THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

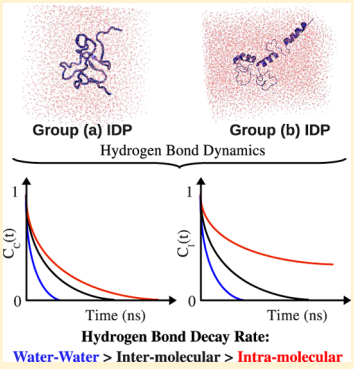
pubs.acs.org/JPCB

Hydrogen Bond Dynamics in Intrinsically Disordered Proteins

Nidhi Rawat and Parbati Biswas*

Department of Chemistry, University of Delhi, Delhi 110007, India

ABSTRACT: Hydrogen bond dynamics is used to investigate the internal motions and structural plasticity of intrinsically disordered (ID) proteins. Group a represent completely disordered proteins, while group b proteins are comprised of regular secondary structures linked by flexible disordered regions. Molecular dynamics simulations of two different groups of ID proteins provide an insight into the hydrogen bond dynamics via the evaluation of the continuous and intermittent time autocorrelation functions. The intermolecular hydrogen bonds between the residues of the ID proteins and water record a short lifetime in both groups of proteins. The intermolecular hydrogen bonds relax faster at a constant rate compared to that of the intramolecular hydrogen bonds whose rate of decay fluctuate during the entire simulation trajectory. The simulations reveal that the intramolecular hydrogen bonds have a longer lifetime in group b proteins compared to those in group a proteins. The hydrophilic residues in ID proteins form stable persistent intramolecular hydrogen bonds as compared to the hydrophobic residues and help to maintain the dynamic equilibrium among the interconvertible conformations.



I. INTRODUCTION

Hydrogen bonds play a pivotal role in governing the stability and structural organization in folded proteins.[1] The energetic contribution of intramolecular hydrogen bonding ranges from partially stabilizing and destabilizing[2,3] to destabilizing[4] thus creating enough controversy in determining the thermodynamic and dynamic properties of proteins in aqueous solutions. The concerted formation and breakage of hydrogen bonds between the residues of proteins or between the protein and water at the surface modulates the flexibility of the protein and controls its specific function. The dynamic behavior of hydrogen bonds in globular proteins and DNA in aqueous solutions provide better insight to understand various aspects of their function and stability.[5,6] The cooperative nature of the hydrogen bond imparts strong local anisotropic interactions between the hydrogen and acceptor forming the bond. The strength of hydrogen bonds may be estimated to be approximately 5−10 times the thermal energy, measured in units of k_BT, where k_B is the Boltzmann constant. For higher fluctuations of the thermal energy, the hydrogen bonds have a finite lifetime and survive only up to the order of picoseconds. However, most hydrogen bonds remain intact at room temperature.

Intrinsically disordered proteins (IDP's) may be represented as a dynamic ensemble of flexible interconverting conformations rather than a well-defined folded structure under physiological conditions.[7,8] Low mean hydrophobicity combined with high net charge is an important prerequisite for the absence of a unique compact folded structure.[9,10] IDP's lack a regular hydrophobic core, with hydrophobic residues being scattered throughout the protein.[11] IDP's are promiscuous binding partners, which serve as flexible linkers between various functional domains to facilitate binding diversity and molecular recognition.[12,13] Recent studies exhibit that while some IDP's undergo a large conformational change upon binding specific targets, others have preformed structural features resembling a considerable number of the bound state conformations for the given structural ensemble.[14] The instantaneous fluctuations of the intramolecular and intermolecular hydrogen bonds of IDP's may help in understanding this rapid conformational rearrangement and the specificity of interactions involved in binding targets, as hydrogen bonds form long-range electrostatic interactions that stabilize the structure of the protein.[15] The strength of hydrogen bonds is context-dependent and depends on the polarity of the hydrogen bond. A recent study shows that the hydrogen bonds are relatively stronger in a nonpolar microenvironment as compared to a polar one.[16] Electrostatic interaction is predominant in hydrogen-bonded systems. Hence the polarization of the hydrogen bond buried in the core of a globular protein would be different from that of a solvent accessible one.

Accurate description of this context-dependence of the hydrogen bond is an important determinant in evaluating the energetics of hydrogen bonds in proteins with well-defined tertiary structures. The fast conformational fluctuations and the internal dynamics of IDP's may be monitored through the hydrogen bond dynamics, where the transient formation and breakage of the hydrogen bonds governs the structural plasticity of the protein and enables it to attain a specific function. The dynamics of the hydrogen bond may be investigated experimentally by different techniques such as ultrafast infrared (IR) spectroscopy, Raman scattering, inelastic neutron scattering, and depolarized light scattering.[17] Molecular dynamics (MD) simulations provide a powerful tool to probe

Received: February 7, 2014
Revised: February 26, 2014
Published: February 26, 2014

   dx.doi.org/10.1021/jp5013544 | J. Phys. Chem. B 2014, 118, 3018−3025

the hydrogen bond dynamics within the atomic resolution by calculating various hydrogen bond time autocorrelation functions.[18−21] The dynamics of hydrogen bonds are classified as either continuous or intermittent, which determines the lifetime and the structural relaxation of particular hydrogen bonds, respectively. In this work, hydrogen bond dynamics is evaluated in terms of the time autocorrelation functions for the intramolecular hydrogen bonds between ID proteins and intermolecular hydrogen bonds between ID proteins and the hydration water at the surface of the protein. The remaining text is organized as follows: Section II describes the methodology for MD simulations and the procedure for calculating the time autocorrelation functions for the hydrogen bonds. Section III presents the results of our investigation and discusses its implications. Section IV summarizes the main conclusions of this study.

II. METHODS

Protein Selection. ID proteins are classified into two groups: (a) proteins which consistently lack well-defined secondary structure, i.e., they are completely disordered proteins; and (b) proteins with regular secondary structures linked by flexible disordered regions.[22−24] The proteins chosen from group a are 4MT2, a metallothionine of 61 residues, and 2ABX, a 74 residue neurotoxin. Group b proteins are comprised of 1HQ1, a 105 residue ribonucleoprotein, and 1CD3 a 120 residue scaffolding protein.[25−28] These proteins have considerably low proline and glycine content, so they are amyloidogenic proteins rather than elastomeric. The function of these proteins are primarily governed by hydrophobic association as 4MT2, 2ABX, 1HQ1, and 1CD3 are composed of 3, 1, 7, and 8% glutamine residues, respectively.[29,30] 1HQ1 is a RNA binding protein and 1CD3 is complexed with single stranded DNA and other scaffolding proteins, respectively. The three-dimensional (3D) structures of these proteins are obtained from the Protein Data Bank (PDB).[31] The position coordinates of the disordered regions of group b proteins are missing in their respective PDB files.[32] The initial template structure of the uncomplexed protein is modeled[33] to obtain these missing coordinates. The 3D structure of the uncomplexed protein is generated in PYMOL with defined secondary structures similar to that of the complexed protein.[34] The resultant structure is made compact, and its energy is minimized by simulating in vacuum at 300 K using AMBER 9 for 50 ns. The radius of gyration (R_G) of the ordered part of the uncomplexed protein has the maximum similarity to that of the complexed protein. This yields the initial template structure for simulation in explicit solvent with the Transferable Intermolecular Potential 3 Point (TIP3P) water model. The initial structure for the group a proteins is obtained from the PDB as the position coordinates of all residues are well-defined.

Molecular Dynamics. Molecular dynamics simulations of the selected proteins were performed by using AMBER 12 with TIP3P water as the explicit solvent.[35] The missing hydrogen atoms were added with LEAP subroutine. Each protein was solvated in a cubic box with TIP3P water. The dimensions of the simulation box used for 4MT2, 2ABX, 1HQ1, and 1CD3 were 56.8 × 53.1 × 63.74, 64.88 × 53.03 × 62.33, 88.59 × 63.13 × 65.36, and 83.53 × 76.66 × 54.41 Å^3, respectively. Charges of the solvated proteins were neutralized by adding sodium or chloride ions, respectively, depending upon the charge of the ID proteins. AMBER99SB force field was used with the periodic boundary conditions. This force field presents a careful reparametrization of the backbone torsion terms in ff99 and achieves a better balance of four basic secondary structure elements.[36] Particle mesh Ewald (PME) in Amber12 was used to treat long-range electrostatic interactions. The cutoffs for electrostatics and van der Waals interactions were set to be 8 and 9 Å, respectively. The motions of the hydrogen atoms were constrained through the SHAKE algorithm by removing the bond stretching motion, consequently allowing a larger time step of 0.002 ps to be used. Each system was energy minimized twice: the solvent is energy minimized first by keeping the protein constrained followed by minimizing the energy of the whole system. The solvated protein is initially simulated at low temperature of 100 K, while the temperature is gradually raised to 300 K for 10 ps at constant volume. Equilibration for 2 ns is performed at a constant temperature of 300 K and a constant pressure of 1 bar (*NPT*) followed by the equilibration for another 5 ns by maintaining constant temperature and volume (*NVT*). Constant temperature is maintained via Langevin thermostat with a coupling constant of 2 ps while constant pressure is maintained through weak coupling to an isotropic pressure bath with a coupling constant of 1 ps.[37] This equilibrated conformation is used for an extended simulation of 25 ns in the microcanonical (NVE) ensemble. This time of simulation is long enough to investigate the sub-nanosecond dynamics of the hydrogen bond. The temperature increases steadily from 298 to 310 K throughout the simulation trajectory. The coordinates of the solvated proteins are saved after every 2 ps for further analysis.

In Figure 1 R_G and root-mean-square deviation (RMSD) are plotted as functions of time for group a and b proteins. For

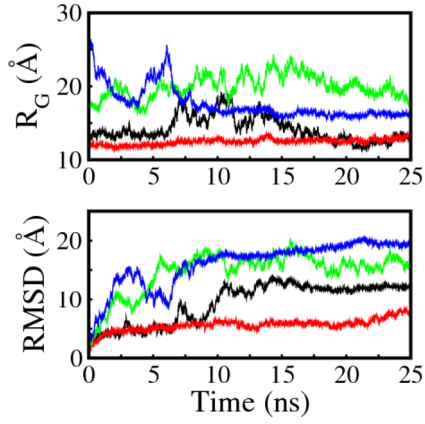


Figure 1. Radius of gyration (R_G) and RMSD as a function of simulation time for 4MT2 (black), 2ABX (red), 1HQ1 (green), and 1CD3 (blue).

4MT2, R_G ranges between 11.85 and 18.88 Å with an average value of 13.5 Å. The values of R_G steadily increase from ∼12 to 13.5 Å for 2ABX throughout the entire simulation. For both group a proteins, the RMSD values are within permissible limits for the disordered proteins.[38] The values of R_G fluctuate between 17.2 and 22.5 Å for 1HQ1, while for 1CD3, R_G decreases from 25.2 to 16.7 Å until 10 ns and remains constant thereafter. The RMSD values reveal a similar overall trend for 1HQ1 and 1CD3. Values of R_G and RMSD reveal that the sizes of group a and group b proteins change continuously throughout the entire duration of simulation.

The size of ID proteins follows Flory's scaling law and varies as $R_G = \langle R_G^2 \rangle^{0.5} = (2.35 \pm 0.09)N^{\alpha}$, with the scaling exponent

dx.doi.org/10.1021/jp5013544 | J. Phys. Chem. B 2014, 118, 3018−3025

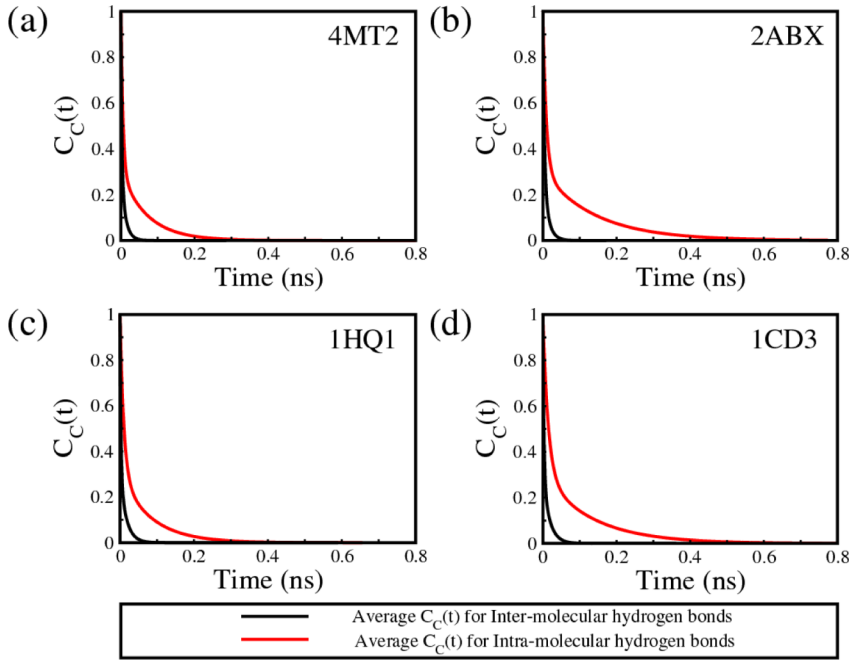


Figure 2. Average continuous autocorrelation function for intermolecular (black) and intramolecular (red) hydrogen bonds in (a) 4MT2, (b) 2ABX, (c) 1HQ1, and (d) 1CD3 as a function of simulation time.

ranging between 0.39 and 0.5, between 0.38 and 0.41, between 0.43 and 0.48, and between 0.5 and 0.41 for 4MT2, 2ABX, 1HQ1, and 1CD3, respectively. The values of these exponents closely match with those of the natively unstructured proteins[39] representing protein molecules in the intermediate solvent conditions.[40,41] This implies that ID proteins exist as an ensemble of conformations which depart considerably from those of an ideal random coil. RMSD values for both groups of proteins range between 5 and 20 Å, which is within the allowed experimental limit for disordered proteins and matches with the results of Monte Carlo simulations.[42,43] Higher RMSD values indicate that these proteins have increased conformational flexibility and undergo continuous rearrangement of the backbone structure as compared to those of the globular proteins.

Hydrogen Bond Dynamics. Molecular dynamics simulations quantitatively characterize the dynamics of hydrogen bonds in ID proteins in terms of different hydrogen bond time autocorrelation functions, initially proposed by Stillinger[44] and developed later by Luzar and Chandler.[45] A hydrogen bond between an acceptor and donor pair in the ID protein is defined in terms of the geometric constraint determined by the maximum and minimum distance between them. The maximum distance between the acceptor and donor forming a hydrogen bond must be less than 3.5 Å, while the distance between the hydrogen and the acceptor should be less than 2.5 Å. The angle between acceptor donor and hydrogen should be less than 35°.[46,47] The mean lifetime of the hydrogen bond is estimated by calculating the time-dependent autocorrelation function which reflects the presence/absence of hydrogen bonds between all possible donor−acceptor residue pairs.[48] Such autocorrelation functions may be defined as

$$C_x(t) = \frac{\sum_{i,j} S_{ij}(t_0)\, S_{ij}(t_0 + t)}{\sum_{i,j} S_{ij}(t_0)} \quad (1)$$

where $S_{ij}(t)$ measures the existence of the hydrogen bond at time t. The value of $S_{ij}(t)$ is unity when a hydrogen bond is present between the ith and jth residues at time t, and zero if the hydrogen bond is absent. The set of values S_{ij} for all donor−acceptor pairs specifies all hydrogen bonds present at any time instant t. The value of the autocorrelation function is unity at the initial time t_0. Two different types of autocorrelation functions, continuous (x = C) and intermittent (x = I), are calculated in this work. The continuous time autocorrelation function evaluates the lifetime of the hydrogen bond until it remains continuously bonded from t_0 to t and does not consider the formation of the same bond, even if it is subsequently re-formed at a later time instant. The intermittent correlation function gives the probability of a particular tagged hydrogen bond being intact at time t provided it was intact at the initial time t_0. Thus $C_I(t)$ is independent of the breaking of hydrogen bonds at any intermediate time instant and allows for the re-formation of broken bonds. The values of the hydrophobicity of the residues of the ID protein are obtained from the Eisenburg hydrophobicity consensus scale as provided in Table 1 of ref 47 for the analysis of the hydrogen bonds formed between hydrophobic and hydrophilic residues in 1CD3. Other interactions which stabilize ID proteins are the salt bridges formed between charged residues. In this study we have included the salt-bridge interactions as hydrogen bonds within the permissible geometric constraints. Thus most of the salt bridges are considered as hydrogen bonds.[49] Hydrogen bonds in ID proteins are expected to exhibit a different dynamics as compared to the globular proteins.

III. RESULTS AND DISCUSSION

The structural organizations of the water molecules at the surface of the ID proteins are correlated to the dynamics of the hydrogen bond network formed between water molecules and the residues of the ID proteins. The dynamics of intermolecular and intramolecular hydrogen bonds of ID proteins are evaluated in terms of the continuous autocorrelation function

dx.doi.org/10.1021/jp5013544 | J. Phys. Chem. B 2014, 118, 3018−3025

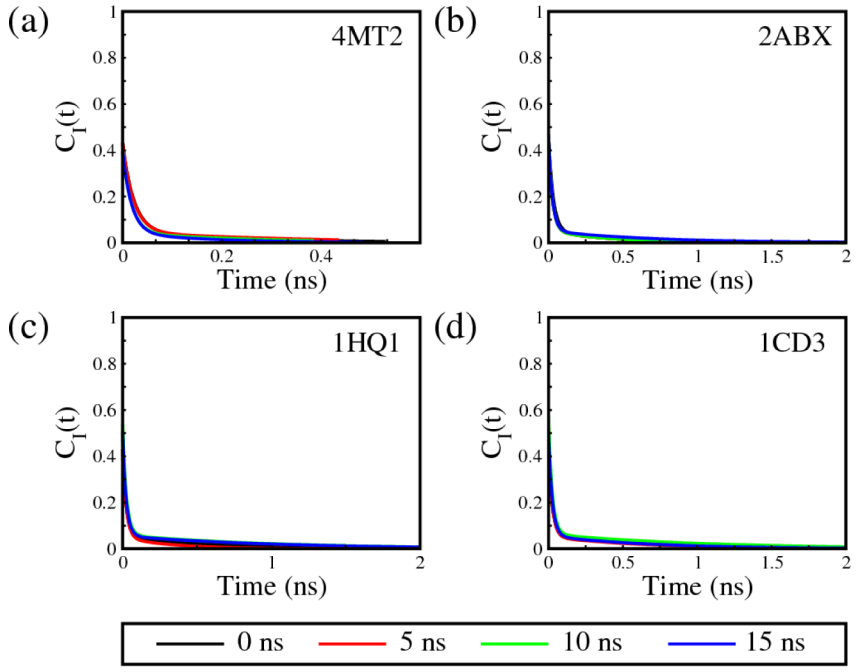


Figure 3. Intermittent autocorrelation function for intermolecular hydrogen bonds in (a) 4MT2, (b) 2ABX, (c) 1HQ1, and (d) 1CD3 as a function of simulation time.

$C_C(t)$ and the intermittent autocorrelation function $C_I(t)$, respectively. The time autocorrelation functions, $C_C(t)$ and $C_I(t)$, for a particular protein are calculated by monitoring specific hydrogen bonds at different time intervals, i.e., 0, 5, 10, and 15 ns respectively in the simulation trajectory. For the continuous time autocorrelation function, all calculations are carried out by averaging over different values of the relaxation time obtained at different time intervals for the same protein.

Parts a−d of Figure 2 represent the comparative decay of $C_C(t)$ for the intermolecular hydrogen bonds between the residue of ID proteins and water and the intramolecular hydrogen bonds between the residues of the protein as a function of the simulation time. For the intermolecular hydrogen bonds, $C_C(t)$ denotes the probability of a hydrogen bond that is formed between a specific pair of donor and acceptor at t_0 and is retained up to time t. The decay of $C_C(t)$ may be expressed as the sum of two exponentials

$$C_C(t) \propto A_1 \exp(-t/\tau_{C1}) + A_2 \exp(-t/\tau_{C2}) \quad (2)$$

where A_1 and A_2 are the time constant amplitude factors, while τ_{C1} measures the short-time relaxation and τ_{C2} accounts for the long-time relaxation of hydrogen bonds, which denote the lifetimes of the hydrogen bonds. The values of τ_{C1} remain constant throughout the simulation run. The values of τ_{C1} for the proteins 4MT2, 2ABX, 1HQ1, and 1CD3 are calculated as 0.002, 0.004, 0.002, and 0.002 ns, respectively. The average values of τ_{C2} are found to be 0.012, 0.013, 0.015, and 0.018 ns for 4MT2, 2ABX, 1HQ1 and 1CD3, respectively. Low values of τ_{C_2} for both group a and b proteins imply that the intermolecular hydrogen bonds formed between protein and water have short lifetimes. Relaxation of the continuous time autocorrelation function is similar for both groups of proteins which implies a similar rate of decay of the intermolecular hydrogen bonds.

The decay curve of the intramolecular hydrogen bonds may be fitted to a sum of two exponentials given as

$$C_C(t) \propto A_1 \exp(-t/\tau_{C1}) + A_2 \exp(-t/\tau_{C2}) \quad (3)$$

which exhibits a trend similar to that of the intermolecular hydrogen bond continuous autocorrelation function. The values of the time constant for the short-time relaxation, τ_{C1}, are estimated to be 0.006, 0.009, 0.01, and 0.014 ns for 4MT2, 2ABX, 1HQ1, and 1CD3, respectively. However, at longer times, the time constant τ_{C2} for intramolecular hydrogen bonds fluctuates more compared to that of the intermolecular hydrogen bonds.

The average value of τ_{C2} is 0.071 ns for 4MT2, 0.143 ns for 2ABX, 0.083 ns for 1HQ1, and 0.133 ns for 1CD3, respectively. These values of τ_{C2} indicate that intramolecular hydrogen bonds persist more in group b proteins as compared to those in the group a proteins. Comparison of the τ_{C1} and τ_{C2} values for the intermolecular and intramolecular hydrogen bonds for both groups of proteins show that the intermolecular hydrogen bonds decay sharply due to hydrogen bond libration and relax faster due to which they have a shorter lifetime relative to the intramolecular hydrogen bonds.

Parts a−d of Figure 3 depict the intermittent time autocorrelation function, $C_I(t)$, of the intermolecular hydrogen bonds as a function of time. $C_I(t)$ denotes the probability of a particular tagged hydrogen bond to remain intact at time t provided it was intact at t_0. $C_I(t)$ may be fitted to a sum of two exponential functions given by

$$C_I(t) \propto A_1 \exp(-t/\tau_{I1}) + A_2 \exp(-t/\tau_{I2}) \quad (4)$$

where A_1 and A_2 are the time constant amplitude factors and τ_{I1} and τ_{I2} denote the short-time relaxation and long-time relaxation of the hydrogen bond. The values of τ_{I1} for 4MT2, 2ABX, 1HQ1, and 1CD3 are almost similar, i.e., 0.020, 0.026, 0.024, and 0.025 ns, respectively. However, the difference in the dynamics of hydrogen bonds of the proteins is observed at longer times.

For 4MT2, the value of τ_{I2} increases from 0.25 to 0.33 ns after 5 ns of simulation time, followed by a steady decrease.

dx.doi.org/10.1021/jp5013544 | J. Phys. Chem. B 2014, 118, 3018−3025

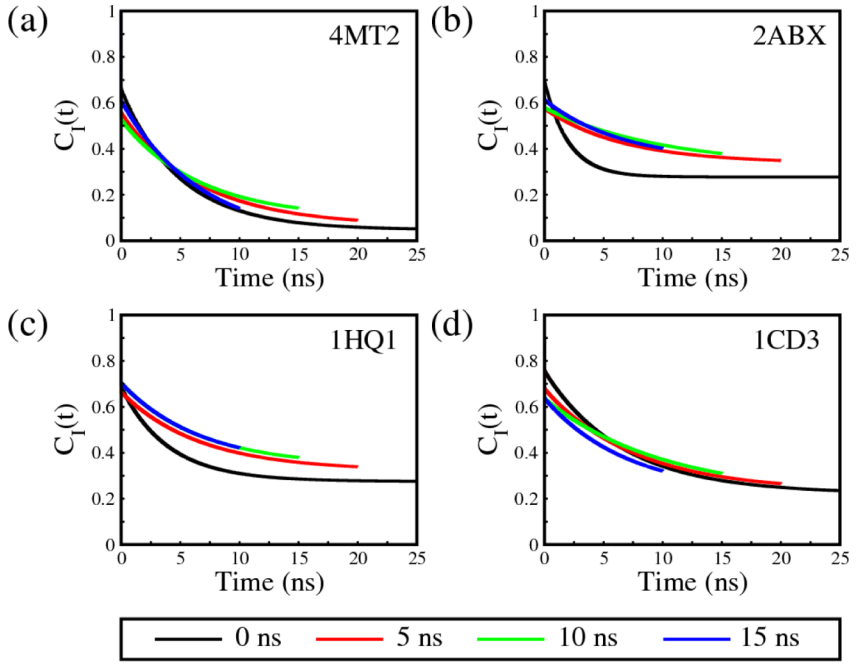


Figure 4. Intermittent autocorrelation function for intramolecular hydrogen bonds in (a) 4MT2, (b) 2ABX, (c) 1HQ1, and (d) 1CD3 as a function of simulation time.

This implies that the rate of hydrogen bond re-formation is highest during the first 5 ns of simulation as compared to the remaining simulation time. For 2ABX the value of τ_{I2} remains constant at 0.4 ns until 15 ns and increases to 6.67 ns after 15 ns, which implies that the rate of re-formation of hydrogen bonds decreases with the increase in simulation time. For the group b proteins, τ_{I2} ranges from 0.91 to 1 ns for 1HQ1 and from 0.667 to 1 ns for 1CD3, respectively. The values of τ_{I2} exhibit a decrease with an increase in the simulation time t, implying that the intermolecular hydrogen bonds present between the protein and water at t_0 are more re-formed compared to the formation of new hydrogen bonds during the simulation. The values of τ_{I1} and τ_{I2} for both groups of proteins indicate that the rate of formation and breakage of the intermolecular hydrogen bonds at intermediate times is similar for both groups of proteins. The re-formation of the intermolecular hydrogen bonds is negligible after 2 ns of simulation time; i.e., they do not persist throughout the entire duration of the simulation. The analysis of intermolecular hydrogen bonds reveals that for group a proteins the hydrogen bonds formed later are less persistent compared to those present at t_0, while for group b proteins the trend is exactly the opposite. The decay of $C_I(t)$ is much slower compared to that of $C_C(t)$, which implies that the intermolecular hydrogen bonds are not continuously present at intermediate times but re-form and break to stabilize the protein throughout the entire simulation.

Parts a–d of Figure 4 depict the intermittent time autocorrelation function of the intramolecular hydrogen bonds as a function of the simulation time. The decay curve for $C_I(t)$ may be fitted to the sum of two exponentials

$$C_I(t) \propto A_1 \exp(-t/\tau_{I1}) + A_2 \exp(-(t/\tau_{I2})^{0.02}) \quad (5)$$

where A_1 and A_2 are the amplitude factors and τ_{I1} and τ_{I2} are the time constants for the short-time and long-time relaxation of $C_I(t)$. The decay of the correlation function $C_I(t)$ is fitted to a stretched exponential function at long times corresponding to the relaxation time constant τ_{I2}. The values of τ_{I2} for all of the proteins is very high; i.e., $\tau_{I2} \approx 10^{50}$ ns as compared to the τ_{I1} values which lie between 2 and 10 ns, indicating that the rate of re-formation is almost similar to the rate of breaking of hydrogen bonds at the longer time regime. Thus the rate of decay of hydrogen bonds approaches zero, making $C_I(t)$ constant in the longer time regime.

For the group a protein 4MT2, the initial value of τ_{I1} is 5 ns which increases to 6.67 ns and remains constant thereafter throughout the entire simulation. This implies a marginal decrease in the initial rate of breaking/re-formation of the hydrogen bonds. For 4MT2, the values of τ_{I2} in the entire duration of simulation reveal that the value of $C_I(t)$ of the hydrogen bonds shows an initial increase to 5 ns after which it remains constant. This indicates that the number of re-formed hydrogen bonds increases as compared to the hydrogen bonds present initially. τ_{I1} for 2ABX increases steadily from 2.5 to 10 ns followed by a decrease to 6.67 ns, while τ_{I2} values increase slightly and remain constant throughout the entire simulation duration. This implies that, at short times, the rate of breaking/re-forming of hydrogen bonds decreases while at longer times the value of $C_I(t)$ becomes constant. The value of τ_{I1} for 1HQ1 increases steadily from 4 to 6.67 ns and attains a constant value after 10 ns, while the value of τ_{I2} increases initially before becoming constant. The initial increase of $C_I(t)$ implies a consequent increase in the rate of breaking/re-forming of the hydrogen bonds at short times, which becomes constant at longer time limits. For 1CD3, τ_{I1} increases from 6.67 to 10 ns followed by a decrease to 7.1 ns, while the τ_{I2} values show a slight increase before attaining a constant value. Thus at short times the rate of breaking/re-forming of hydrogen bonds decreases, while at longer times the value of $C_I(t)$ increases. For both time regimes these rates become almost constant after the first 5 ns of simulation time in the NVE ensemble. This implies that the rate of breakage/re-formation of hydrogen bonds is higher at small times as compared to the longer times, where it approaches zero as the number of hydrogen bonds breaking is

dx.doi.org/10.1021/jp5013544 | J. Phys. Chem. B 2014, 118, 3018–3025

Table 1. Values of Relaxation Times for the Inter- and Intramolecular Hydrogen Bonds Formed by the Group a and b ID Proteins

	continuous (ns)				intermittent (ns)			
	intermolecular		intramolecular		intermolecular		intramolecular	
protein	τ_{C1}	τ_{C2}	τ_{C1}	τ_{C2}	τ_{I1}	τ_{I2}	τ_{I1}	τ_{I2}
4MT2	0.002	0.012	0.006	0.071	0.020	0.25−0.33	5−6.67	10×10^{70}
2ABX	0.004	0.013	0.009	0.143	0.026	0.4−0.67	2.5−10	1.125×10^{65}
1HQ1	0.002	0.015	0.010	0.083	0.024	0.91−1	4−6.67	1.125×10^{65}
1CD3	0.002	0.018	0.014	0.133	0.025	0.67−1	6.67−10	1.39×10^{26}

almost equal to the number of hydrogen bonds re-forming. However, this rate is different for both groups of ID proteins throughout the simulation trajectory. At short-time regime, the rate of decay of hydrogen bonds decreases for all proteins. At longer times, the change in value of $C_I(t)$ implies an increase or decrease in the number of hydrogen bonds forming and breaking at the same rate. The number of such hydrogen bonds is a minimum for 4MT2, which shows a slight increase until 15 ns and decreases thereafter to the initial value. For 2ABX, 1HQ1, and 1CD3 the number of such hydrogen bonds increases as a function of the simulation time. This implies that the new hydrogen bonds formed during simulation have a higher contribution in stabilizing the ID proteins as compared to the hydrogen bonds present initially. The values of relaxation times are summarized in Table 1.

Figure 5 depicts the continuous and intermittent time autocorrelation function of the water−water hydrogen bonds formed by water molecules around the surface of proteins (distance < 5 Å) as a function of simulation time. The decay curve of the water−water hydrogen bonds may be fitted to a sum of two exponentials given as

$$C_{C/I}(t) \propto A_1 \exp(-t/\tau_{C1/I1}) + A_2 \exp(-t/\tau_{C2/I2}) \quad (6)$$

A_1 and A_2 are the amplitude factors and τ_{C1}/τ_{I1} and τ_{C2}/τ_{I2} are the time constants for the short-time and long-time relaxation of $C_C(t)$ and $C_I(t)$, respectively. For these hydrogen bonds, τ_{C1} and τ_{C2} values are calculated to be 0.003 and 0.007 ns, respectively. This implies that at short times the rate of decay of water−water and water−protein hydrogen bonds is almost the same but at longer times the water−water hydrogen bonds decay approximately two times faster compared to the intermolecular hydrogen bonds. The intermittent τ values are 0.008 and 0.083 ns for the short- and long-time regions. Comparison of these values with the intermolecular hydrogen bonds reveal that at short times the rate of decay of water−water hydrogen bonds is almost three times faster than that of the intermolecular hydrogen bonds, while at longer times it is almost 10 times faster. This implies that the intermolecular hydrogen bonds persist more as compared to the water−water hydrogen bonds. For the water−water hydrogen bonds on the surface of the protein, group a and group b proteins exhibit similar values of τ, implying similar hydrogen bond dynamics for both groups of ID proteins at the surface. This depicts that these proteins exhibit similar surface hydrophobicity.

The role of hydrophobicity in the formation of intramolecular hydrogen bonds between hydrophobic (H) and hydrophilic (P) residues is analyzed for the group b protein 1CD3, which contains 48 hydrophobic and 72 hydrophilic residues. Present in 1CD3 are 715 hydrogen atoms, which form

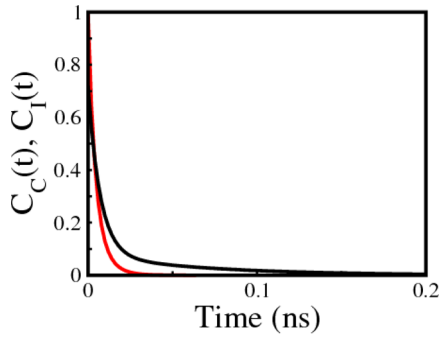


Figure 5. Continuous (red) and intermittent (black) autocorrelation function for water−water hydrogen bonds formed by water molecules around the surface of proteins as a function of simulation time.

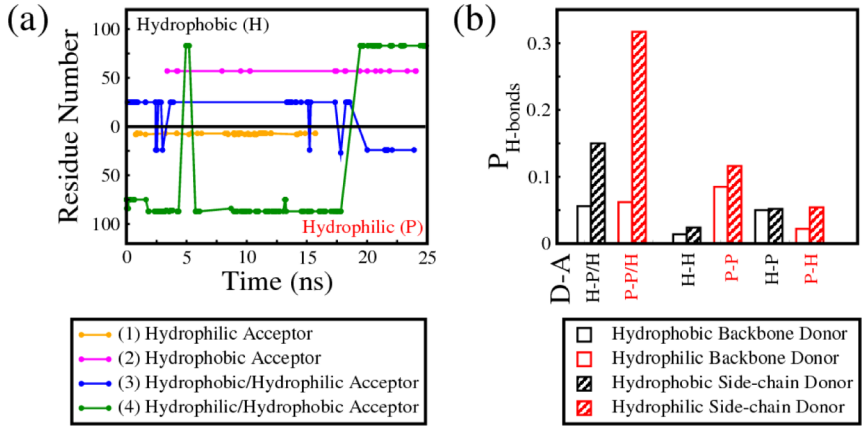


Figure 6. (a) Profile of the partners of hydrogen bonds for 1CD3 according to the acceptor hydrophobicity as function of simulation time. (b) Probability distribution of hydrogen bonds formed for 1CD3 throughout the simulation time of 25 ns.

dx.doi.org/10.1021/jp5013544 | J. Phys. Chem. B 2014, 118, 3018−3025

intramolecular hydrogen bonds during the simulation time, of which 245 are from hydrophobic and 470 are from hydrophilic donors. Hydrophobic donors consist of 85 backbone and 160 side-chain hydrogen atoms, while hydrophilic donors are comprised of 121 backbone and 349 side-chain hydrogen atoms, respectively. Figure 6a depicts the time profile of the hydrogen bonds with respect to a particular acceptor. According to the hydrophobicity/hydrophilicity of the specific donor and acceptor, the hydrogen bonds are classified into four groups H−H, P−P, H−P, and P−H. The hydrogen bond fluctuates between acceptors of similar hydrophobicity. Hydrogen bonds formed in group P−P persist more as compared to the hydrogen bonds of groups H−P and P−H followed by the H−H hydrogen bonds. Higher occurrence of less stable P−P hydrogen bonds result in more interconvertible structures for the native state of 1CD3. Plot 1 of Figure 6a represents two groups of hydrogen bonds H−H and P−H where the donor is hydrophobic and hydrophilic, respectively, but the acceptor is hydrophobic throughout the simulation time. Similarly, plot 2 represents the hydrogen bond groups where the acceptor is hydrophilic but the donor is either hydrophobic or hydrophilic, thus representing two groups H−P and P−P, respectively. If the hydrogen bond formed is fluctuating between hydrophobic and hydrophilic acceptors, then they are represented by separate groups H−P/H and P−P/H according to the donor hydrophobicity. Plots 3 and 4 in Figure 6a represent hydrogen bonds which fluctuate from hydrophobic to hydrophilic acceptors or vice versa. In Figure 6b the probability of hydrogen bonds formed by backbone and side-chain hydrogen atoms are classified for six different groups of hydrogen bonds classified in Figure 6a.

It is evident from Figure 6b that more hydrogen bonds are formed by hydrophilic hydrogen atoms as compared to the hydrophobic hydrogen atoms. The maximum number of hydrogen bonds (58%) formed during the simulation time belong to H−P/H and P−P/H groups. This is exemplified in 1CD3, in which more hydrogen bonds fluctuate between hydrophobic and hydrophobic acceptors. The backbone hydrogen atoms for both hydrophobic and hydrophilic donors form hydrogen bonds with the same probability. The hydrophilic donors form more hydrogen bonds compared to that of the hydrophobic donors for the side-chain hydrogens. 20.1% of hydrogen bonds are in the P−P group, 3.8% of hydrogen bonds belong to H−H group, while 10% are in H−P and 7.6% are in P−H group, respectively. This shows that the maximum hydrogen bonds are formed between hydrophilic donor−acceptor/s, while the least is between hydrophobic donor−acceptor/s for which the hydrophobicity of the acceptor does not fluctuate between hydrophilic/hydrophobic throughout the simulation duration. Hydrophobic hydrogen atoms form more hydrophobic−hydrophilic bonds as compared to the hydrophilic hydrogen atoms. This study implies that hydrogen bonds formed by hydrophilic residues are predominant in ID proteins and the side-chain hydrogen atoms form more hydrogen bonds compared to that of the backbone hydrogen atoms.

IV. CONCLUSION

This study aims to characterize the hydrogen bond dynamics in ID proteins by measuring the different time autocorrelation functions. The relaxation time of the continuous and intermittent autocorrelation functions for the intermolecular hydrogen bonds depict that the decay rates of the hydrogen bonds are almost similar for both group a and group b proteins, implying that both groups of proteins display similar surface hydrophobicity, while the core region of group a proteins may differ from that of group b proteins. Intermolecular hydrogen bonds relax faster compared to that of the intramolecular hydrogen bonds, thus exhibiting shorter lifetimes relative to the intramolecular hydrogen bonds. The analysis of the intermittent autocorrelation function shows that the breakage/reformation of hydrogen bonds occurs throughout the entire duration of the simulation for the intramolecular hydrogen bonds at different rates, while for the intermolecular hydrogen bonds re-formation of hydrogen bond ceases after some time. This implies that the intermolecular hydrogen bonds of ID proteins are very weak as compared to the intramolecular hydrogen bonds. The intermolecular hydrogen bonds are not continuously present at intermediate times but re-form and break to stabilize the protein throughout the simulation. More hydrophilic residues interact in ID proteins to form intramolecular hydrogen bonds as compared to the hydrophobic residues which form less stable hydrogen bonds and help to maintain the dynamic equilibrium among the interconvertible conformations.

AUTHOR INFORMATION

Corresponding Author

*E-mail: pbiswas@chemistry.du.ac.in.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial assistance from Delhi University. We also acknowledge Bioinformatics Resources and Application Facility (BRAF) of Centre for Development of Advanced Computing (CDAC), Pune, India, for providing the computational facility in the Biogene cluster. N.R. acknowledges the financial support from CSIR India in the form of a Senior Research Fellowship.

REFERENCES

(1) Rose, G. D.; Wolfenden, R. Hydrogen Bonding, Hydrophobicity, Packing, and Protein Folding. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 381−415.

(2) Dill, K. A. Dominant Forces in Protein Folding. *Biochemistry* **1990**, *29*, 7133−7155.

(3) Bolen, D. W.; Rose, G. D. Structure and Energetics of the Hydrogen-Bonded Backbone in Protein Folding. *Annu. Rev. Biochem.* **2008**, *77*, 339−362.

(4) Honig, B. Protein folding: From the Levinthal paradox to structure prediction. *J. Mol. Biol.* **1999**, *293*, 283−293.

(5) Pal, S.; Maiti, P. K.; Bagchi, B. Exploring DNA Groove Water Dynamics Through Hydrogen Bond Lifetime and Orientational Relaxation. *J. Chem. Phys.* **2006**, *125*, 234903.

(6) Jana, B.; Pal, S.; Bagchi, B. Hydrogen Bond Breaking Mechanism and Water Reorientational Dynamics in the Hydration Layer of Lysozyme. *J. Phys. Chem. B* **2008**, *112*, 9112−9117.

(7) Tompa, P. Unstructural Biology Coming of Age. *Curr. Opin. Struct. Biol.* **2011**, *21*, 419−425.

(8) Uversky, V. N.; Dunker, A. K. Understanding Protein Non-folding. *Biochim. Biophys. Acta* **2010**, *1804*, 1231−1264.

(9) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. Why are Natively Unfolded Proteins Unstructured under Physiologic Conditions? *Proteins* **2000**, *41*, 415−427.

(10) Müller-Spath, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Rüegger, S.; Reymond, L.; Nettels, D.; Schuler, B. Charge Interactions

Can Dominate the Dimensions of Intrinsically Disordered Proteins. *Proc. Natl. Aacd. Sci. U. S. A.* **2010**, *107*, 14609−14614.
(11) Rawat, N.; Biswas, P. Hydrophobic Moments, Shape, and Packing in Disordered Proteins. *J. Phys. Chem. B* **2012**, *116*, 6326−6335.
(12) Follis, A. V.; Galea, C. A.; Kriwacki, R. W. Intrinsic Protein Flexibility in Regulation of Cell Proliferation: Advantages for Signaling and Opportunities for Novel Therapeutics. *Adv. Exp. Med. Biol.* **2012**, *725*, 27−49.
(13) Thakur, J. K.; Yadav, A.; Yadav, G. Molecular recognition by the KIX domain and its role in gene regulation. *Nucleic Acids Res.* **2014**, *42*, 2112−2125.
(14) Cino, E. A.; Wong-ekkabut, J.; Karttunen, M.; Choy, W.-Y. Microsecond Molecular Dynamics Simulations of Intrinsically Disordered Proteins Involved in the Oxidative Stress Response. *PLoS One* **2011**, *6*, e27371.
(15) Thirumalai, D.; Reddy, G.; Straub, J. E. Role of Water in Protein Aggregation and Amyloid Polymorphism. *Acc. Chem. Res.* **2012**, *45*, 83−92.
(16) Gao, J.; Bosco, D. A.; Powers, E. T.; Kelly, J. W. Localized Thermodynamic Coupling between Hydrogen Bonding and Microenvironment Polarity Substantially Stabilizes Proteins. *Nat. Struct. Mol. Biol.* **2009**, *16*, 684−691.
(17) Gallat, F. X.; Laganowsky, A.; Wood, K.; Gabel, F.; van Eijck, L.; Wuttke, J.; Moulin, M.; Härtlein, M.; Eisenberg, D.; Colletier, J. P.; et al. Dynamical Coupling of Intrinsically Disordered Proteins and Their Hydration Water: Comparison with Folded Soluble and Membrane Proteins. *Nature* **2012**, *103*, 129−136.
(18) Luzar, A.; Chandler, D. Hydrogen-Bond Kinetics in Liquid Water. *Nature* **1996**, *379*, 55−57.
(19) Laage, D.; Hynes, J. T. A Molecular Jump Mechanism of Water Reorientation. *Science* **2006**, *311*, 832−835.
(20) Tay, K. A.; Bresme, F. Kinetics of Hydrogen-Bond Rearrangements in Bulk Water. *Phys. Chem. Chem. Phys.* **2009**, *11*, 409−415.
(21) Xia, Z.; Qiang, Z.; Dong-Xia, Z. Hydrogen Bond Lifetime Definitions and the Relaxation Mechanism in Water Solutions. *Acta Phys.-Chim. Sin.* **2011**, *27*, 2547−2552.
(22) Ptitsyn, O. B. Molten Globule and Protein Folding. *Adv. Protein Chem.* **1995**, *47*, 83−229.
(23) Plaxco, K. W.; Gross, M. Unfolded, yes, but random? Never! *Nat. Struct. Biol.* **2001**, *8*, 659−660.
(24) Tanford, C. Protein Denaturation. *Adv. Protein Chem.* **1968**, *23*, 121−282.
(25) Braun, W.; Vasak, M.; Robbins, A. H.; Stout, C. D.; Wagner, G.; Kagi, J. H. R.; Wüthrich, K. Comparison of the NMR Solution Structure and the X-ray Crystal Structure of Rat Metallothionein-2. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 10124−10128.
(26) Love, R. A.; Stroud, R. M. The Crystal Structure of α-Bungarotoxin at 2.5 Å Resolution: Relation to Solution Structure and Binding to Acetylcholine Receptor. *Protein Eng.* **1986**, *1*, 37−46.
(27) Batey, R. T.; Sagar, M. B.; Doudna, J. A. Structural and Energetic Analysis of RNA Recognition by a Universally Conserved Protein from the Signal Recognition Particle. *J. Mol. Biol.* **2001**, *307*, 229−246.
(28) Dokland, T.; Bernal, R. A.; Burch, A.; Pletnev, S.; Fane, B. A.; Rossmann, M. G. The Role of Scaffolding Proteins in the Assembly of the Small, Single-Stranded DNA Virus φX174. *J. Mol. Biol.* **1999**, *288*, 595−608.
(29) Uversky, V. N.; Dunker, A. K. The Case for Intrinsically Disordered Proteins Playing Contributory Roles in Molecular Recognition without a Stable 3D Structure. *F1000Rep.: Biol.* **2013**, *5*, 1.
(30) Rauscher, S.; Baud, S.; Miao, M.; Keeley, F. W.; Pomes, R. Proline and Glycine Control Protein Self-Organization into Elastomeric or Amyloid Fibrils. *Structure* **2006**, *14*, 1667−1676.
(31) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(32) Sickmeier, M.; Hamilton, J. A.; LeGall, T.; Vacic, V.; Cortese, M. S.; Tantos, A.; Szabo, B.; Tompa, P.; Chen, J.; Uversky, V. N.; et al. DisProt: the Database of Disordered Proteins. *Nucleic Acids Res.* **2007**, *35* (Suppl 1 (database issue)), D786−D793.
(33) Battisti, A.; Tenenbaum, A. Molecular Dynamics Simulation of Intrinsically Disordered Proteins. *Mol. Simul.* **2012**, *38*, 139−143.
(34) *The PyMOL Molecular Graphics System*, Version 1.3r1; Schrödinger LLC: Portland, OR, USA, 2010.
(35) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; et al. *AMBER 12*; University of California: San Fransisco, CA, USA, 2012.
(36) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65*, 712−725.
(37) Sindhikara, D. J.; Kim, S.; Voter, A. F.; Roitberg, A. E. Bad Seeds Sprout Perilous Dynamics: Stochastic Thermostat Induced Trajectory Synchronization in Biomolecules. *J. Chem. Theory Comput.* **2009**, *5*, 1624−1631.
(38) Nicolau-Junior, N.; Giuliatti, S. Modeling and Molecular Dynamics of the Intrinsically Disordered e7 Proteins from High- and Low-Risk Types of Human Papillomavirus. *J. Mol. Model.* **2013**, *19*, 4025−4037.
(39) Rawat, N.; Biswas, P. Size, Shape, and Flexibility of Proteins and DNA. *J. Chem. Phys.* **2009**, *131*, 165104.
(40) Hong, L.; Lei, J. Scaling Law for the Radius of Gyration of Proteins and Its Dependence on Hydrophobicity. *J. Polym. Sci., Part B: Polym. Phys.* **2009**, *47*, 207−214.
(41) Arteca, G. A. Scaling Behavior of Some Molecular Shape Descriptors of Polymer Chains and Protein Backbones. *Phys. Rev. E* **1994**, *49*, 2417−2428.
(42) Zhang, W.; Ganguly, D.; Chen, J. Residual Structures, Conformational Fluctuations, and Electrostatic Interactions in the Synergistic Folding of Two Intrinsically Disordered Proteins. *PLoS Comput. Biol.* **2012**, *8*, e1002353.
(43) Verkhivker, G. M.; Bouzida, D.; Gehlhaar, D. K.; Rejto, P. A.; Freer, S. T.; Rose, P. W. Simulating disorder-order transitions in molecular recognition of unstructured proteins: Where folding meets binding. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 5148−5153.
(44) Stillinger, F. H. Water Revisited. *Science* **1980**, *209*, 451−457.
(45) Luzar, A.; Chandler, D. Hydrogen-Bond Kinetics in Liquid Water. *Nature* **1996**, *379*, 55−57.
(46) Baker, E. N.; Hubbard, R. E. Hydrogen Bonding in Globular Proteins. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97−179.
(47) Levitt, M.; Sharon, R. Accurate Simulation of Protein Dynamics in Solution. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 7557−7561.
(48) Rapaport, D. C. Density Fluctuations and Hydrogen Bonding in Supercooled Water. *Mol. Phys.* **1983**, *48*, 23−31.
(49) Barlow, D. J.; Thornton, J. M. Ion-Pairs in Proteins. *J. Mol. Biol.* **1983**, *168*, 867−885.