Solvent Sensitivity of Protein Unfolding: Dynamical Study of Chicken Villin Headpiece Subdomain in Water–Ethanol Binary Mixture

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · OCTOBER 2013

Impact Factor: 3.3 · DOI: 10.1021/jp406255z

3 AUTHORS:

Rikhia Ghosh
Indian Institute of Science
11 PUBLICATIONS **45** CITATIONS
SEE PROFILE

Susmita Roy
Rice University
23 PUBLICATIONS **164** CITATIONS
SEE PROFILE

Biman Bagchi
Indian Institute of Science
456 PUBLICATIONS **10,989** CITATIONS
SEE PROFILE

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

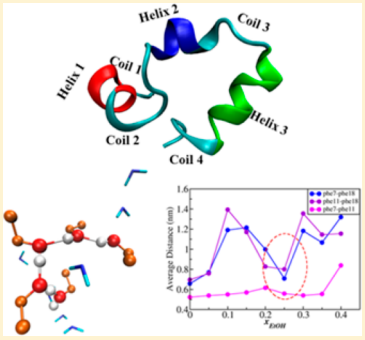
Solvent Sensitivity of Protein Unfolding: Dynamical Study of Chicken Villin Headpiece Subdomain in Water−Ethanol Binary Mixture

Rikhia Ghosh, Susmita Roy, and Biman Bagchi*

Solid State and Structural Chemistry Unit, Indian Institute of Science, C. V. Raman Avenue, Bangalore 560012, India

Supporting Information

ABSTRACT: We carry out a series of long atomistic molecular dynamics simulations to study the unfolding of a small protein, chicken villin headpiece (HP-36), in water−ethanol (EtOH) binary mixture. The prime objective of this work is to explore the sensitivity of protein unfolding dynamics toward increasing concentration of the cosolvent and unravel essential features of intermediates formed in search of a dynamical pathway toward unfolding. In water−ethanol binary mixtures, HP-36 is found to unfold partially, under ambient conditions, that otherwise requires temperature as high as ∼600 K to denature in pure aqueous solvent. However, an interesting course of pathway is observed to be followed in the process, guided by the formation of unique intermediates. The first step of unfolding is essentially the separation of the cluster formed by three hydrophobic (phenylalanine) residues, namely, Phe-7, Phe-11, and Phe-18, which constitute the hydrophobic core, thereby initiating melting of helix-2 of the protein. The initial steps are similar to temperature-induced unfolding as well as chemical unfolding using DMSO as cosolvent. Subsequent unfolding steps follow a unique path. As water−ethanol shows composition-dependent anomalies, so do the details of unfolding dynamics. With an increase in cosolvent concentration, different partially unfolded intermediates are found to be formed. This is reflected in a remarkable nonmonotonic composition dependence of several order parameters, including fraction of native contacts and protein−solvent interaction energy. The emergence of such partially unfolded states can be attributed to the preferential solvation of the hydrophobic residues by the ethyl groups of ethanol. We further quantify the local dynamics of unfolding by using a Marcus-type theory.



I. INTRODUCTION

Whereas a general picture of protein folding and unfolding is beginning to emerge using the concept of energy−entropy funnel with intrinsic ruggedness,[1,2] the details of kinetics and pathways are still far from our understanding at present. The reason is the absence of information of local changes involved in these transitions. Recent advances in two-dimensional infrared (2D-IR) vibrational echo spectroscopic techniques, pioneered by, among others, Fayer and coworkers, have developed a reliable method to measure dynamics of the proteins that are sensitive to the details of spatial arrangement of amino acid residues around a given peptide group.[3−8] 2D-IR spectroscopy allows unprecedented accuracy of time scale (in the picosecond range) of local motion that is required to study protein unfolding, in particular, whereas conventional techniques (like NMR) are reliable only on longer time scales[8] (on the order of nanoseconds). We report molecular dynamics (MD) simulation study of unfolding of a small protein with particular emphasis on melting/unfolding of secondary structure. The novelty of the present work is to understand the sensitivity of suitable aqueous binary mixture whose composition can be varied to induce unfolding under ambient conditions. This allows a direct comparison between experiments and theory aided by the simulations performed.

The funnel picture of protein folding was essentially introduced by Bryngelson and Wolynes in their landmark work in 1989.[1] The folding funnel landscape describes the protein-folding phenomenon as folding of protein to the native state through any of the possible large number of pathways and intermediates rather than choosing a single mechanism.[9] Although some proteins seem to show a unique folding path through well-defined intermediate states, many aspects of the folding picture have been accepted by the community at large. It is expected that different pathways could be accessed under different conditions to achieve the same native state.

Recently, there has been substantial insight regarding the knowledge of native states as well as the stable intermediate states of several proteins.[10−13] However, less information is available about the path followed by protein to fold to correct, native state and the characteristics of the intermediates along the way. These intermediates are mostly short-lived and thus very difficult to trap.[14] Because of the formation of such unstable intermediates, in the pathway of unfolding, most of the studies on protein folding focus on the native state conformation rather than looking at the unfolded states.[15,16]

The energy landscape of a particular protein is perturbed by various external factors that include temperature, pressure, as

Special Issue: Michael D. Fayer Festschrift

Received: June 25, 2013
Revised: October 29, 2013
Published: October 29, 2013

well as nature of solvents. Water is found to be a universal solvent for proteins, stabilizing the native state in most of the cases.[17,18] In contrast with the theoretical studies, experimental studies of protein folding and other kinetic studies involving enzymes seldom employ pure water, and a variety of cosolvents are being used for solvating such biomolecular systems before undergoing experiments. These cosolvents are found to play diverse roles in influencing the structural transitions of such systems. While some cosolvents are famous for stabilizing the native structures (glycerol, trimethyl amine N-oxide), some bring about denaturation in the same (urea, guanidium chloride).

Although an impressive number of studies have been devoted to explore the dynamics of water around a biomolecule,[19−22] relatively fewer studies have been performed on proteins in binary mixtures. Binary mixtures of solvents have always served as promising ones in terms of large change in physical and chemical behaviors that they exhibit compared with the individual components.[23−26] Water−ethanol (EtOH) is one of such very important binary solvents, widely used in biology as essential solvent because of its unique behavior. Water−ethanol binary mixture is famous for exhibiting striking anomalies at various concentrations that essentially arise due to structural transformation of ethanol through hydrophobic as well as hydrophilic interactions. Recent simulation studies have shown that at very low concentration ($x_{EtOH} \approx 0.06-0.10$) a bicontinuous phase forms by aggregation of ethanol molecules, leading to weak phase transition in the system.[27] In another significant experimental work, Juurinen et al. have performed synchrotron X-ray Compton scattering technique to show the structural transformation in solution of ethanol from a new perspective.[28] The experiments reveal two distinct structural regimes, one at low concentration ($x_{EtOH} \approx 0.05$) characterized by increase in intramolecular bond lengths and another at high concentration range ($x_{EtOH} > 0.15$), characterized by excess density. In fact, ethanol is found to be a very well known solvent for biomolecules. It has been, of course, the universal cosolvent in various liquors for ages. A recent experimental work successfully provides quantitative solvation properties of lysozyme in water−ethanol binary mixtures, where preferential binding of ethanol molecules is obtained for both the native state as well as unfolded states.[29] In another work, it has been shown that fluorinated ethanol (trifluoroethanol, TFE) brings in conformational transition in proteins to form new stable conformational states that resemble "molten globule" intermediate characterized by high α-helical content and disrupted tertiary structure.[30,31] Various studies of biomolecules in water−ethanol have revealed that because of structural transformation from the typical tetrahedral-like water to the chain-like alcohol clusters in alcohol−water mixtures, secondary structures of proteins change from β-strand to α-helical strand near the transition concentration. A recent study of the protein chymotrypsin inhibitor-2 (CI2) in aqueous ethanol shows an increase in radius of gyration up to mole fraction of ethanol $x_{eth} \approx 0.2$, and after that it decreases significantly.[32] The universal feature of alcohol-induced structural change of the protein from β-strand to α-helix is demonstrated in this work.

These interesting studies have driven us to explore how the structure of a small protein rearranges with the progressive addition of ethanol and what is the microscopic reason behind the same. In a very recent study we could identify the sequence of steps taking place during complete unfolding of a protein in water−DMSO binary mixture as the DMSO concentration is increased.[33] However, no systematic theoretical study seems to exist on unfolding in water−ethanol binary mixture. Because unfolding can be induced in mixed solvents by varying composition of the solute or cosolvent, it provides us with a great opportunity to study various aspects of unfolding, especially the sensitivity to varying environment and hence a glimpse of the energy landscape of the protein. For this purpose, we have taken a 36 residue protein, chicken villin headpiece (HP 36), and looked at its change of dynamics with varying ethanol concentration. We have further compared the results with those obtained for water−DMSO binary mixture. HP-36 is a small globular protein that represents the thermostable subdomain present at extreme C terminus of the 76-residue chicken villin headpiece domain.[34,35] Villin is a unique protein that can both assemble and disassemble actin structures. HP-36 contains one of the two F-actin binding sites in villin necessary for F-actin bundling activity.[35,36]It is a very interesting protein that although being very small, comprises three α-helices as well as one compact hydrophobic core, thereby bearing characteristics of large proteins. The three α-helices are denoted as Helix 1 (Asp-4 to Lys-8), Helix 2 (Arg-15 to Phe-18), and Helix 3 (Leu-23 to Glu-32). The biological activity is believed to be centered on helix 3.[36] Simulation studies with this protein are computationally less expensive owing to its small size. In previous works, a model of the protein was extensively studied by constructing a hydropathy scale for the constituting amino acids, thereby exploring the correlation between energy landscape and folding topology of the same.[37,38] Further atomistic simulation study of HP 36 in water revealed that the protein is extremely stable in water, and partially unfolded molten globule state was achieved on increasing the temperature as high as 600 K.[39]

The following article reveals several interesting results based on the molecular dynamics study of HP 36 in aqueous ethanol solutions of increasing ethanol concentration. We demonstrate unusual dynamical variation of structure of the protein along with change in ethanol concentration and correlate the same with the anomalous behavior of water−ethanol binary mixture at particular critical concentrations, which is brought about by the aggregation of ethanol molecules. Through detailed study of protein dynamics, we show that with progressive addition of ethanol the protein undergoes a structural transition, initially unfolding partially, followed by refolding of the same. Time evolution of several order parameters provides us with the same anomalous structural changeover. We further provide understanding at a molecular level to apprehend the reason for such striking features of protein at different cosolvent concentrations.

Initial steps of unfolding are found to be similar to those of unfolding in water−DMSO binary mixture as well as temperature-induced unfolding, that is, the separation of the hydrophobic core (instigated by the separation of Phe-18 from Phe-7 and Phe-11). Such separation is possible due to cumulative interaction of hydrophobic core with hydrophobic part of the cosolvent molecules (CH_3-CH_2-). However, as ethanol concentration increases, the distance between the phenylalanine residues decreases considerably, which is unlike the other cases of unfolding studies. At $x_{eth} \approx 0.25$, the protein achieves a partially folded state that is native-like. We find from the corresponding snapshots that the α-helices are reformed in this stage, which is similar to the results obtained for chymotrypsin inhibitor 2, which have been discussed beforehand. Thus the results obtained here confirm the propensity of

ethanol molecules to bring about stabilization of α-helices. On further increasing ethanol concentration, another set of abnormal structural variation of HP 36 is observed that is again related to the localized availability of free ethanol molecules driving the partial unfolding.

The rest of the article is arranged as follows. Details of the system and simulation are discussed in Section II. Section III elaborates the main results obtained in this work, which include anomalous variation of several order parameters with change of ethanol concentration, such as radius of gyration, root-mean-square deviation, average fraction of native contacts formed at equilibrium, and comparative maps of hydrophobic native contacts at different ethanol concentration. The later part of this section provides the microscopic details of the reason behind such anomalous structural transition of the protein, which is mainly attributed to the aggregation of ethanol molecules at higher concentrations. Section IV presents a simple theoretical analysis of solvent dependence of unfolding that is derived by combining the aspects of Bryngelson–Wolynes theory with Marcus theory of electron transfer. The brief summary of the results obtained is given in Section V.

II. SIMULATION DETAILS

We have performed molecular dynamics simulation of the protein HP-36 in water ethanol binary mixture at various ethanol concentrations. All simulations are done at 300 K temperature and 1 bar pressure. The simulations are initiated with the particular configuration of HP-36 that is obtained from NMR data of the villin headpiece Subdomain.[34,35] The coordinates are collected from Protein Data Bank (PDB code 1VII). The terminal residues of the protein (Met-1 and Phe-36) are amidated and acelyated, respectively, to avoid technical difficulties during simulation. We have used the extended simple point-charge model (SPC/E) for water.[40] We have treated the ethanol molecules as united atoms within the GROMOS53a6 force field;[41] that is, full atomistic details have been retained, except for the hydrogen atoms attached to carbon atoms. At first, water–ethanol binary solution was prepared for various concentrations in cubic boxes, with sides of length 3.0 nm, although minor variations were there in all cases to maintain proper mole fraction ratio (concentrations of ethanol taken are 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, and 0.40, respectively). After performing steepest descent energy minimization, equilibration of the system was done for 2 ns, keeping temperature and volume constant. After that, we again performed an equilibration for 2 ns, keeping pressure and temperature constant. Finally, production run was executed for 20 ns in a NPT ensemble. Temperature was kept constant using a Nosé–Hoover thermostat,[42,43] and a Parinello–Rahman barostat[44] was used for pressure coupling. After the binary mixtures were prepared, the protein was dissolved in each of them and again, the same procedure was followed for energy minimization. The box size taken in this case is larger (~7 nm) to accommodate the protein as well as permitting better solvation by housing more solvent molecules in the box. The solvent was further equilibrated for 20 ns by restraining the position of protein while allowing the solvent to move. This method is popularly known as position-restrained molecular dynamics simulation. Finally, the production run was performed for 100 ns in NPT ensemble. Periodic boundary conditions were applied, and nonbonded force calculations were employed by applying a grid system for neighbor searching.[45] The cutoff radius taken for the neighbor list and van der Waals interaction was 1.4 nm. To calculate electrostatic interactions, we used particle mesh Ewald (PME) with a grid spacing of 0.16 nm and an interpolation order of 4.

III. STUDY OF PROTEIN UNFOLDING IN WATER–ETHANOL BINARY MIXTURE

A. Solvent Composition Dependent Structural Changes of Protein: Variation of C-αRMSD and Radius of Gyration. The standard way of analyzing structural stability of a protein in a MD simulation is to monitor the root-mean-square deviation (RMSD) from the initial structure along the simulation. The dynamical change of RMSD of C-α backbone of the protein along with change in ethanol concentration is shown in Figure 1a. Figure 1b shows the corresponding average

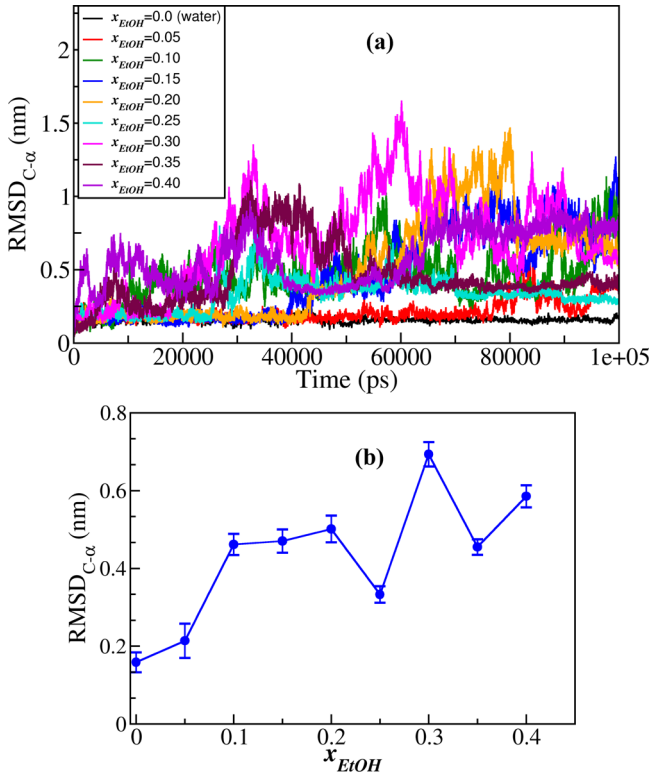


Figure 1. (a) Time evolution of root-mean-square deviation (RMSD) of C-αbackbone and (b) average RMSD over the whole time trajectory bearing a clearer picture of solvent composition dependent structural changes. The error bars characterize variance from the mean value, calculated over the entire trajectory. Lines are drawn to guide the eye of the viewer.

values at each individual mole fraction. We find that on adding ethanol up to mole fraction, $x_{eth} \approx 0.10$, gradual unfolding of the protein structure takes place. However, on further addition of ethanol, the process of unfolding is somewhat hindered, as is seen from Figure 1b ($x_{eth} \approx 0.10$~0.20). After that, there is sudden drop in the value of average RMSD at, $x_{eth} \approx 0.25$, signifying the transition of the structure toward the folded state, which again jumps to a very high value at $x_{eth} \approx 0.30$, thus indicating the formation of the most unfolded state in this case. At higher concentrations, such oscillations between the partially folded and unfolded states prevail ($x_{eth} \approx 0.35$~0.40).

Although RMSD serves as a good measure of the degree of conservation of a structure, it is still limited, and analysis of radius of gyration is helpful to bring further insight toward the

dx.doi.org/10.1021/jp406255z | J. Phys. Chem. B 2013, 117, 15625–15638

same. We have calculated the time evolution of the radius of gyration of HP 36 by changing the cosolvent concentration from $x_{eth} \approx 0.0$ to 0.40, which is shown in Figure 2a, along with

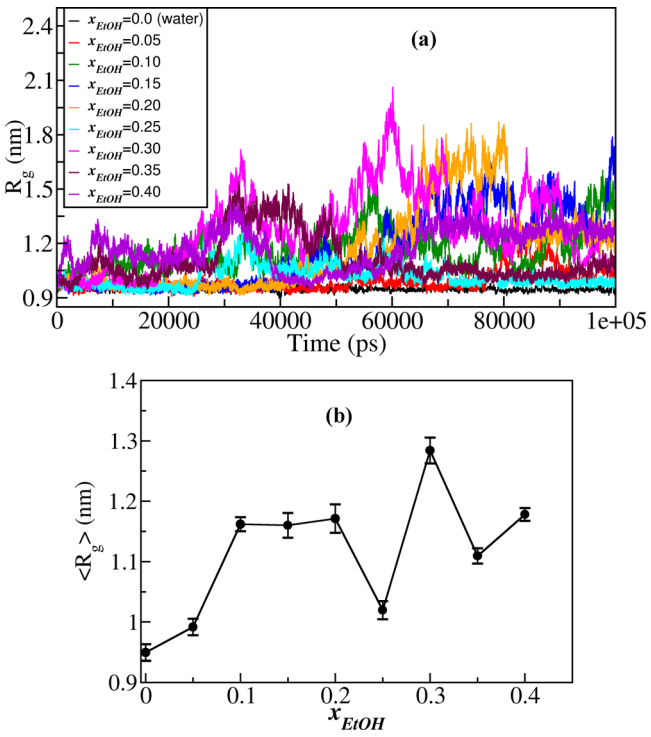


Figure 2. (a) Time trajectory of radius of gyration of the protein at different mole fraction of ethanol. (b) Equilibrium average of radius of gyration over the whole trajectory plotted against increasing ethanol concentration. The error bars represent variance from the mean value, calculated over the entire trajectory. Lines are shown to guide the eye of the viewer.

equilibrium average values shown in Figure 2b. In this case, we also find the same signature as that obtained from RMSD values, that is, gradual unfolding up to x_{eth}~0.10, followed by less significant change of structure in the concentration $x_{eth} \approx$ 0.10~0.20 range. On further increasing ethanol concentration, the protein structure fluctuates between partially folded and unfolded states ($x_{eth} \approx$ 0.20~0.40).

B. Study of Fraction of Native Contacts Formed and Broken. The formation or rupture of specific contacts between respective residues in a protein serves as an important parameter in the study of protein dynamics. Proteins are often known to fold to the correct native structure, making some precise contacts among the hydrophobic and hydrophilic residues, respectively. While the process of unfolding is being executed, these important contacts break up in the first place, making the course of unfolding smoother for the same. To understand the structural changes induced by increasing ethanol concentration, we have evaluated the average value of fraction of contacts ruptured or restored in the protein ($\langle\eta\rangle$) relative to that of the native, correctly folded structure (Figure 3). The fraction of native contacts survived is calculated in the following way. The native structure is taken in the first place, and the distance between each pair of the backbone residue is measured. If that distance falls within a cutoff, then that contact is defined to be a "native" contact. Now the time evolution of the protein at particular ethanol concentration is studied in the same way to see how many of the "native" contacts are restored

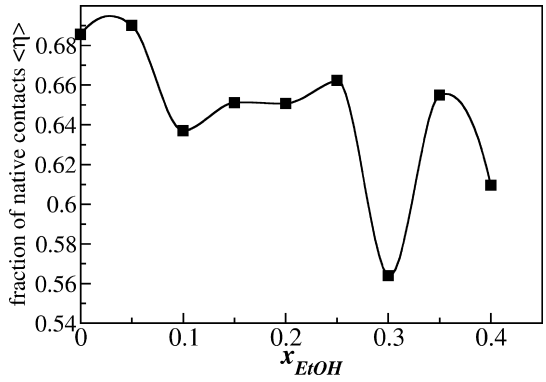


Figure 3. Equilibrium average of fraction of native contacts as a function of increasing ethanol concentration. Lines are shown to guide the eye of the viewer.

at each step by calculating the distance among the backbone residues in the same manner. The cut off for the formation of native contacts between the residues has been taken as 4 Å, as per the standards.[46]

In this case, we find that on adding a very small quantity of ethanol ($x_{eth} \approx 0.05$) the relative number of native contacts formed in the protein undergoes minor change compared with the number of contacts formed in bulk water. However, on increasing the ethanol concentration to $x_{eth} \approx 0.10$, the average number of fractional contacts decreases significantly, indicating the emergence of partial unfolding of the structure. The important contacts are again restored on further increase in ethanol concentration ($x_{eth} \approx$ 0.10~0.25), which signifies the phenomenon of refolding. At $x_{eth} \approx 0.30$, there is huge decrease in the value of $\langle\eta\rangle$, followed by oscillation between a higher and lower value of the same at higher ethanol concentration ($x_{eth} \approx$ 0.35~0.40) that is eventually a measure of relative folding and unfolding of the structure.

C. Snapshots of the HP-36 at Different Ethanol Concentration. To see the structural changes induced by varying concentration of ethanol, we follow the snapshots of the protein during different stages of simulation. In Figure 4, we provide the structures obtained after equilibrium is reached in the system. The structures are essentially taken after ~70 ns run is over, subsequent to which not much of change in average configuration is observed. We find that at $x_{eth} \approx 0.05$, partial unfolding is initiated by disruption of Helix 2 region as well as deformation of hydrophobic core that comprises Phe-7, Phe-11, and Phe-18. The same initial pathway toward unfolding was found to be followed by the protein in the case of previous simulation studies.[33] In the ethanol concentration range of $x_{eth} \approx 0.10$ to 0.20, not much structural variation is observed. In all of the cases, we find melting of the second helix, accompanied by melting of coil-2 and coil-3, along with slight deformation of first helix at $x_{eth} \approx 0.20$. This is also found from the radius of gyration and RMSD plots (Figures 1 and 2). Inconsistency arises at $x_{eth} \approx 0.25$, where the protein is again found to acquire a native like structure, with reformation of Helix-2. This remarkable result is highly unexpected, as with increasing ethanol concentration protein is expected to unfold gradually as a result of increasing protein−solvent hydrophobic interaction. Further increase in ethanol concentration leads to the formation of different partially unfolded intermediates. At $x_{eth} \approx 0.30$, Helix-2 and Helix-3 melt, whereas Helix-1 remains intact. At $x_{eth} \approx 0.35$, all three helices undergo partial melting, and at $x_{eth} \approx 0.40$, Helix-1 melts along with partial melting of

dx.doi.org/10.1021/jp406255z | J. Phys. Chem. B 2013, 117, 15625−15638

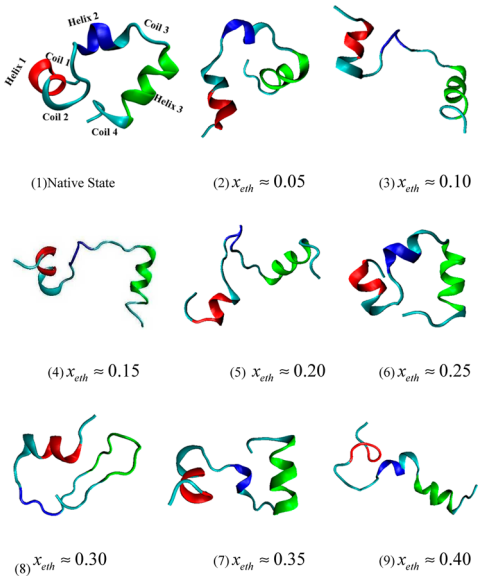


Figure 4. Equilibrium snapshots of HP-36 at different mole fractions of ethanol. Three different helices are assigned three different colors, namely, Helix-1 (4−8 amino acid residues) is assigned red, Helix-2 (15−18 amino acid residues) is assigned blue, and Helix-3 (23−32 amino acid residues) is assigned green. Coil-1, Coil-2, Coil-3, and Coil-4 are marked by cyan color. Snapshot (1) demonstrates native structure of the protein in water, (2) at $x_{eth} \approx 0.05$, (3) at $x_{eth} \approx 0.10$, (4) at $x_{eth} \approx 0.15$, (5) at $x_{eth} \approx 0.20$, (6) at $x_{eth} \approx 0.25$, (7) at $x_{eth} \approx 0.30$, (8) at $x_{eth} \approx 0.35$, and (9) at $x_{eth} \approx 0.40$.

Helix-2, keeping Helix-3 intact. These unpredictable structural fluctuations provide highly interesting results, and perhaps this work reports the first ever systematic study to show such fluctuating structural changes, along with increasing cosolvent concentration to the best of our knowledge. To ensure that the reformation of native-like state at $x_{eth} \approx 0.25$ is a general phenomenon, we have performed several simulations at this particular concentration, taking a high-resolution crystal structure (PDB code 1YRF), and encountered similar observations. Details of these results are provided in the Supporting Information. These initial results demand further detailed analysis of the system that is discussed in the next sections.

D. Study of Structural Change of HP-36 through Contact Map. From the previous discussions on dynamical variation of HP 36, it is realized that the pathway of unfolding can be envisaged by a number of partially unfolded intermediates. This can be interpreted as the presence of a number of local free-energy minima along the pathway of unfolding. To better understand the nature of these intermediates, we map the hydrophobic contacts present in the protein at each ethanol concentration after equilibrium is reached and compare them with the corresponding native hydrophobic contacts (Figure 5)

A close inspection of Figure 5 reveals that also at very low concentration of ethanol ($x_{eth} \approx 0.05$) the tertiary structure starts breaking with the removal of long-distant contacts, marked by the removal of points placed at the farthest corner (1−36, 2−36, 1−6). The initial steps of unfolding involve separation of the two significant hydrophobic contacts (7−18 and 11−18). An important point to be noted is that in the case of study of HP-36 in water−DMSO also the initial steps of unfolding were found to be the separation of hydrophobic core through the separation of Phe-7−Phe-18 and Phe-11−Phe-18 contacts.[33] In fact, in the case of $x_{eth} \approx 0.05$, some new tertiary contacts are found to be formed that are not present in the native structure, which implies the formation of some stable partially folded intermediate with different type of contacts present in the system (e.g., 9−36, 17−29).

At $x_{eth} \approx 0.10$, all tertiary contacts are lost along with the loss of some secondary contacts (e.g., 2−19, 18−29). This signifies further unfolding of the protein. As ethanol concentration is increased from $x_{eth} \approx 0.15$ to 0.20 it is found that Phe-11−Phe-18 contact reappears, with the recurrence of some secondary contacts. This is in accordance with the distribution of distances between these two residues shown in the next section. From Figure 1b, we find that in this region the average R_g value of the protein remains almost constant. However, at $x_{eth} \approx 0.25$, it is seen that almost all secondary and tertiary contacts are restored, thereby establishing the fact that at this particular concentration the protein is again folded to a give a native-like structure. Here we find that all of the hydrophobic core contacts are reformed (7−18 and 11−18); important tertiary contacts also reappear (7−36, 11−35). At $x_{eth} \approx 0.30$, protein acquires a significantly extended state, designated by the least number of contacts present in the corresponding map. At $x_{eth} \approx 0.35$, some secondary and tertiary contacts redevelop, which is also evident from Figure 4. (The second helix is reformed to some extent.) At $x_{eth} \approx 0.40$, no tertiary contacts are present, and the formation of a partially unfolded structure takes place.

E. Breakdown of Tertiary Structure: Initial Separation Followed by Aggregation of Hydrophobic Core. HP-36, although being a small protein (36 amino acid residues), consists of a central hydrophobic core that comprises three phenylalanine groups (Phe-7, Phe-11, and Phe-18). As seen from the contact map of the native structure, these three hydrophobic residues form important tertiary contacts that contribute to correct folding of the protein. Because protein unfolding is universally observed to be associated with structural changes in hydrophobic cores, we follow the distance distribution of the three hydrophobic pairs, namely, Phe-7−Phe-18, Phe-11−Phe-18, and Phe-7−Phe-11, along the course of changing ethanol concentration (Figure 6). We find that all pairs are in close contact in water. However, as ethanol concentration increases, these tertiary contacts start disappearing ($x_{eth} \approx 0.10$) for the two hydrophobic pairs Phe-7−Phe-18 and Phe-11−Phe-18. In distance distribution plots, we find that the first peak almost vanishes and a second peak at a larger distance starts emerging. This phenomenon essentially signifies the breakage of tertiary structure that is also reflected in the snapshot of the protein at $x_{eth} \approx 0.10$, where the disruption of the second helix is found to be initiated. On further increasing ethanol concentration ($x_{eth} \approx 0.15$ to 0.20), these hydrophobic tertiary contacts start reappearing, which is manifested in the gradual disappearance of the second peak at larger distance and recurrence of first peak at smaller distance. Thus we state that initial melting of the protein is associated with initial separation, followed by nucleation of the hydrophobic core. In fact, we have found the same phenomenon happening in the case of water−DMSO also in this particular concentration range.[33] The formation of Phe-7−Phe-18 and Phe-11−Phe-18 contacts continues up to $x_{eth} \approx 0.25$, where the distribution again produces a single peak at a small separation value. Accordingly, we find from the snapshots that the protein refolds at this concentration, giving average R_g and RMSD values very close to those in water. This produces an astonishing result, as mixed

dx.doi.org/10.1021/jp406255z | J. Phys. Chem. B 2013, 117, 15625−15638

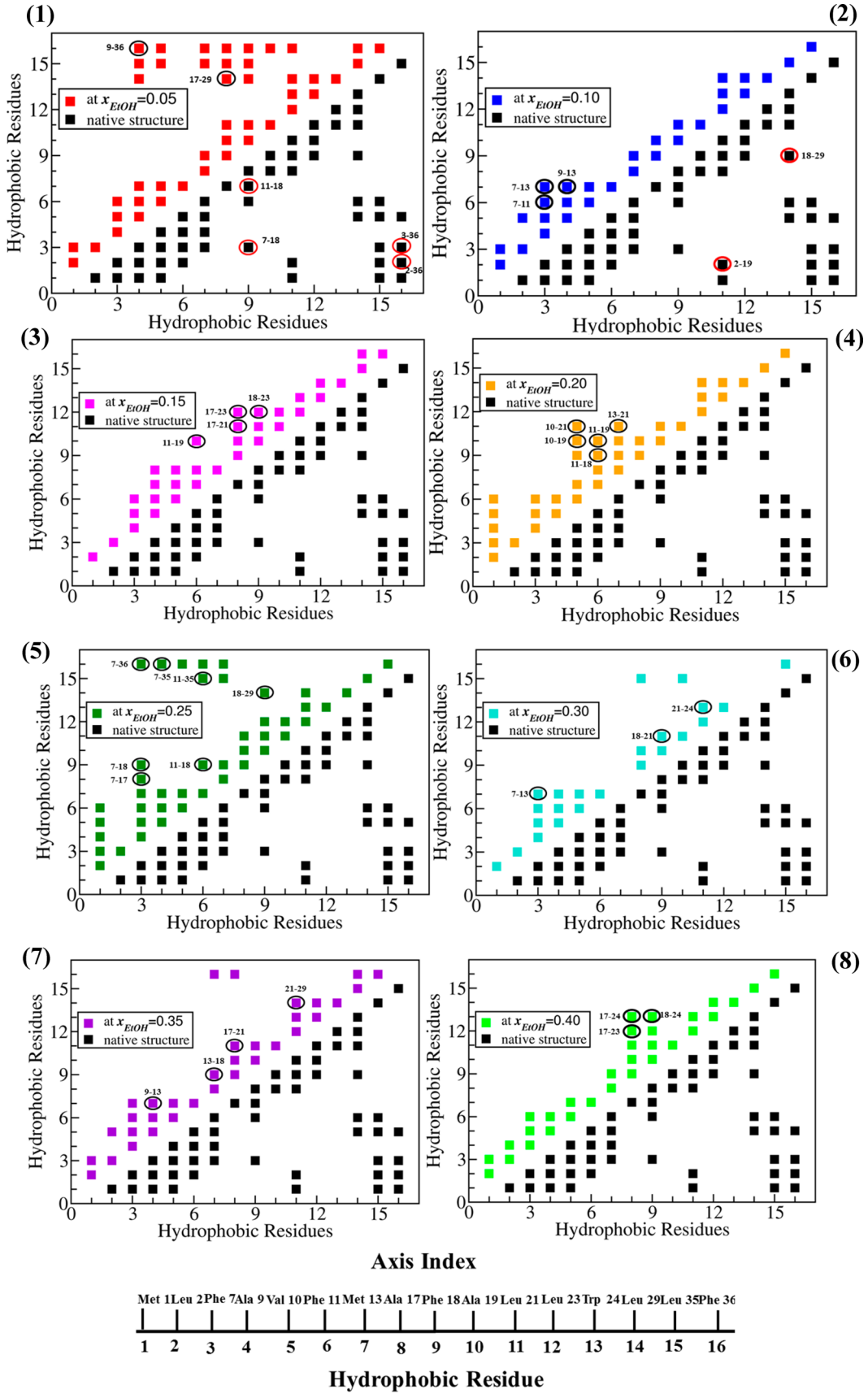


Figure 5. Hydrophobic contact mapping of HP-36 at different mole fractions of ethanol showing evolution of new contacts and breakage of old contacts, along with comparison with the native hydrophobic contacts present at equilibrium (right diagonal). The axis index given above shows the hydrophobic residue number. It is to be noted that deformation of second helix starts with separation of hydrophobic core (Phe-7, Phe-11, and Phe-18). As the protein refolds again ($x_{eth} \approx 0.25$), these contacts are restored once more.

dx.doi.org/10.1021/jp406255z | J. Phys. Chem. B 2013, 117, 15625−15638

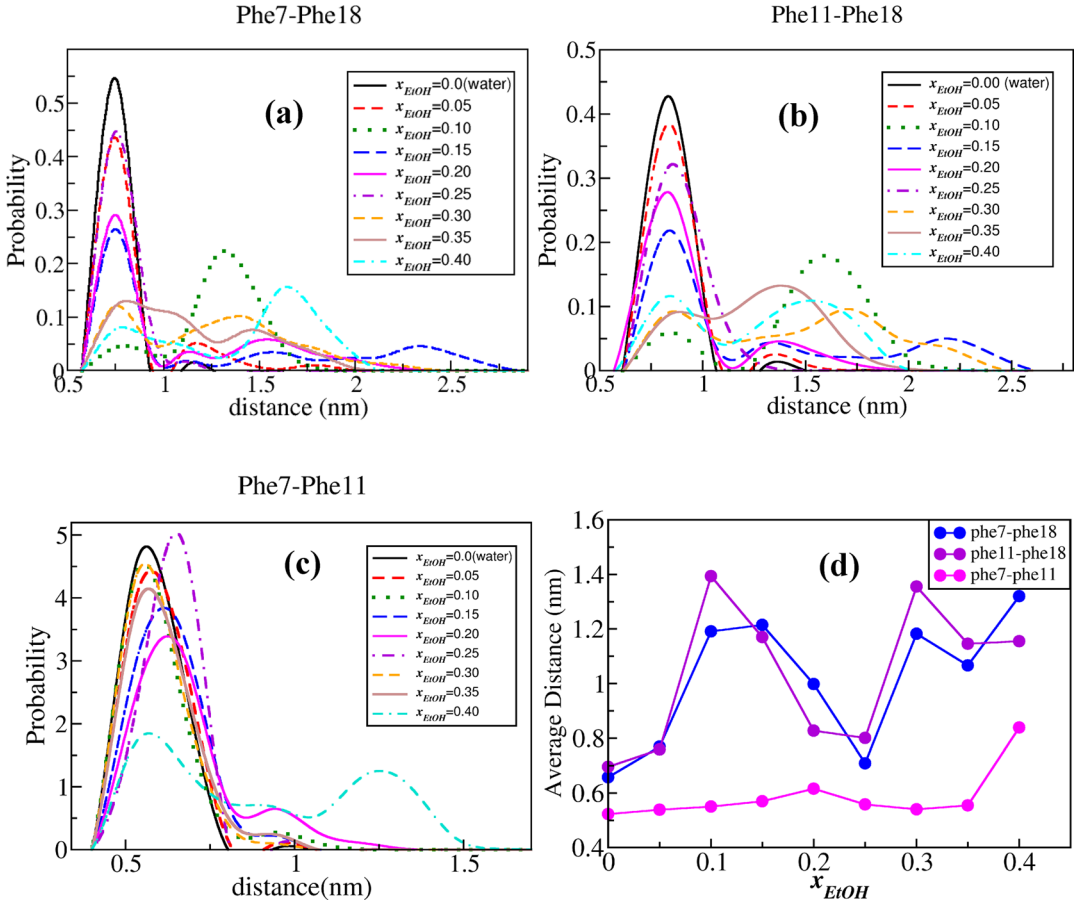


Figure 6. Comparison of probability distribution of distances between the residues of hydrophobic core (comprising of phenylalanine 7, phenylalanine 11, phenylalanine 18) with changing ethanol concentration. Distance distribution of (a) Phe-7−Phe-18 pair, (b) Phe-11−Phe-18 pair, and (c) Phe-7−Phe-11 pair. (d) Average value of distances between these three consecutive pairs along with change of ethanol concentration.

solvents are known to substantiate unfolding process at higher cosolvent concentration. On further increasing ethanol concentration (x_{eth} ≈ 0.30 to 40), we again observe the appearance of bimodal distribution with oscillatory values of the respective maxima, indicating corresponding appearance of partially unfolded states. The distance distribution of Phe-7−Phe-11 does not produce any such anomalies at lower concentrations, with the first peak height decreasing gradually along with the formation of a second peak continuing up to x_{eth} ≈ 0.20. At x_{eth} ≈ 0.25, the height of the first peak increases, followed by a small decrease in the same at x_{eth} ≈ 0.30. After that, the peak height again decreases, giving a bimodal distribution at x_{eth} ≈ 0.40. To further verify the connection between unfolding pathway and that of the hydrophobic core, we plot the average distance of the three hydrophobic pairs, respectively. In this case, we also find the same trend, thereby further strengthening the fact.

To compare the nature of structural changes of protein in water−ethanol with those in water−DMSO by varying cosolvent concentration, we plot the distribution of distances of the groups forming the hydrophobic core in Figure S2 (provided in the Supporting Information) at various DMSO concentrations. We find that on increasing DMSO concentration from x_{DMSO} = 0.05 to 0.10 the distance between both the pairs Phe-7−Phe-18 and Phe-11−Phe-18 increases, with the shift of distribution to a larger distance. On further increasing DMSO concentration (x_{DMSO} = 0.15), the distance between the hydrophobic pairs decreases to a large extent, indicated by the movement of the distribution again toward a smaller distance. These results are in excellent agreement with that obtained for water−ethanol. We come to the conclusion that the initial steps of unfolding induced by cosolvents pursue the same path, which is the primary separation of the hydrophobic core, followed by recoalition of the same. However, at further higher concentration the scenario takes a different approach for the two cases. In water−DMSO, with progressive addition of DMSO, the distance between the hydrophobic pairs starts increasing, leading to complete unfolding of the protein. We also plot the comparative average distance between these three residues to perceive a prominent picture of the corresponding unfolding pathway.

F. Interaction Energies: Understanding the Role of Solvent in Protein Unfolding. To understand the sensitivity of protein unfolding transition toward the change of solvent concentration, it is important to see how the interaction energy between the two varies along the path. We follow average of protein−protein interaction energy (Figure 7a) as well as average protein−solvent interaction energy (Figure 7b) over the whole time trajectory. It can be observed that on adding a small amount of ethanol (x_{eth} ≈ 0.05) in water both protein−protein and protein−solvent interaction energies show an insignificant change. However, on increasing the ethanol concentration to x_{eth} ≈ 0.10, protein−protein interaction energy decreases, thereby destabilizing the protein. Protein−solvent interaction energy increases, indicating the initialization of unfolding. In the range of ethanol concentration, x_{eth} ≈ 0.15

dx.doi.org/10.1021/jp406255z | J. Phys. Chem. B 2013, 117, 15625−15638

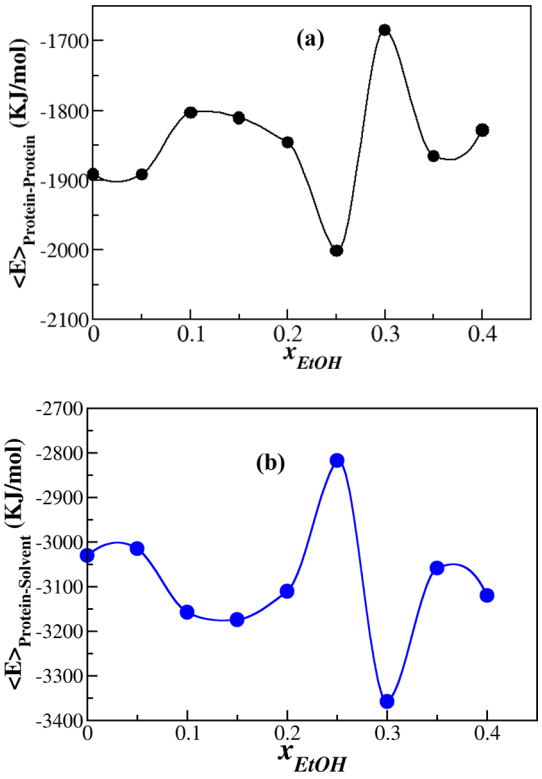


Figure 7. Change of average (a) protein−protein and (b) protein−solvent interaction energy with change in ethanol concentration. Nonmonotonous behavior of both protein−protein and protein−solvent interaction energy indicate dynamic structural fluctuation of the protein between partially folded and unfolded states. Lines are shown to guide the eye of the viewer.

to 0.20, protein−protein interaction energy increases to a small extent and protein−solvent interaction energy decreases accordingly. This means the unfolding that started by adding a small amount of cosolvent now ceases to occur. At $x_{eth} \approx 0.25$, protein−protein interaction energy increases enormously, leading to huge decrease in protein−solvent interaction energy, signifying the formation of a folded structure that is also observed in other results. On further increase in ethanol concentration, at $x_{eth} \approx 0.30$, the structural transition remarkably changes by stabilizing a highly unfolded state, signaling the very high average value of protein−solvent interaction energy and corresponding smaller value of protein−protein interaction energy. At $x_{eth} \approx 0.35$ to 0.40, both the average energies again achieve a moderate value that is close to the energy values obtained at $x_{eth} \approx 0.15$ to 0.20, thereby marking the appearance of a partially unfolded state all over again.

All of the results obtained in the preceding sections indicate a very interesting phenomena happening with increase in ethanol concentration. Obviously such unique observation demands explanation at a molecular level. In the next section, we try to develop the plausible rationalization of the same through deeper understanding of the problem.

G.1. Study of Water−Ethanol Binary Mixture at Different Ethanol Concentration: Ethanol Molecules Associate through Hydrogen Bonding at Moderate Concentration of Ethanol in Solution. To perceive a clearer picture of the phenomena at a molecular level, we looked at the equilibrated snapshots of the box. A very interesting scenario came out from this study. We have found that at lower concentration of ethanol ($x_{eth} \approx 0.05$ to 0.20), $2EtOH{\cdot}1H_2O$ clusters are prevalent in the system, formed by hydrogen bonding between hydrogen of water and oxygen of ethanol (Figure 8a). However, at $x_{eth} \approx 0.15$ to 0.20, ethanol molecules are seen to gradually come closer to each other and aggregate around the protein. This phenomenon is essentially a local one, bringing in microheterogeneous phase separation in the system. Through a closer inspection of the system in this particular concentration range, we find the emergence of a greater number of aggregated ethanol molecules, comprising three to five units of ethanol (Figure 8b,c). At $x_{eth} \approx 0.25$, it is found that the ethanol molecules are clustered together, forming three to five coordinated species, which are connected to each other through an ethanol−ethanol hydrogen bond. We also find that $2EtOH{\cdot}1H_2O$ species get numbered, with very few present in the bulk and near the surface of the aggregated ethanol molecules. The structural transformation taking place in water ethanol solution at moderately high concentrations is distinctly different from that occurring at low concentration of ethanol[28] and has already been reported in several works, as discussed in the Introduction. At $x_{eth} \approx 0.35$ to 0.40, similar molecular arrangements of ethanol are found to persist. We have also looked into the snapshots of the neat water−ethanol binary mixture at different concentration to make sure whether such self-association of ethanol is induced by protein. There also we find the same aggregation phenomena taking place with increase in ethanol concentration.

To quantify the presence of such aggregated molecular species in solution, we have calculated the average number of water−ethanol and ethanol−ethanol hydrogen bonds in neat water−ethanol binary mixture (Figure 9a). We indeed find that with increase in ethanol concentration the average number of water−ethanol hydrogen bonds decreases significantly, accompanied by a marked increase in number of ethanol−ethanol hydrogen bond. Thus our observations are consolidated. We have also plotted the diffusion coefficient of ethanol as a function of ethanol concentration to see whether such aggregation of ethanol molecules has any effect on the system as a whole (Figure 9b). We find that at concentration range $x_{eth} \approx 0.25$ to 0.30 diffusion coefficient decreases by a significant amount, followed by a further increase at $x_{eth} \approx 0.35$ to 0.40. Thus there exists an extensive anomaly in the concentration range of $x_{eth} \approx 0.25$ to 0.30 that can be attributed to the formation of three-coordinated and five-coordinated ethanol clusters, thus making the diffusion of the ethanol molecules through the system a very slow process. The reason for anomalous values of diffusion coefficient at low concentration is already explained in ref 27.

G.2. Aggregation of Ethanol Molecules Serves as the Main Reason behind the Anomalous Structural Variation of Protein. Clustering of ethanol molecules with increasing concentration alone explains most of the riddles that we faced in the abnormal structural variation of the protein. When free ethanol molecules are present in the solution, they serve as a good medium of interaction with the protein through strongly hydrophobic ethyl groups, resulting in separation of the hydrophobic core as well as initialization of melting of second helix. Now with increasing ethanol concentration free ethanol molecules become less available due to self-association. Thus the structural change in protein does not vary much in the concentration range of $x_{eth} \approx 0.15$ to 0.20. However, the refolding of the protein at $x_{eth} \approx 0.25$ is

dx.doi.org/10.1021/jp406255z | J. Phys. Chem. B 2013, 117, 15625−15638

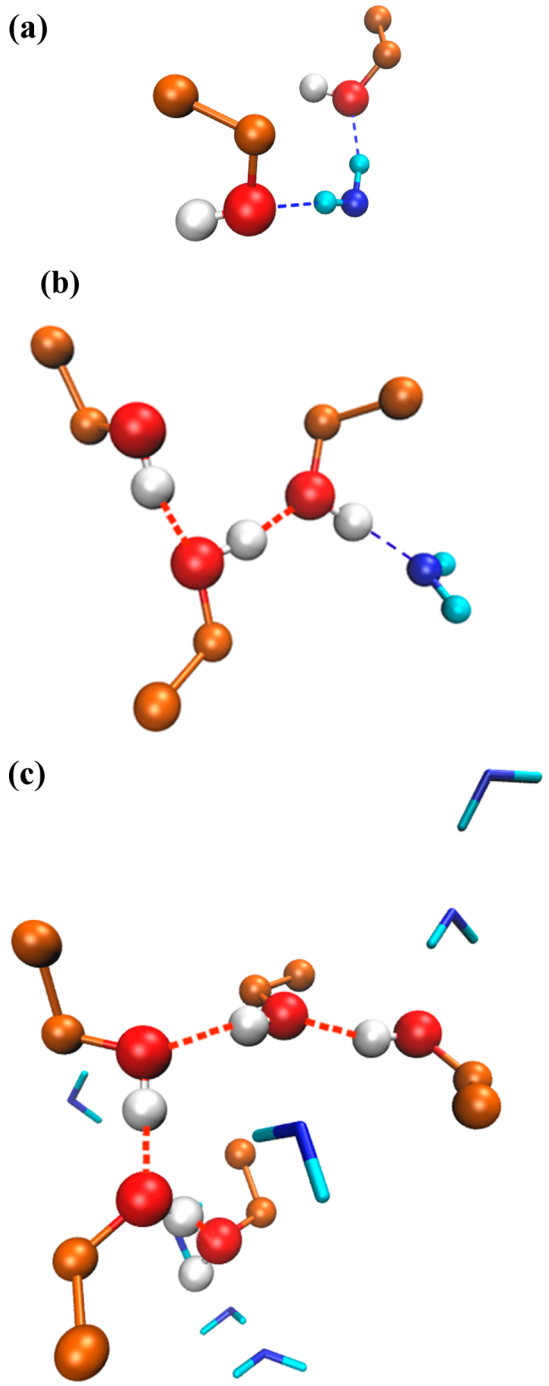


Figure 8. Structural formations in water ethanol binary mixture observed at different concentrations of ethanol. The bronze balls demonstrate ethyl groups, red balls show oxygen, and white balls are hydrogen atoms of ethanol molecules. The dark-blue balls and lines are oxygen, and light-blue ones are hydrogen of water, respectively. (a) $2EtOH{\cdot}1H_2O$ unit formed by hydrogen bond between hydrogen of water and oxygen of ethanol. These species are abundant in the solution at low concentration of ethanol. (b) Association of three ethanol molecules through hydrogen bonding between oxygen of ethanol molecule and hydrogen of another ethanol molecule. These species start forming once the ethanol concentration is gradually increased. (c) Five-coordinated ethanol molecules that are ubiquitous at and onward $x_{eth} \approx 0.25$. In this concentration range, species (a) is scarce and solution is rich of (b) and (c). The water molecules that are hydrogen-bonded to ethanol are represented by ball and stick, while the free ones are indicated by dynamic bonds.

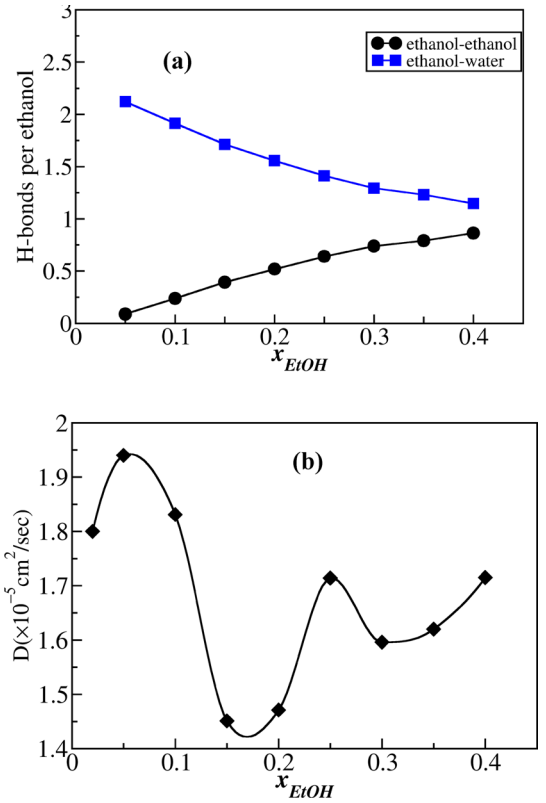


Figure 9. (a) Average number of water−ethanol and ethanol−ethanol hydrogen bonds per ethanol molecule as a function of increasing ethanol concentration. Association of ethanol molecules through hydrogen bond takes place at the expense of water−ethanol hydrogen-bonded species with increasing ethanol concentration. (b) Anomalous variation of diffusion coefficient of ethanol with varying concentration. The decrease in diffusion coefficient value at $x_{eth} \approx 0.30$ strongly supports the phenomena of aggregation of the cosolvent molecules. Lines are shown to guide the eye of the viewer.

attributed to the fact that formation of the ethanol−ethanol cluster comes to saturation at this level, thereby making a much smaller number of free ethanol molecules available. Thus the environment around the protein becomes water-like as a whole, resulting in the stabilization of a partially folded state. This is further supported by Figure 10, where we have plotted the order parameter η that is the percentage of water in the first hydration shell around the protein (within ~0.55 nm cutoff) as a function of ethanol concentration. We find that with increasing ethanol concentration η decreases significantly, demonstrating that ethanol molecules come in the close vicinity of the protein, promoting higher hydrophobic interaction. However, at $x_{eth} \approx 0.25$, water content increases markedly in the close proximity of the protein, signifying the creation of a water-like environment around the protein. This in turn promotes better solvation of the protein, thereby stabilizing a native-like structure.

On adding a little more ethanol to the solution ($x_{eth} \approx 0.30$), suddenly some free ethanol molecules are accommodated in the system, which again promotes hydrophobic solvation of the protein, thereby destabilizing the system and generating a highly unfolded state. At higher concentration of $x_{eth} \approx 0.35$ to 0.40, again the ethanol clustering starts forming in the system, making free ethanol molecules less abundant, thus stabilizing partially unfolded structures. In this context, it should be mentioned that the partially unfolded structures obtained at

dx.doi.org/10.1021/jp406255z | J. Phys. Chem. B 2013, 117, 15625−15638

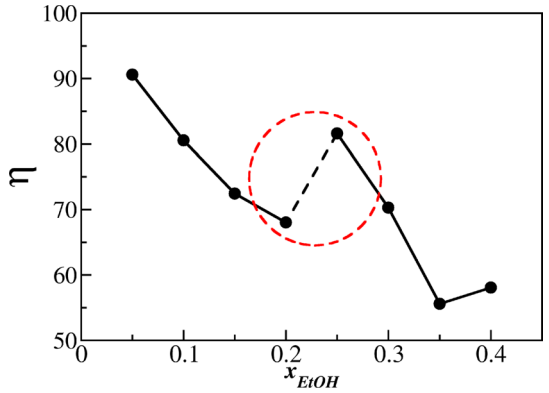


Figure 10. Plot of order parameter η as a function of mole fraction of ethanol, where η is defined as percentage of water in the first hydration shell around the protein. It is to be noted that with increasing ethanol concentration water content in the first hydration shell (within ∼0.55 nm cutoff) significantly decreases. However, at $x_{eth} \approx 0.25$, the percentage of water in the first hydration shell markedly increases, signifying the reappearance of water-like environment around the protein, thereby promoting better solvation of the same.

each concentration are quite different from each other, with the partially melted third helix and stable first helix at some concentration and vice versa in others. The presence of a number of such stable partially unfolded states can be attributed to the fact that in protein−solvent system a large number of local minima are encountered by the protein while traversing the pathway toward unfolding. In binary mixtures, due to diverse interactions of the protein and cosolvents, these minima are not separated by large free-energy barriers. In such cases, a small change in composition of the solvent is prone to bring about large conformational fluctuation in the protein, which gets easily trapped in such local minima. Such incident is expected to occur in this particular system, thereby stabilizing different conformational forms of protein at different concentrations. It is worth mentioning that for being known as the fastest folding protein, HP-36 often serves as the bridge between experimental and theoretical studies of protein folding. In recent experimental studies, the presence of intermediates for this protein is investigated by monitoring conformational energy fluctuations using triplet−triplet energy-transfer method.[47,48] They determine the rate-determining step as well as the corresponding rate constants for the folding process. The folding free-energy landscape observed in this case also consists of multiple minima, signifying the presence of intermediates in the folding pathway. However, to the best of our knowledge, no significant experimental study on conformational dynamics of HP-36 in binary mixtures has been reported to date.

IV. THEORETICAL ANALYSIS OF SOLVENT DEPENDENCE OF UNFOLDING

Bryngelson−Wolynes theory of folding was developed in terms of two order parameters, namely, the fraction of native contacts, η, and the radius of gyration of the protein, R_g.[1,2] Corresponding free-energy surface is characterized by multiple minima, with the deepest minima ideally corresponding to the native state under normal stability conditions of the folded protein. For example, for HP-36 under ambient conditions, the deepest minimum is the native state, with η close to unity and small value of R_g. As the protein is made to unfold, in the present cases, through increase in DMSO or EtOH concentration in the binary mixtures, the free-energy surface changes, with relative stability of the native state first decreasing with the increase in solute composition. The first major act of unfolding is the melting of the hydrophobic cluster formed by phenylalanines at 7, 11, and 18 positions. However, the details seem to depend critically on solvent conditions. In particular, the sequence of unfolding depends on relative depths of various minima and is hard to quantify. Solvent composition dependence allows one to probe the relative minima.

We have recently proposed a theoretical scheme to describe unfolding that combines aspects of Bryngelson−Wolynes theory with Marcus theory of electron transfer.[33,49−52] The basic idea is to introduce (i) a set of local structural order parameters $\{\eta_i\}$ to describe formation of native secondary structures like helices and beta strands and (ii) another set of order parameters, $\{R_i\}$, to describe pair separation distances between amino acid residues that form tertiary native contacts in the folded state.

For HP-36, we may retain only three order parameters, η_1, η_2, and η_3, to describe the formation of three helices. Simulations show that these three helices form and break at different stages.

Determination of the distance set $\{R_i\}$ is a bit more difficult. Simulations again show that at least for HP-36 only a few such distances are practically important, like the distance between Phe-11 and Phe-18 and that between Ala-9 and Leu-35.[33] One can thus describe the unfolding process in terms of two sets of order parameters: $\{\eta_i\}$and $\{R_i\}$.

The advantage of introducing such an extended set of order parameters is that we can now describe relative stability of different secondary structures in terms of the relative minima of free energy, $F(\{\eta_i\}, \{R_i\})$. These minima are a function of the solvent conditions. We can also now describe coupling between the two order parameters, within and across the set. Melting of HP-36 in the water−DMSO mixture involves cooperative interaction between the breaking of helix 2 and separation between Phe-11 and Phe-18.

We next extend the two-state model of protein folding/unfolding in terms of two state models for each of the order parameters, with minima located at$\{\eta_i^{UF}\}$and $\{R_i^{UF}\}$ for the unfolded or extended state and $\{\eta_i^F\}$ and $\{R_i^F\}$ for the folded state, respectively.

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

We now follow the celebrated Marcus theory of electron transfer to incorporate essential features of the free-energy surface in a rate constant.[49−52] Let us consider the first step as separation between Phe-11 and Phe-18, accompanied by melting of the second helix. Because these two order parameters $(\{R_i^F\},\{\eta^F\})$ seem to be coupled and seem to vary in unison, we now consider only the separation $R_{11,18}$ as the order parameter to describe the initial stages of unfolding. The relevant free-energy gap between the open state and associated compact state where phenylalanines are in contact is denoted by$\Delta G^0(\{R'\},\{\eta'\})$. The primes on R and η indicate that other values remain the same as those in the native state. Marcus theory requires a change in the value of the order parameter that is here equal to $R_{11,18} - a$, where a is the separation in closed, native state. Marcus theory then provides the following expression for the activation energy of the initial melting process[49−51]

dx.doi.org/10.1021/jp406255z | J. Phys. Chem. B 2013, 117, 15625−15638

$$E_{act} = \frac{[\Delta G^0(\{R'\}, \{\eta'\}) + \lambda_{11,18}]^2}{4\lambda_{11,18}} \quad (1)$$

where we assume

$$\lambda_{11,18} = \frac{1}{2}\omega^2(R_{11,18} - a)^2 \quad (2)$$

As mentioned, ΔG^0 is the free-energy gap between two minima when the other values of the two-order parameters are kept fixed at respective values in the native state and hence contains the effect of composition dependence. ω is the harmonic frequency of the free-energy surface.

In neat water, the native state configuration is more stable. Thus, the barrier is large. Also, the extended state is metastable. So, even if it forms by fluctuation, it goes back to the native, compact state.

In the case of water−DMSO binary mixture, as DMSO concentration increases, the unfolded state becomes more stable, and the barrier toward the formation of this state decreases. At large ΔG^0 (at $x_{DMSO} = 0$), we envisage a crossover to barrierless dynamics, as in Marcus theory of electron transfer.

We can now understand the difference between the DMSO and EtOH mixtures by using the above theory. Both of the two species lead to stabilization of the unfolded state, albeit more for DMSO than for EtOH, but significant difference exists at larger composition. For DMSO, the other secondary structures also melt, as the extended or molten states gain more stability in this case. However, for EtOH, the story is different. Here the molten or coiled state of the amino acid residues forming the third helix, for example, never becomes more stable than the native state in different EtOH concentrations. (See the schematic energy profile in Figure 11.) Thus, it melts only partially.

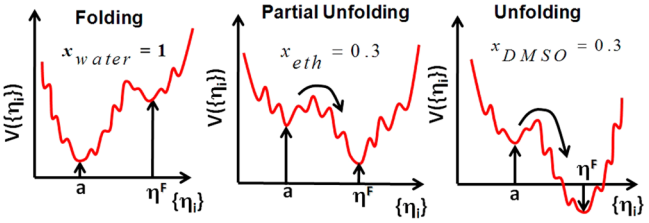


Figure 11. Schematic energy profile of fraction of native contact involved in folding−unfolding transition. $V(\{\eta_i\})$ represents the potential of mean force (PMF), where $\{\eta_i\}$ is the distinct fraction of native contact order parameter. $\eta_i = a$ is the separation in native state, while the $\eta_i = \eta_F$ gives the separation at the final state during unfolding. Note that in the case of ethanol the final state is a partially unfolded state. However, for DMSO, the final stable state is a completely unfolded state. The above scheme is given for the two cosolvents when their concentration reaches 0.30 mol fraction range.

We now develop a simple theory to quantify the above discussions. Because 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 time evolution of the probability distribution[53] $P(\{R_i\},\{\eta_i\},t)$:

$$\frac{\partial}{\partial}P(\{R_i\}, \{\eta_i\}, t) = \sum_{i=1}^{N_{sp}} D_i \frac{1}{R_i^2}\frac{\partial}{\partial R_i} e^{-\beta V_i(\{R_i\},\{\eta_i\})}\frac{\partial}{\partial R_i}$$
$$e^{\beta V_i(\{R_i\},\{\eta_i\})}P - \sum_{i=1}^{N_{ij}} \Gamma_{\eta_i}\frac{\partial V(\{R_i\}, \{\eta_i\})}{\partial \eta_i} \quad (3)$$

where we have assumed decoupling between the order parameters. This is certainly an approximation but seems reasonable as a first step. However, the potential functions retain a dependence on other distances implicitly.

The potential functions, $V(\{R_i\},\{n_i\})$, are to be regarded as potential of mean force (PMF) of the type previously described.[33] Thus, in the presence of DMSO, the energy surface,$(R_{11,18},\eta_1,\eta_2)$, undergoes a sharp change compared with that of water or ethanol, where a minimum at contact $(R_{11,18})_{min} = a$ for mole fraction is $x_{DMSO} = 0$ replaced by a minimum at a larger distance. Therefore, unfolding at large denaturant concentration is, in general, a relaxation in a potential energy surface with a composition-dependent activation barrier.

Adaptation of Marcus theory in 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, at different cosolvent environment. This in turn explores the tuning of the barrier height that holds the native state.

The progression of unfolding often begins with the step that induces, under appropriate conditions, subsequent separation of other contacts, signifying a high degree of cooperation in the unfolding process. For folding to unfolding transition of HP-36, the sequential steps of unfolding of Helix-2 hold the distinct characteristics of global unfolding in both DMSO and ethanol. Nevertheless, in both cases, we compare the partial unfolding process of the full protein, where at least the Helix-2 undergoes complete unfolding. To observe the barrier sensitivity to the secondary structure of HP36, we evaluate the qualitative free-energy landscape in (η, R_g) order plane (Figure 12). The Figure shows the second helix melting process at $x_{DMSO} = 0.25$ and $x_{EtOH} = 0.30$. In the case of DMSO, the second helix melting process involves different stable intermediates (I_D1, I_D2) expressing two distinct minima separated by a barrier (>1 kJ/mol). On the contrary, in the case of ethanol, different intermediates (I_E1, I_E2, I_E3) evolve but are separated by a shallow barrier (<0.4 kJ/mol). This signifies that barrier separation of the intermediates directs the pathway of unfolding. Here, for DMSO, the barrier separation between the intermediates reveals that once it reaches I_D2 from I_D1, it is unlikely that it can again go back toward I_D1. However, for ethanol, the small separation of barrier among the intermediates signifies that I_E2 has a significant probability to shift toward I_E1 or I_E3. Thus DMSO makes the second helix melting somewhat more efficient than ethanol, where the latter makes several traps.

V. CONCLUSIONS

The pathway of unfolding of a protein from stable native state to the extended unfolded (or, partially unfolded) state is often a complex one, comprising unstable as well as partially stable intermediates that constitute the large number of local minima present in the free-energy landscape. These intermediates are very hard to capture in general by means of experiments because of their very short lifetimes. Recent developments in 2D-IR spectroscopy are able to detect the fast dynamics of proteins beautifully[3] and thus can be expected to identify the

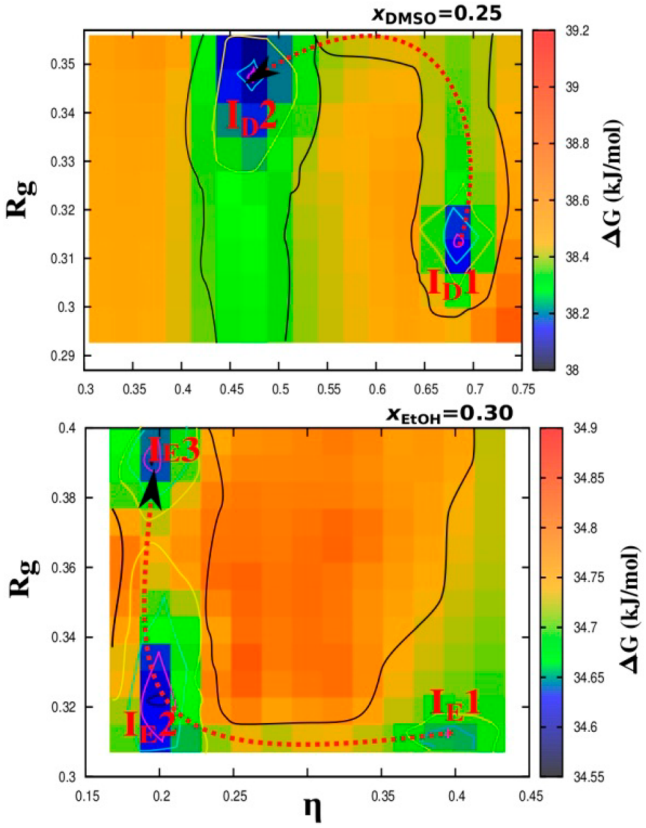


Figure 12. Contour map of the two-order parameter (η, R_g)-based free-energy landscape of the secondary structure involving the second helix of HP-36. For folding to unfolding transition of HP-36, the sequential step of unfolding of Helix-2 holds distinct characteristics of global unfolding both in DMSO and ethanol. In the two cases, we compare the partial unfolding process of the full protein where Helix-2 undergoes complete unfolding. The upper panel of the Figure shows the 2nd helix melting process at x_{DMSO} = 0.25. The presence of two distinct minima separated by a barrier corresponds to the emergence of different intermediates (I_D1, I_D2) from folding to unfolding pathway, as indicted in the Figure (arrowed line). The Free energy separation (>1 kJ/mol) between two minima is evident from the energy landscape. The lower contour plot also demonstrates distinguishable unfolding path (arrowed line) evolving different intermediates (I_E1, I_E2, I_E3) but separated by a shallow barrier (<0.4 kJ/mol). This plot shows the second helix melting process at x_{EtOH} = 0.30.

metastable intermediates formed in the unfolding pathway. The unfolding facilitated by different external factors may give rise to different intermediates. While thermal denaturation brings deformation of the protein as well as modification of water structure that incorporate further conformational changes; chemical denaturation by using varied composition of cosolvents brings in static and dynamical changes in the protein by employing more subtle changes in intermolecular interactions among the protein and solvent(s).

In this work, we study the effects of an important cosolvent ethanol on a 36 residue protein, chicken villin headpiece subdomain, commonly known as HP-36. We vary the composition of the solvent and intend to see what type of dynamical changes in the structure of the protein take place. We indeed find that with change in ethanol concentration, considerable structural change is incorporated in the system. At low concentration of ethanol, partial unfolding of the protein occurs, accompanied by the signature of deformation of second helix. At $x_{eth} \approx 0.05$ to 0.10, we also find that the hydrophobic core groups (Phe-7, Phe-11, Phe-18) are considerably separated from each other, which primarily forms the basis of unfolding. In fact, the initial pathway of unfolding is found to be similar to that of thermal denaturation as well as in water−DMSO. On increasing ethanol concentration further to $x_{eth} \approx 0.15$ to 0.20, we see that the structural change is not significant, and here we find that the largely separated hydrophobic contacts formed by the three phenylalanine groups start coming closer again. At $x_{eth} \approx 0.25$, surprisingly a native-like state appears, with the consequent signatures in the R_g, RMSD values, as well as in the average fraction of native contacts present in the structure. At this concentration, the hydrophobic core is associated in a similar way as that of native structure. This gives rise to a very unusual phenomenon, as we expect gradual unfolding of the protein with increasing ethanol concentration due to interaction between the hydrophobic groups of protein and cosolvent. After that, on further increase in ethanol concentration, different partially unfolded states of the protein are obtained, and full denaturation is not accomplished, even at ethanol concentration as high as $x_{eth} \approx 0.40$. Thus, we show that although initiation of unfolding of HP-36 follows a universal path, further progress in the unfolding landscape may take on different routes depending on the environment. We compare the results with that obtained from unfolding study of the same protein in water−DMSO binary mixture to show the similarity and differences in the pathway.

The reason for the occurrence of such unusual phenomenon at $x_{eth} \approx 0.25$ is attributed to the microheterogeneous phase separation in the water−ethanol solvent system itself that is brought about by the self-association of ethanol molecules through hydrogen bonding at that concentration. Evidence of such phenomena arising has been manifested by a decreasing trend in diffusion coefficient of ethanol as well as an increasing number of ethanol−ethanol hydrogen bonds in that concentration range. We also show the appearance of different partially unfolded states at different ethanol concentration, giving signature of multistage dynamics. These particular states signify various local minima in the free-energy landscape of the protein−solvent system that are generated due to diverse mode of interactions between them. We have also provided a theoretical understanding of the phenomenon by employing Marcus theory combined with the vital aspects of Bryngelson−Wolynes theory. Our results essentially suggest that by tuning solvent concentration the dynamical evolution in the structure of protein can be significantly altered. This might effectively help in regulating enzymatic activity by controlling concentration of cosolvents, although further verification of the results is needed.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Plot of average radius of gyration of high resolution crystal structure of chicken villin headpiece. Equilibrium snapshots of chicken villin headpiece simulated at x_{eth}= 0.25 (PDB code 1YRF). Distribution of distances between residues of the hydrophobic core with increasing DMSO concentration. Probability distribution of distance between Phe-7 and Phe-18, Phe-11 and Phe-18, and Phe-7 and Phe-11. Variation of average distance between these three pairs with changing DMSO concentration. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

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

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was partially supported by grants from BRNS and DST. We thank JC Bose Fellowship for a partial support.

REFERENCES

(1) Bryngelson, D. J.; Wolynes, G. P. Intermediates and Barrier Crossing in a Random Energy Model (with Applications to Protein Folding). *J. Phys. Chem.* **1989**, *93*, 6902−6915.
(2) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167−195.
(3) Chung, J. K.; Thielges, M. C.; Fayer, M. D. Dynamics of the Folded and Unfolded Villin Headpiece (HP35) Measured with Ultrafast 2D IR Vibrational Echo Spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 3578.
(4) Chung, J. K.; Megan, C.; Thielges, M. C.; Lynch, S. R.; Fayer, M. D. Fast Dynamics of HP35 for Folded and Urea-Unfolded Conditions. *J. Phys. Chem. B* **2012**, *116*, 11024−11031.
(5) Thielges, M. C.; Fayer, M. D. Protein Dynamics Studied with Ultrafast Two-Dimensional Infrared Vibrational Echo Spectroscopy. *Acc. Chem. Res.* **2012**, *45*, 1866−1874.
(6) Kim, S.; Chung, J. K.; Kwak, K.; Bren, K. L.; Bagchi, B.; Fayer, M. D. Native and Unfolded Cytochrome c—Comparison of Dynamics using 2D-IR Vibrational Echo Spectroscopy. *J. Phys. Chem. B.* **2008**, *112*, 10054−10063.
(7) Chung, H. S.; Ganim, Z.; Jones, K. C.; Tokmakoff, A. Transient 2D IR Spectroscopy of Ubiquitin Unfolding Dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14237−14242.
(8) Zanni, M.; Hochstrasser, R. M. Two-Dimensional Infrared Spectroscopy: A Promising New Method for the Time Resolution of Structures. *Curr. Opin. Struct. Biol.* **2001**, *11*, 516.
(9) Leopold, P. E.; Montal, M.; Onuchic, J. N. Protein Folding Funnels: a Kinetic Approach to the Sequence-Structure Relationship. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8721−8725.
(10) Karplus, M.; Weaver, D. L. Protein-Folding Dynamics. *Nature* **1976**, *260*, 404−406.
(11) Kim, P. S.; Baldwin, R. L. Specific Intermediates in the Folding Reactions of Small Proteins and the Mechanism of Protein Folding. *Annu. Rev. Biochem.* **1982**, *51*, 459−489.
(12) Weissman, J. S.; Kim, P. S. Reexamination of the Folding of BPTI: Predominance of Native Intermediates. *Science* **1991**, *253*, 1386.
(13) Radford, S. E.; Dobson, C. M.; Evans, P. A. The Folding of Hen Lysozyme Involves Partially Structured Intermediates and Multiple Pathways. *Nature.* **1992**, *358*, 302.
(14) Jackson, S. E.; Fersht, A. R. Folding of Chymotrypsin Inhibitor 2. 1. Evidence for a Two-State Transition. *Biochemistry.* **1991**, *30*, 10428.
(15) Koizumi, M.; Hirai, H.; Onai, T.; Inoue, K.; Hirai, M. Collapse of the Hydration Shell of a Protein Prior to Thermal Unfolding. *J. Appl. Crystallogr.* **2007**, *40*, 175−178.
(16) Li, W.; Zhou, R.; Mu, Y. Salting Effects on Protein Components in Aqueous NaCl and Urea Solutions: Toward Understanding of Urea-Induced Protein Denaturation. *J. Phys. Chem. B.* **2012**, *116*, 1446−1451.
(17) Zhou, R. Free Energy Landscape of Protein Folding in Water: Explicit vs. Implicit Solvent. *Proteins: Struct., Funct., Genet.* **2003**, *53*, 148−161.
(18) Levy, Y.; Onuchic, J. N. Water Mediation in Protein Folding and Molecular Recognition. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 389−415.
(19) Pizzitutti, F.; Marchi, M.; Sterpone, F.; Rossky, P. J. How Protein Surfaces Induce Anomalous Dynamics of Hydration Water. *J. Phys. Chem. B* **2007**, *111*, 7584−7590.
(20) Jha, S. K.; Dhar, D.; Krishnamoorthy, G.; Udgaonkar, J. B. Continuous Dissolution of Structure during the Unfolding of a Small Protein. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 11113−8.
(21) Bhattacharya, K. Nature of Biological Water: a Femtosecond Study. *Chem. Commun.* **2008**, 2848−2857.
(22) Bagchi, B.; Jana, B. Solvation Dynamics in Dipolar Liquids. *Chem. Soc. Rev.* **2010**, *39*, 1936−1954.
(23) Chandra, A.; Bagchi, B. Molecular Theory of Dielectric Relaxation in a Dense Binary Dipolar Liquid. *J. Phys. Chem.* **1991**, *95*, 2529−2535.
(24) Chandra, A.; Bagchi, B. Molecular Theory of Solvation and Solvation Dynamics in a Binary Dipolar Liquid. *J. Chem. Phys.* **1991**, *94*, 8367.
(25) Roy, S.; Banerjee, S.; Biyani, N.; Jana, B.; Bagchi, B. Theoretical and Computational Analysis of Static and Dynamic Anomalies in Water−DMSO Binary Mixture at Low DMSO Concentrations. *J. Phys. Chem. B* **2010**, *115*, 685−692.
(26) Banerjee, S.; Roy, S.; Bagchi, B. Enhanced Pair Hydrophobicity in the Water−Dimethyl sulfoxide (DMSO) Binary Mixture at Low DMSO Concentrations. *J. Phys. Chem. B* **2010**, *114*, 12875−12882.
(27) Banerjee, S.; Ghosh, R.; Bagchi, B. Structural Transformations, Composition Anomalies and a Dramatic Collapse of Linear Polymer Chains in Dilute Ethanol−Water Mixtures. *J. Phys. Chem. B* **2012**, *116*, 3713−3722.
(28) Juurinen, I.; Nakahara, K.; Ando, N.; Nishiumi, T.; Seta, H.; Yoshida, N.; Morinaga, T.; Itou, M.; Ninomiya, T.; Sakurai, Y.; et al. Measurement of Two Solvation Regimes in Water-Ethanol Mixtures Using X-Ray Compton Scattering. *Phys. Rev. Lett.* **2011**, *107*, 197401−5.
(29) Ortore, G. M.; Mariani, P.; Carsughi, F.; Cinelli, S.; Onori, G.; Teixeira, J.; Spinozzi, F. Preferential Solvation of Lysozyme in Water/Ethanol Mixtures. *J. Chem. Phys.* **2011**, *135*, 245103−9.
(30) Bhakkuni, V. Alcohol-Induced Molten Globule Intermediates of Proteins: Are They Real Folding Intermediates or Off Pathway Products? *Arch. Biochem. Biophys.* **1998**, *357*, 274.
(31) Goodwin, A. C.; Allen, T. J.; Oslick, S. L.; McClure, K. F.; Lee, J. H.; Kemp, D. S. Mechanism of Stabilization of Helical Conformations of Polypeptides by Water Containing Trifluoroethanol. *J. Am. Chem. Soc.* **1996**, *118*, 3082−3090.
(32) Yoshida, K.; Kawaguchi, J.; Lee, S.; Yamaguchi, T. On the Solvent Role in Alcohol-Induced α-Helix: Formation of Chymotrypsin Inhibitor 2*. *Pure Appl. Chem.* **2008**, *80*, 1337−1347.
(33) Roy, S.; Bagchi, B. Chemical Unfolding of Chicken Villin Headpiece in Aqueous Dimethyl Sulfoxide Solution: Co-solvent Concentration Dependence, Pathway and Microscopic Mechanism. *J. Phys. Chem. B.* **2013**, *117*, 4488−4502.
(34) McKnight, C. J.; Matsudaira, P. T.; Kim, P. S. NMR Structure of the 35-Residue Villin Headpiece Subdomain. *Nat. Struct. Biol.* **1997**, *4*, 180−184.
(35) McKnight, C. J.; Doering, D. S.; Matsudaira, P. T.; Kim, P. S. A Thermostable 35-Residue Subdomain within Villin Headpiece. *J. Mol. Biol.* **1996**, *260*, 126−134.
(36) Doering, D. S.; Matsudaira, P. Cysteine Scanning Mutagenesis at 40 of 76 Positions in Villin Headpiece Maps the F-Actin Binding Site and Structural Features of the Domain. *Biochemistry* **1996**, *35*, 12677−12685.
(37) Srinivas, G.; Bagchi, B. Foldability and the Funnel of HP-36 Protein Sequence: Use of Hydropathy Scale in Protein Folding. *J. Chem. Phys.* **2002**, *116*, 8579−8587.
(38) Mukherjee, A.; Bagchi, B. Correlation between Rate of Folding, Energy Landscape, and Topology in the Folding of a Model Protein HP-36. *J. Chem. Phys.* **2003**, *118*, 4733−4747.
(39) Bandyopadhyay, S.; Chakraborty, S.; Bagchi, B. Atomistic Simulation Study of the Coupled Motion of Amino Acid Residues and Water Molecules around Protein HP-36: Fluctuations at and around the Active Sites. *J. Phys. Chem. B.* **2004**, *108*, 12608−12616.

(40) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(41) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656.
(42) Hoover, W. G. Canonical dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A.* **1985**, *31*, 1695.
(43) Nose, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, *81*, 511.
(44) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182.
(45) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed.; Academic Press: San Diego, CA, 2002.
(46) Rajgaria, R.; McAllister, S. R.; Floudas, C. A. A Novel High Resolution C^{α}- C^{α} Distance Dependent Force Field Based on a High Quality of Decoy Set. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 726−741.
(47) Reinera, A.; Henkleinb, P.; Kiefhabera, T. An Unlocking/Relocking Barrier in Conformational Fluctuations of Villin Headpiece Subdomain. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 4955−4960.
(48) Beauchampa, K. A.; Ensignb, D. L.; Das, R.; Pande, V. S. Quantitative Comparison of Villin Headpiece Subdomain Simulations and Triplet−Triplet Energy Transfer Experiments. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 12734−10.
(49) Marcus, R. A. Chemical and Electrochemical Electron-Transfer Theory. *Annu. Rev. Phys. Chem.* **1964**, *15*, 155−196.
(50) Marcus, R. A. On the Theory of Electron-Transfer Reactions. VI. Unified Treatment for Homogeneous and Electrode Reactions. *J. Chem. Phys.* **1965**, *43*, 679−701.
(51) Marcus, R. A.; Sutin, N. Electron Transfers in Chemistry and Biology. *Biochim. Biophys. Acta* **1985**, *811*, 265−322.
(52) Sumi, H.; Marcus, R. A. Dynamical Effects in Electron Transfer Reactions. *J. Chem. Phys.* **1986**, *84*, 4894−4914.
(53) Bagchi, B. *Molecular Relaxation in Liquids*; Oxford University Press: New York, 2012.