cosolvent in driving the native state to unfolded state, we monitor the time trajectory of fraction of native contact and protein−protein and protein−DMSO interaction energy (see Figure 9). The multistage dynamics in also observed in both the interaction energy evolution and the steps are strongly coupled with the fraction of native contact dynamics. In Figure 9 we can map a one-to-one correlation between the native contact lose and protein−DMSO interaction energy. It signifies that important native contacts are tuned by the surrounding solvent molecules. Most importantly, the native contacts formation or breaking

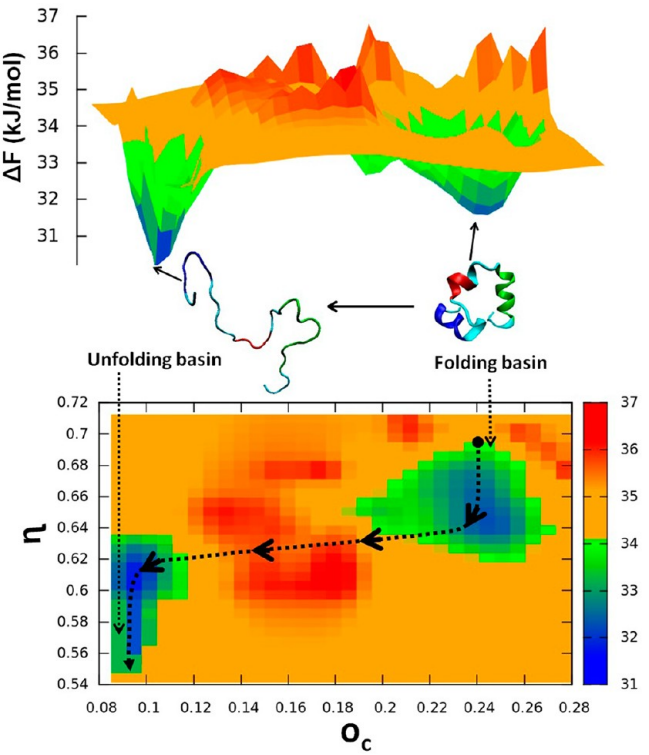


Figure 8. Contact order parameter (O_c, η) based free energy landscape is shown, along with a contour map, for folding to unfolding transition of HP-36 at 30% mole fraction of DMSO. The upper panel shows the presence of two minima that corresponds to the folding and unfolding basins as indicted in the figure. The low free energy separation (<1 kJ/mol) between folding and unfolding minima is evident from the energy landscape. The below contour plot also demonstrates distinguishable population distribution of the folded and unfolded basin. The black arrowed dashed line is the schematic guide to show the most probable path from native folded state to unfolded protein.

dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

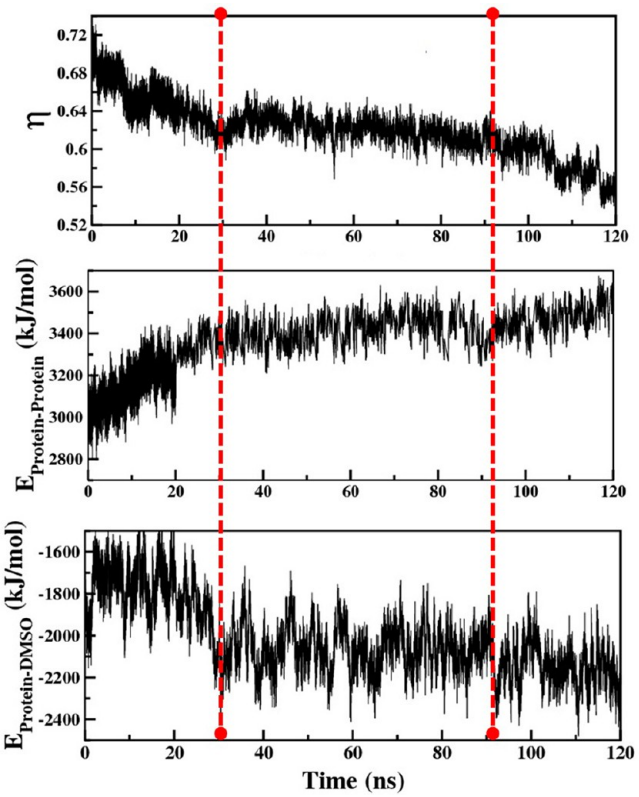


Figure 9. Cosolvent-induced native contact dynamics. Multistage unfolding trajectories showing computation of (a) fraction of native contact (η) (upper panel), (b) protein−protein interaction energy (middle panel), and (c) protein−DMSO interaction energy (lower panel). Different dynamical stages are indicated by the red dividers. These three trajectories propagate simultaneously, which signifies a strong coupling among them. Note that both the loss of important native contacts and protein−protein interaction energy are dominantly controlled by the protein−DMSO interaction energy.

predominantly depends on the arrangement and orientation of DMSO molecules.

4.5. Local Melting and Cooperativity in Unfolding of HP-36. The central hydrophobic core of the native structure of HP-36 contains three phenyl alanine groups. The initial step of DMSO induced unfolding of HP-36 involves the separation of two important hydrophobic native contacts, namely, between Phe11 and Phe18 and between Phe7 and Phe18. These initial separations seem to induce subsequent separation of other hydrophobic native contacts. Therefore, these hydrophobic contact separations can provide significant entropy gain.[37,38] In addition, the potential of mean force of such contact pair plays an important role to process the unfolding.[24]

Contact map presented below shows the stepwise loss of the important contacts (see Figure 10). In the first stage (after 15 ns) we observe that the native hydrophobic core contacts particularly involving Phe7-Phe18 and Phe11-Phe18, Val10-Leu29 are untying to initiate the melting of second helix. In fact, the hydrophobic core deformation and second helix melting are synchronized processes and occur quite simultaneously. In the second stage (starts after 30 ns) of unfolding, several other long distance contacts separate (e.g., 2−35, 9−35, 2−36, etc). This stage also involves expansion of the protein.

The third stage of unfolding is the helix melting processes. The first helix melting process is initiated after the second stage (as evidenced by 2−11, 2−9 contacts separation). At the end of the third stage, helix-3 melting is visible in the separation of 21−29, 29−35 contacts. This essentially completes global unfolding that proceeds in cooperation with the local melting steps.

To monitor such stepwise residual level transitions we evaluate the time trajectory of the distance between few important contacts and observed the first level of local melting process facilitates the second level of unfolding (see Figure 11). For example, in the first stage, the Phe11-Phe18 distance increases by about 32 ns, whereas in the second stage, the long distance contacts (2−35, 9−35) get separated and the protein size expands at around 40 ns. The melting of first helix starts after 50 ns and the final third helix melting process prolongs beyond 110 ns.

4.6. Comparison with Folding Transition. The folding pathway of HP-36 has been investigated by several research groups, mostly in neat water. The landmark simulation study of folding pathway search by Duan and Kollman revealed that the initial phase of folding of villin headpiece subdomain evolved from the unfolded state first by hydrophobic collapse and helix formation. This was followed by conformational adjustments in the second phase.[39] However, this pathway was revised by Lee et al. who reported that the folding was initiated by hydrophobic collapse, after which the concurrent formation of full tertiary structure and α-helical secondary structure is observed.[40] After survey of the folding of 12 structurally diverse proteins, Shaw et al. explained that the folding follows a single dominant route in which elements of the native structure have a high tendency to become ordered from the largely disordered unfolded state.[41] In turn, one can expect the folded to unfolded transition to follow the same reverse pathway as in protein folding at equilibrium condition. Among multiple possible pathways of folding/unfolding, only few routes are preferred for a specific sequential protein. Thus, we find that for HP-36 the unfolding event starts with the disruption of the second helix, followed by first helix and finally third helix. The secondary structure formation of HP-36 initiates with the C-terminal helix (3rd helix) followed by the N-terminal helix (1st helix). The short middle helix is shown to form last.[40] However, in unfolding process, the second helix melting and the tertiary structure unraveling are observed to occur quite concomitantly and in a cooperative way. Then the first helix and, finally, the third helix melting events occur accordingly.

5. THEORETICAL ANALYSIS OF UNFOLDING

Denaturation in water−DMSO binary mixture is a non-equilibrium process that proceeds through a discrete number of intermediates that are characterized by distinct pair separation distances, $\{R_i\}$, between groups that form contacts in the native state, and the fraction (or, number) of native contacts, η. Simulations show that may be (at least for HP-36) only a few such distances are practically important, like the distance in the present case between Phe11 and Phe18, Ala9-Leu35, and Leu2-Phe11. One can thus describe the unfolding process in terms of two sets of order parameter: $\{R_i\}$ and η, where η now contains only those contacts that exclude long distance tertiary native contacts.

Here we discuss how a sequence of unfolding steps can be ordered by cooperative interactions. The free energy barrier of the subsequent step of melting faces a high barrier unless the previous unfolding process is executed. Thus, the free energy surface and the relevant barriers are path-dependent.

dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

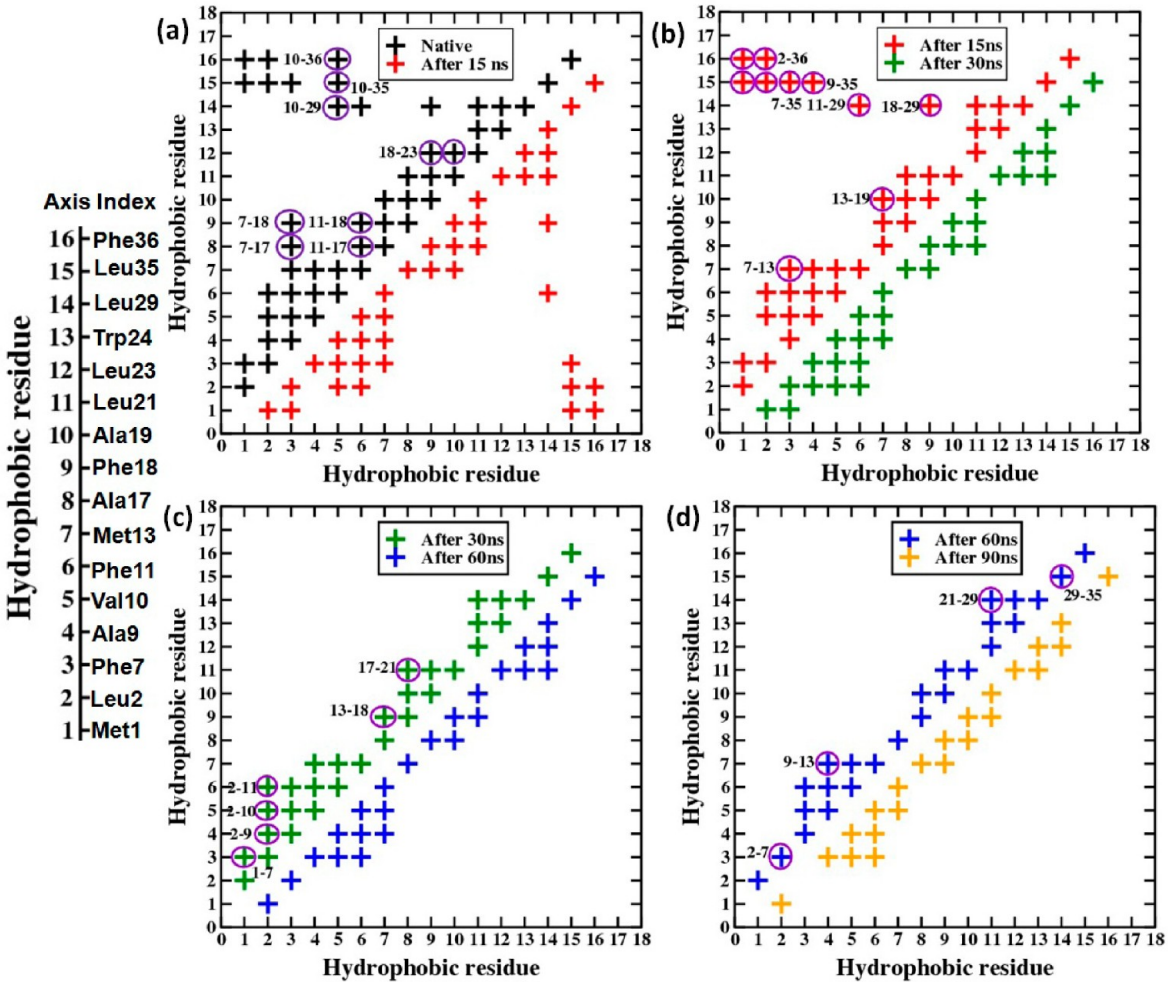


Figure 10. Contact map of HP-36 at 30% DMSO mole fraction, at different simulation time frame, showing evolution of unfolded structure. The missing native contacts from the lower panel of each contact map are indicated in the upper panel (violet circle). Axis index is scaled according to the hydrophobic residue number. (a) First event is observed after 15 ns time steps: 2nd helix deformation starts with Phe11-Phe18, Phe7-phe18 contacts separation. (b) Second event after 30 ns: Long range hydrophobic contacts are observed to disappear. (c) Third event after 60 ns: Hydrophobic contact separation detected largely in the first helix. Additionally, 2nd helix and 3rd helix tertiary contacts break. (d) Fourth event after 90 ns: Hydrophobic contact disconnection mostly identified in the 3rd helix.

Let us further define a probability distribution by $P(\{R_i\},\eta,t)$, that has separation values, $\{R_i\}$, and the fraction of native contacts, η at time, t. The final state (F) is characterized by values $\{R_i^F\}$ and η^F.

As unfolding is a process much slower than solvent motion (or, any fast small amplitude motion of the protein), we can write down the following Smoluchowski equation for the time evolution of the probability distribution, $P(\{R_i\},\eta,t)$:

$$\frac{\partial}{\partial t}P(\{R_i\},\eta,t) = \sum_{i=1}^{N_{sp}} D_i \frac{1}{R_i^2}\frac{\partial}{\partial R_i} e^{-\beta V_i(\{R_i\})}\frac{\partial}{\partial R_i} e^{\beta V_i(\{R_i\})} P - \Gamma_\eta \frac{\partial V(\eta)}{\partial \eta} \quad (2)$$

where we have assumed decoupling between the order parameters. Here D_i is the mutual diffusion coefficient of the two groups involved in the contact pair. This is certainly an approximate description but seems reasonable as a first step. The potential functions, $V(\{R_i\})$ are to be regarded as potential of mean force (PMF) of the type described earlier.[23] Thus, in presence of DMSO, the energy surface, $V(R_{2,11})$ undergoes a sharp change where a minimum at contact $V(R_{2,11})_{min} = a$, for DMSO free solution ($x_{DMSO} = 0$) is replaced by a deeper minimum at a larger separation. Therefore, unfolding at large denaturant concentration is, in general, a nonequilibrium relaxation in a potential energy surface with several traps or activation barriers separating the intermediate states, as shown in Figure 8.

Figure 12 is the energy surface of separation for any one given contact separation. Similarly, potential energy surface of all the separations undergo qualitative changes. If the denaturant is removed ($x_{DMSO} = 0$) the reverse process will start ending in the formation of the native contact when $V(R_{2,11})_{min} = a$. As separation is assumed to occur sequentially, each $V(R_i)$ depends implicitly on other separations.

One can now construct a Marcus-type free energy surface of the separation, which gives the following expression for the activation energy,[42]

$$E_{act} = \frac{[\Delta G^0 + (R_{2,11}^F - a)^2]^2}{4(R_{2,11}^F - a)^2} \quad (3)$$

dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

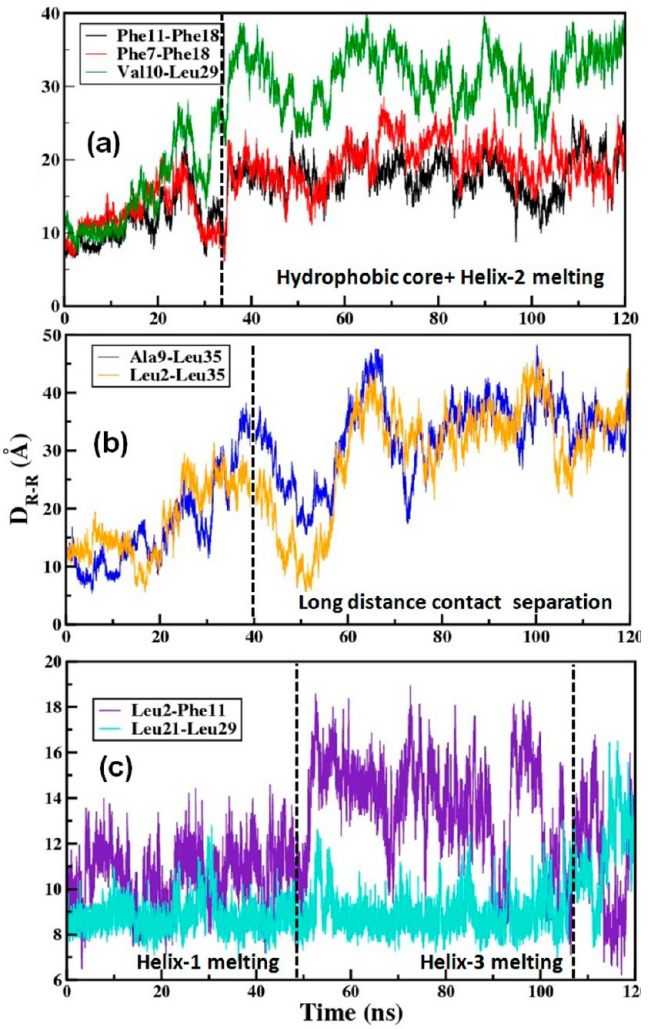


Figure 11. Time trajectory of the distance between important contacts. (a) In the first level of local melting process involving Phe11-Phe18, Phe7-Phe18, and Val10-Leu29, contact pair distances enlarge beyond 30 ns (b) The event facilitates the long-range hydrophobic contacts separation between contact pairs Ala9-Leu35, Leu2-Leu35 after 40 ns. (c) The concurrent event occurs by contact separation of Leu2-Phe11 associated with helix-1 melting and Leu21-Leu29 associated with helix-3 melting.

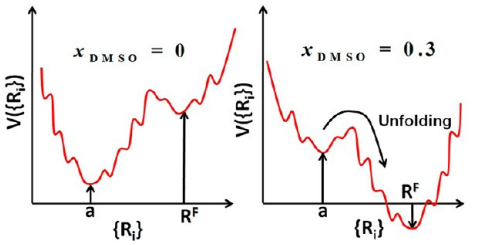


Figure 12. Schematic energy profile of a native contact separation involved in folding-unfolding transition. $V(\{R_i\})$ represents the potential of mean force (PMF) where $\{R_i\}$ are the distinct pair separation distances. $R_i = a$ is the separation in native state while the $R_i = R^F$ gives the separation at the final state of unfolding.

where ΔG^0 is the free energy gap between the two minima and, hence, contains the effect of composition dependence. At large ΔG^0 (at $x_{DMSO} = 0$), we envisage a crossover to barrier less dynamics, as in Marcus theory of electron transfer.[42,43]

The potential energy surface for the native contact fraction, $V(\eta)$ also undergoes significant change as x_{DMSO} is increased. Simulations show that a large change occurs near $x_{DMSO} = 0.3$ when the helices finally disappear. One can therefore understand some aspects regarding the composition dependence of $V(\eta)$. We expect $V(\eta)$ also to depend on the nature of denaturant.

As seen from Figure 8, the largest activation free energy for unfolding is about 5 kJ/mol and this is the opening of contacts between Leu2 and Phe11. The time scale of unfolding as measured by simulations is ~120−140 ns and is determined by sequential processes consisting of three to four distinct steps. This has been indicated in Figure 8.

Adaptation of Marcus theory to the present case allows us to explain the onset of the nonequilibrium process of unfolding in terms of the increase in Gibbs' free energy gap, ΔG^0 with an increase of composition. This in turn leads to disappearance of barriers that holds the native state. At $x_{DMSO} = 0.3$, the unfolding is found to be fast and over by ~120 ns. The slowest steps are the breaking of the first helix and third helix associated hydrophobic contacts such as Leu2 and Phe11 (near helix-1).

In eq 2 the mutual diffusion coefficient is derived from the relative contact separation of the tagged contact pair. An estimate of its value can be obtained by considering both the folded and unfolded states. In Figure 13, we show the velocity auto correlation function of the separation between Leu2 and Phe11. The corresponding mutual diffusion coefficient has been computed in 30% DMSO solution.

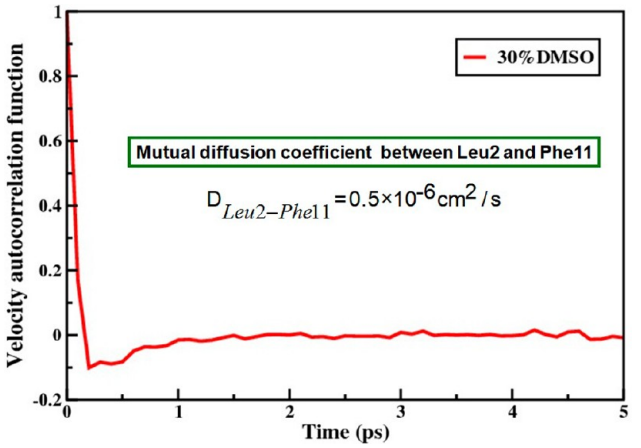


Figure 13. Mutual velocity auto time correlation function (VACF) of the contact pair, Leu2 and Phe11. Leu2-Phe11 contact is associated with helix-1 melting process and is accounted as the slowest step among the other stages of contact break. The mutual VACF of the contact pair is computed in 30% mole fraction of DMSO. The mutual diffusion coefficient is calculated following eq 6.

We can now use this value to obtain an estimate of the time needed for the expansion step in unfolding (stage 4). This is the comparably slower stage than the other stages. In order to obtain a rough estimate we assume that the diffusion during unfolding happens over a rugged free energy landscape which reduces the value of the diffusion coefficient by Zwanzig's expression $D(\varepsilon) = D \exp[-(\varepsilon/k_BT)^2]$.[44] The value of the ruggedness (ε) is of the order of unity. If we now combine all the parameter values to obtain the time needed to diffuse a distance of 7 Å (see lower panel of Figure 11), then we find a value of ~160 ns which is in the right ballpark.

dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

The above analysis however does not include the activated steps, like the initial melting of the second helix needed to start the expansion discussed above. Here we need to take into account the activation barrier of the order of $2-3k_BT$. Here also, we need the value of the mutual diffusion coefficient of the groups involved. One can use the Smoluchowski rate expression to obtain an estimate of the time constant of this process.[45]

$$k_f \equiv k_f^{SL} = \frac{\omega_R \omega_b}{2\pi\zeta} \exp(-E_f/k_BT) \quad (4)$$

where ω_R and ω_b are the frequencies that confine the harmonic reactant well and barrier top respectively. In such Smoluchowski limit the value of the friction coefficient (ζ) is much larger than that of the barrier frequency (ω_b), then rate is predicted to be inversely proportional to the friction. E_f corresponds to the activation energy required to cross the free energy barrier.

The friction along the reaction coordinate has been obtained by using Einstein's relation (DSE) between diffusion and friction,

$$\zeta = \frac{k_BT}{D} \quad (5)$$

Mutual diffusion coefficient (D) has been obtained from the velocity autocorrelation function considering the separation between the contact pair (Leu2-Phe11) following the well-known relation (see Figure 13).

$$D = \frac{1}{3}\int_0^\infty \langle v(0)v(t)\rangle dt \quad (6)$$

The free energy well and the barrier top in the free energy profile are fitted to a harmonic well and then we obtain the harmonic reactant well frequency (ω_R) and barrier top frequency (ω_b), respectively. These evaluations make possible to assess the rate of unfolding of HP-36 in DMSO denaturant media obeying eq 4. The residence time calculated from this theoretical rate estimate is around ~114 ns and correlates favorably well with our simulation results.

As already mentioned, at 30% DMSO concentration, unfolding is a nonequilibrium process with a definite initial and a fairly definite final configuration. As this process involves large amplitude motions of bulky groups, akin to what is often seen in isomerization processes, we expect unfolding rate to be strongly viscosity dependent and should, in fact, exhibit inverse viscosity dependence.[46] It should be possible to study this dependence.

6. MICROSCOPIC MECHANISM OF DMSO-INDUCED UNFOLDING PROCESS

In a temperature-induced unfolding process, Bandyopadhyay et al. showed that Phe-18 of helix-2 is the central residue that initiates unfolding dynamics while other amino acid side chains of helix-2 remain largely unperturbed.[23] In the present DMSO-induced chemical denaturation, Phe-18 of helix-2 also serves as the first deformation point from which the native state to molten globule like compact state transformation is originated. The effects of the exposure of this residue to solvent then propagate and eventually unfolding of helix-2 occurs. Once the helix-2 melts, it remains in the unfolded state during the rest of the simulation. The rapid second helix opening process is followed by the late stage unfolding of first and third helix.

The similarity of the first step of unfolding between thermal denaturation and DMSO induced unfolding is striking. In the present case, DMSO induced conformational changes can be easily attributed to the preferential solvation of exposed protein hydrophobic residues by the methyl groups of DMSO. However, our simulation studies show that the initial unfolding by this preferential solvation by the methyl groups of DMSO is accompanied, or closely followed, by the solvation of the protein backbone via hydrogen bonding through the oxygen of DMSO. Initially the favorable electrostatic interaction with hydrophilic residues by polar end (S═O) of DMSO, and dispersion interaction with already exposed hydrophobic residues by nonpolar end (CH_3) of DMSO play a dominant role to maintain the native like fluctuation. However, the key unfolding step is reached (by a fluctuation) when the solvation of Phe-18 (and Phe-11) by methyl groups reaches a critical amount to expose the hydrophobic core which then offer the necessary access to the backbone for hydrogen bonding of the >N−H hydrogen by >S═O oxygen.

Identification of this initial melting at the second helix allows us to understand the microscopic aspects of protein−DMSO interaction. We have extensively studied the side chain and backbone dynamics near the second helix and their interaction with DMSO. We find that DMSO induced unfolding mechanism of HP-36 follows two elementary molecular steps: (1) In the first step the methyl groups of DMSO solvate the hydrophobic side chain groups that are partly buried, in the process displacing the water molecules from the surface of the hydrophobic region (here the cluster of Phe-7, Phe-11 and Phe-18 groups). As a result of such preferential solvation, hydrophobic side chain atoms become exposed (see snap 2 of Figure 5). (2) In the next step the backbone >N−H hydrogen is accessed. The exposure of the side chain now allows the solvation of the backbone through the oxygen of DMSO via hydrogen bonding. This stabilizes the unfolded structure (snap 4 of Figure 5).

In Figure 14 it is clearly evident that the enhancement in number of backbone ("NH")-DMSO ("O") hydrogen bonds occurs just immediately after the jump observed in the trajectory of interaction energy between the methyl groups of DMSO and the side chain residues of protein.

To emphasize the critical step of DMSO-induced unfolding process consists of both the hydrophobic methyl-side chain interaction as well as the backbone−DMSO hydrogen bond formation.

As shown in Figure 7, the initial separation between phenylalanines at positions 18 and 11 (or 7) is followed by an expansion of the protein.

7. CORRELATION WITH SOLVATION DYNAMICS STUDIES

Solvation dynamics has been developed into a powerful tool to study dynamics of complex systems.[47−49] Fortunately, solvation dynamics of HP-36 (and several other proteins) have been studied in detail by simulations[50,22] and here we attempt to correlate those results with the present study. It is found that water molecules near the hydrophobic groups are nearly free and exhibit fast solvation dynamics, comparable to molecules in bulk water. Free energy calculations show that there is hardly any barrier for escape into the bulk for these (often termed "free") water molecules. Thus, it is relatively easy for the methyl groups of DMSO to displace such water molecules. Simulation studies also show, surprisingly, that water molecules hydrogen

dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

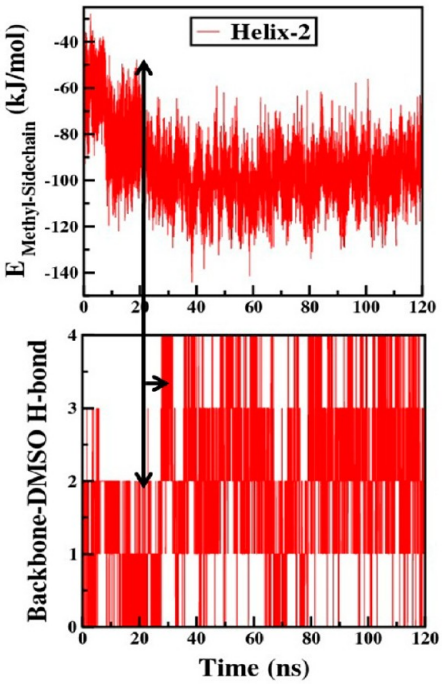


Figure 14. Role of the hydrophobic side chain−DMSO and backbone-DMSO interaction underlying the unfolding mechanism. Gradual decrease in the interaction energy between methyl groups of DMSO and the side chain residues of protein (up to the vertical black line) points toward the initial step of the mechanism that involves preferential solvation of the hydrophobic side chain atoms though the methyl groups of DMSO. Followed by the preferential solvation, after a certain interval of time, enhancement of the backbone−DMSO hydrogen bond formation is noticed. The exposure of side chains allows the oxygen of DMSO to access the backbone NH groups by hydrogen bond interaction.

bonded to the backbone polar atoms (oxygen, nitrogen, and hydrogen) also remain relatively free as these bonds are weak.[51] Therefore, the oxygen atom of >S═O group of DMSO can displace these water molecules, also.

Solvation dynamics studies and accompanied free energy calculations also show that there are a few water molecules near the hydrophilic surface that are strongly bound and face large free energy barriers for escape to bulk.[22] We find that these few water molecules remain bound to the surface and now play important role in stabilizing the unfolded state. From the simulation trajectory we found few water molecules that are doubly H-bonded to the protein residues named Glu5 and also to Asp6 in helix-1.

Zewail et al. and Bhattacharyya et al.[52,53] observed that solvation dynamics of protein groups can slow down slightly in the unfolded state. This rather paradoxical result can now be understood from the configuration of the unfolded state. The slower component can come from contribution of the side chains which are now free to interact with water and some of the side chains themselves are polar and/or charged. Simulations of Singer et al.[54] showed that solvation dynamics of Trp becomes faster if the motions of side chain atoms are quenched.

8. DISCUSSIONS AND CONCLUSIONS

Transition from an entropy dominated extended state to an enthalpy dominated folded (and/or ordered) native state and vice versa are complex processes where water and cosolvent molecules combine together to induce large conformational changes. Unlike in thermal denaturation where simultaneous weakening of water structure and entropic driving force at higher temperatures drive unfolding, here in the case of chemical denaturation, intermolecular interactions are tuned to cause the damage. These processes are yet to be understood in a detailed way. In an illuminating study, Barbara et al.[1] showed how interactions among different monomers in conjugated polymers can give rise to ordered defect structures. They also pointed out how arrangements of polymer segments can give rise to a photochemical funnel in these organic diodes, thus establishing a correlation between structure and functionality.

One of the main results of the present work is the retardation in the progress of unfolding with DMSO concentration at about 15%. All the properties of the protein (as shown in Figures 1, 3, and 4) undergo dramatic nonmonotonic composition dependence in the said range. As emphasized elsewhere,[15] this is a manifestation of a percolation-driven aggregation among methyl groups and that is a consequence of the amphiphilic character of DMSO. If this aggregation phenomenon did not intervene, we could have had much of unfolding over by 25%. Therefore, we need to have 30% molefraction of DMSO (which corresponds to 63% volume fraction) for the unfolding to be completed. Note that this amphiphilic nature of DMSO is essential, as already discussed, for hydrophobic solvation of side chains by the methyl groups and backbone solvation by oxygen of DMSO.

While folding of polymers and proteins has drawn enormous attention in the last 3−4 decades, the phenomenon of unfolding has seen less activity, with certain notable exceptions. It is not yet clear to what extent folding and unfolding pathways overlap. In this paper we have studied a small globular protein, HP-36, which has been studied extensively both in experiments and in simulations. The present work is different from earlier ones in the use of DMSO as a chemical denaturant. While urea and GdmCl have been used earlier, no theoretical or simulation studies have been carried out where unfolding has been investigated by increasing concentration of DMSO. Experimentally chemical denaturation has been studied and it has been found that unfolding is completed by 30−40% mole fraction of DMSO for most of proteins. Our simulations broadly reproduce the experimental observations.

Our reason for choosing this system is many-fold. First and foremost, we have discovered recently that water−DMSO binary mixture shows a surprising structural transformation around 15% of DMSO mole fraction where the methyl groups associate to form a percolating network. In the 10−15% of DMSO concentration range, polymer conformation also shows anomalous variation in the radius of gyration.[55] In another study it was observed that the active site of lysozyme shows a remarkable quenching of protein motion in the said composition range.[15]

Our study of unfolding of HP-36 reveals certain features perhaps not fully anticipated before. First, we find a nucleation-like mechanism that triggers unfolding. This is the separation/melting of Phe18 from Phe11 and Phe7. This step appears to be a common initial step toward unfolding, also observed in thermal denaturation. In the present case this initial unfolding step is facilitated by solvation of phenylalanines by the methyl groups of DMSO. As mentioned in the Introduction, the facilitation of unfolding by the methyl groups can be understood at least partly by considering the potential of

dx.doi.org/10.1021/jp308589b | J. Phys. Chem. B XXXX, XXX, XXX−XXX

mean force between two phenyl alanine groups.[24] The subsequent steps of unfolding are clearly coupled to this initial step as the second step is the separation between Ala9 and Leu35, and Leu2 and Leu35. Separation of these long distance tertiary contacts is accompanied by a significant expansion of the protein. The last stage is the melting of the helices which retain their structure till the last, and also remain virtually unaltered at lower DMSO concentrations.

Thus, unfolding of such a simple protein exhibits distinct, multistage stage (roughly four) dynamics. As our figures demonstrate, there is a correlation between these steps. Such a correlation seems to confirm the existence of a dominant pathway of unfolding. Quantitative understanding of chemical denaturation is a formidable problem because molecular theory of dynamics in a binary dipolar liquid itself is yet to be developed.[56] In the present case (water−DMSO mixture) one needs to additionally consider the amphiphilic nature of the DMSO molecule. In fact, in such a complex situation even computer simulations might lead to erroneous results because of the inadequacy of the employed force field. In the present case, we find, fortunately, that many aspects of chemical unfolding of HP-36 in water−DMSO mixture are consistent with results from experiments and also with results from earlier thermal denaturation and water−urea induced denaturation studies.

While denaturation by temperature, urea, and guanidium hydrochloride has been investigated extensively by simulation and theoretical studies, ours is perhaps the first theoretical/simulation analyses of unfolding of a globular protein in water−DMSO mixture. While some of the results appear to be in agreement with experiments,[14] much of the predictions of the present study remains to be verified.

AUTHOR INFORMATION

Corresponding Author

*E-mail: bbagchi@sscu.iisc.ernet.in.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This paper is dedicated to the memory of Professor Paul Barbara. To one of us (B.B.), he was a dear friend and valued collaborator whom B.B. knew from Ph.D. days at Brown University. Paul influenced in many ways ongoing research in our group at Bangalore. Work reported in this paper was supported in part by DST and BRNS, India. B.B. acknowledges support from Sir J. C. Bose Fellowship, DST.

REFERENCES

(1) (a) Hu, D.; Yu, J.; Wong, K.; Bagchi, B.; Rossky, P. J.; Barbara, P. F. *Nature* **2000**, *405*, 1030−1033. (b) Hu, D.; Yu, J.; Barbara, P. F. *J. Am. Chem. Soc.* **1999**, *121*, 6936−6937. (c) Yu, J.; Hu, D.; Barbara, P. F. *Science* **2000**, *289*, 1327−1330.

(2) (a) Levinthal, C. *J. Chim. Phys.* **1968**, *65*, 44−45. (b) Levinthal, C. How to fold graciously. In *Mossbauer Spectroscopy in Biological Systems*, Proceedings of a meeting held at Allerton house, Monticello, Illinois; De-Brunner, P., Tsibris, J., Munck, E., Eds.; University of Illinois Press: Urbana, IL, 1969; pp 22−24.

(3) Zwanzig, R.; Szabo, A.; Bagchi, B. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 20−22.

(4) (a) Chan, H. S.; Dill, K. A. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 447−490. (b) Dill, K.; Chan, H. S. *Nat. Struct. Biol.* **1997**, *4*, 10−19.

(5) Sali, A.; Shakhnovich, E.; Karplus, M. *Nature (London)* **1994**, *369*, 248−251.

(6) (a) Bryngelson, J. D.; Wolynes, P. G. *J. Phys. Chem.* **1989**, *93*, 6902−6915. (b) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167−195. (c) Onuchic, J. N.; Wolynes, P. G. *Curr. Opin. Struct. Biol.* **2004**, *14*, 70−75.

(7) Honeycutt, J. D.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3526−3529.

(8) (a) Wales, D. J. *Philos. Trans. R. Soc. A* **2005**, *363*, 357−377. (b) Wales, D. J.; Bogdan, T. V. *J. Phys. Chem. B* **2006**, *110*, 20765−20776. (c) Carr, J. M.; Wales, D. J. *J. Chem. Phys.* **2005**, *123*, 234901/1−12.

(9) Dill, K. A.; Shortle, D. *Annu. Rev. Biochem.* **1991**, *60*, 795−825.

(10) Cho, J. H.; Raleigh, D. P. *Methods Mol. Biol.* **2009**, *490*, 339−351.

(11) Koizumi, M.; Hirai, H.; Onai, T.; Inoue, K.; Hirai, M. *J. Appl. Crystallogr.* **2007**, *40*, 175−178.

(12) Li, W.; Zhou, R.; Mu, Y. *J. Phys. Chem. B* **2012**, *116*, 1446−1451.

(13) Yang, Z. W.; Tendian, S. W.; Carson, W. M.; Brouillette, W. J.; Delucas, L. J.; Brouillette, C. G. *Protein Sci.* **2004**, *13*, 830−841.

(14) Bhattacharjya, S.; Balaram, P. *Proteins: Struct., Funct., Genet.* **1997**, *29*, 492−507.

(15) Roy, S.; Jana, B.; Bagchi, B. *J. Chem. Phys.* **2012**, *136*, 115103/1−10.

(16) (a) Chung, J. K.; Thielges, M. C.; Fayer, M. D. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 3578−3583. (b) Chung, J. K.; Thielges, M. C.; Fayer, M. D. *J. Am. Chem. Soc.* **2012**, *134*, 12118−12124.

(17) (a) Wallqvist, A.; Covell, D. G.; Thirumalai, D. *J. Am. Chem. Soc.* **1998**, *120*, 427−428. (b) O'Brien, E. P.; Dima, R. I.; Brooks, B.; Thirumalai, D. *J. Am. Chem. Soc.* **2007**, *129*, 7346−7353.

(18) Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Saboungi, M. L.; Dempsey, C. E.; MacKerell, A. D., Jr.; Brady, J. W. *J. Am. Chem. Soc.* **2004**, *126*, 11462−11470.

(19) Lim, W. K.; Rosgen, J.; Englander, S. W. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 2595−2600.

(20) Shortle, D.; Ackerman, M. S. *Science* **2001**, *293*, 487−489.

(21) England, J. L.; Haran, G. *Annu. Rev. Phys. Chem.* **2011**, *62*, 257−277.

(22) Roy, S.; Bagchi, B. *J. Phys. Chem. B* **2012**, *116*, 2958−2968.

(23) Bandyopadhyay, S.; Chakraborty, S.; Bagchi, B. *J. Chem. Phys.* **2006**, *125*, 084912/1−11.

(24) (a) Mukherjee, A.; Bagchi, B. *Biochemistry* **2006**, *45*, 5129−5139. (b) Mukherjee, A.; Bhimalapuram, P.; Bagchi, B. *J. Chem. Phys.* **2005**, *123*, 014901/1−11.

(25) (a) McKnight, C. J.; Matsudaira, P. T.; Kim, P. S. *Nat. Struct. Biol.* **1997**, *4*, 180−184. (b) McKnight, C. J.; Doering, D. S.; Matsudaira, P. T.; Kim, P. S. *J. Mol. Biol.* **1996**, *260*, 126−134.

(26) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269−6271.

(27) (a) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656−1676. (b) Liu, H.; Mueller−Plathe, F.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **1995**, *117*, 4363−4366. (c) Geerke, D. P.; Oostenbrink, C.; van der Vegt, N. F. A.; van Gunsteren, W. F. *J. Phys. Chem. B* **2004**, *108*, 1436−1445.

(28) (a) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1697. (b) Nose, S. *J. Chem. Phys.* **1984**, *81*, 511−519.

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

(30) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed.; Academic Press: San Diego, CA, 2002.

(31) (a) Roy, S.; Banerjee, S.; Biyani, N.; Jana, B.; Bagchi, B. *J. Phys. Chem. B* **2011**, *115*, 685−692. (b) Banerjee, S.; Roy, S.; Bagchi, B. *J. Phys. Chem. B* **2010**, *114*, 12875−12882.

(32) Zagrovic, B.; Snow, C. D.; Khaliq, S.; Shirts, M. R.; Pande, V. S. *J. Mol. Biol.* **2002**, *323*, 153−164.

(33) (a) Srinivas, G.; Bagchi, B. *J. Chem. Phys.* **2002**, *116*, 8579−8587. (b) Srinivas, G.; Bagchi, B. *Theor. Chem. Acc.* **2003**, *109*, 8−21.

(34) Mukherjee, A.; Bagchi, B. *J. Chem. Phys.* **2004**, *120*, 1602−1612.

(35) Plaxco, K. W.; Simons, K. T.; Baker, D. *J. Mol. Biol.* **1998**, *277*, 985−994.
(36) Wales, D. J. *Energy Landscapes: with Applications to Clusters, Biomolecules and Glasses (Cambridge Molecular Science)*; Cambridge University Press: U.K., 2004
(37) Flory, P. J. *Principles of Polymer Chemistry*; Cornell University Press: New York, 1953; pp 519−538.
(38) Bryngelson, J. D.; Wolynes, P. G. *Biopolymers* **1990**, *30*, 177−188.
(39) Duan, Y.; Kollman, P. A. *Science* **1998**, *282*, 740−744.
(40) Lee, I. H.; Kim, S. Y.; Lee, J. *J. Comput. Chem.* **2010**, *31*, 57−65.
(41) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. *Science* **2011**, *334*, 517−520.
(42) (a) Marcus, R. A. *Annu. Rev. Phys. Chem.* **1964**, *15*, 155−196. (b) Marcus, R. A. *J. Chem. Phys.* **1965**, *43*, 679−701. (c) Marcus, R. A.; Sutin, N. *Biochim. Biophys. Acta* **1985**, *811*, 265−322.
(43) Sumi, H.; Marcus, R. A. *J. Chem. Phys.* **1986**, *84*, 4894−4914.
(44) Zwanzig, R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 2029−2030.
(45) Bagchi, B. *Molecular Relaxation in Liquids*; Oxford University Press: New York, 2012.
(46) Bagchi, B. *Int. Rev. Phys. Chem.* **1987**, *6*, 1−33.
(47) Barbara, P. F.; Jarzeba, W. *Acc. Chem. Res.* **1988**, *21*, 195−199.
(48) (a) Nandi, N.; Bhattacharyya, K.; Bagchi, B. *Chem. Rev.* **2000**, *100*, 2013−2046. (b) Nandi, N.; Bagchi, B. *J. Phys. Chem. A* **1998**, *102*, 8217−8221.
(49) Fleming, G. R.; Wolynes, P. G. *Phys. Today* **1990**, *43*, 36−43.
(50) Bandyopadhyay, S.; Chakraborty, S.; Balasubramanian, S.; Bagchi, B. *J. Am. Chem. Soc.* **2005**, *127*, 4071−4075.
(51) Jana, B.; Pal, S.; Bagchi, B. *J. Chem. Sci.* **2012**, *124*, 317−325.
(52) Othon, C. M.; Kwon, O. H.; Lin, M. M.; Zewail, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 12593−12598.
(53) Das, D. K.; Das, A. K.; Mandal, A. K.; Mondal, T.; Bhattacharyya, K. *Chem. Phys. Chem.* **2012**, *13*, 1949−1955.
(54) Li, T.; Hassanali, A. A.; Kao, Y. T.; Zhong, D.; Singer, S. J. *J. Am. Chem. Soc.* **2007**, *129*, 3376−3382.
(55) (a) Ghosh, R.; Banerjee, S.; Chakrabarty, S.; Bagchi, B. *J. Phys. Chem. B* **2011**, *115*, 7612−7620. (b) Ghosh, R.; Roy, S.; Bagchi, B. *J. Indian Chem. Soc.* **2011**, *88*, 1223−1231.
(56) (a) Chandra, A.; Bagchi, B. *J. Chem. Phys.* **1991**, *94*, 8367−8377. (b) Chandra, A.; Bagchi, B. *J. Chem. Phys.* **1989**, *91*, 1829−1842.