JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Structural and Thermodynamic Characteristics That Seed Aggregation of Amyloid-β Protein in Water

Song-Ho Chong, Mirae Park, and Sihyun Ham*

Department of Chemistry, Sookmyung Women's University, Hyochangwon-gil 52, Yongsan-gu, Seoul, 140-742, Korea

ABSTRACT: Amyloid-β (Aβ) proteins undergo conformational transitions leading to aggregation-prone structures, which can initiate self-assembly to form soluble oligomers and eventually insoluble amyloid fibrils when transferred from the transmembrane phase to the physiological aqueous phase. Yet, how Aβ proteins acquire an aggregation-prone nature during the conformational transitions in water remains elusive. Here, we investigate key structural and thermodynamic features of a 42-residue Aβ (Aβ42) protein that seed aggregation based on the fully atomistic, explicit-water molecular dynamics simulations as well as on the integral-equation theory of liquids for solvation thermodynamic analysis. We performed a structure-based analysis on the solvation free energy, a major determinant of the protein hydrophobicity/solubility that influences the aggregation propensity of Aβ42 protein in water. In addition, the Gibbs free energy and its constituents including protein internal energy, protein configurational entropy, solvation enthalpy, and solvation entropy were computed to elucidate thermodynamic driving forces for the conformational transitions of Aβ42 protein in water. On the basis of the atomic-decomposition analysis of these thermodynamic functions, we demonstrate how N-terminal (residues 1−11) and C-terminal (39−42) regions as well as the central region (16−18) contribute significantly to decreasing the solubility of Aβ42 protein upon its conformational transitions in water. These results are consistent with the recent experimental and computational implications and further provide the molecular origin for why the C terminus may serve as an "internal seed" for aggregation and the N-terminal segment may act as a "catalyst" in inducing the Aβ42 self-assembly. This work takes a step forward toward the identification of structural and thermodynamic features of the Aβ42 monomer that seed the aggregation process in water.

INTRODUCTION

A number of human diseases, such as Alzheimer's disease, Parkinson's disease, and type II diabetes, are associated with the deposition of insoluble protein aggregates.[1] Protein aggregation occurs in multistep processes which involve initial conformational changes to form aggregation-prone structures capable of initiating self-assembly, sometimes called protein misfolding.[2] Misfolded proteins can self-assemble into a variety of neurotoxic aggregate species, ranging from soluble oligomers to amyloid fibrils.[3] Understanding the driving factors and mechanistic information of the initial misfolding and subsequent aggregation has been one of the most actively investigated issues in relation to the development of therapeutics and medical applications for the cure of protein aggregation diseases.[4,5] Recently, an increasing number of evidence has accumulated, indicating that soluble oligomers rather than mature amyloid fibrils are the primary toxic agent,[6,7] leading to more focus on early stages of oligomerization involving monomers and small oligomers.[8]

Alzheimer's disease is linked to the aggregation of amyloid-β (Aβ) proteins.[6,9] Deducing the aggregation mechanism of Aβ proteins requires studies on the structures of monomeric and oligomeric species as well as on the kinetics and thermodynamics of their interconversions. However, experimental investigation on the atomistic structures of the monomer and each oligomer species has been hindered by their strong propensity to aggregate in water.[10] Molecular dynamics (MD) simulations have therefore been extensively used to characterize atomic-level structural features of Aβ monomers, the starting point for aggregation. To enable efficient sampling of Aβ conformations, numerous groups have studied Aβ fragments in implicit or explicit solvent and full-length Aβ proteins in implicit solvent.[11−39] To fully take into account the detailed protein−water interactions, explicit-solvent MD simulations on full-length Aβ proteins have also been performed.[40−50]

Despite these numerous computational efforts, key structural features of Aβ monomers that seed aggregation in water remain elusive. In this regard, the knowledge of the protein solvation free energy is an essential ingredient since the change in the hydrophobicity/solubility of a protein is closely correlated with the propensity to form protein aggregates[51,52] and the solvation free energy is a major determinant of the hydrophobicity/solubility.[53,54] The solvation free energy of a protein in water depends on the protein conformation, chemical composition of the protein surface, and the protein−solvent interaction. For example, it has been observed that, when transferred from the organic phase to the aqueous phase, Aβ protein acquires a strong propensity to aggregate in water.[41] This occurs since the protein conformational transitions take place in response to the change in the surrounding physicochemical environment, and the protein surface acquires a more hydrophobic nature, i.e., becomes more prone to aggregate, during those conformational transitions. Therefore, in elucidating the key structural features that seed aggregation, it is crucial to follow how the change in the hydrophobicity/solubility of Aβ protein occurs during the course of the conformational transitions in water. On the other hand, it is also important to note that it is not the solvation free

Received: October 26, 2011
Published: January 7, 2012

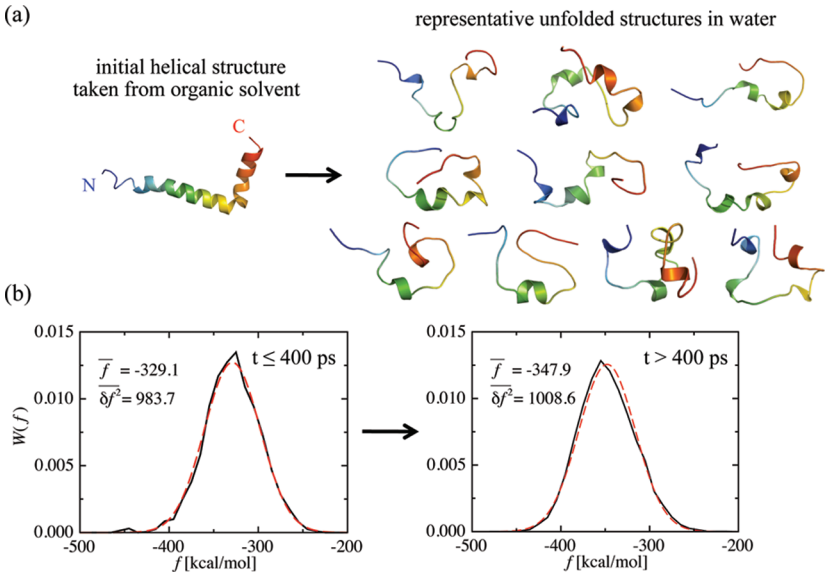

Figure 1. (a) NMR structure of the Aβ42 monomer determined in an 80:20 hexafluoroisopropanol/water mixture at neutral pH (PDB ID: 1IYT)[57] (left) and representative aqueous-phase Aβ42 monomer structures determined in ref 46 from the clustering analysis of each of 10 independent simulation trajectories (right). Each structure is color-coded according to sequence, ranging from blue to red at the N and C termini, respectively. (b) Probability distribution function $W(f)$ (solid curves) of the solvent-averaged protein potential energy f of Aβ42 in water for the short-time regime ($t \leq 400$ ps) related to the initial helical configurations representing organic phase structure (left panel) and the one for the long-time regime ($t > 400$ ps) for the aqueous-phase configurations (right panel). Dashed curves denote the fit by the Gaussian function. Numerical values of $\bar{f}$ and $\overline{\delta f^2}$ in units of kcal/mol and (kcal/mol)^2 are reported in each panel.

energy, but the Gibbs free energy, that determines whether such conformational transitions occur spontaneously. Therefore, a detailed thermodynamic analysis in terms of the Gibbs free energy and its constituents—protein internal energy, protein configurational entropy, and solvation free energy along with solvation enthalpy and solvation entropy components—is necessary to probe driving forces for protein misfolding and aggregation.

In this paper, we report a detailed thermodynamic investigation on the full-length, 42-residue form of Aβ protein (Aβ42) in water based on explicit-water MD simulations as well as on the solvation thermodynamic analysis using the integral-equation theory of liquids.[55,56] The aim of this work is to address key structural and thermodynamic features of the Aβ42 monomer that seed the subsequent oligomerization process in water. We used the helical structure determined by NMR in a membrane-mimicking organic solvent[57] as the initial structure and carried out fully atomistic MD simulations of the Aβ42 monomer with explicit water. This closely resembles the *in vivo* situation that the Aβ42 protein is produced through the endoproteolysis of the transmembrane amyloid precursor protein and then is released to an extracellular aqueous environment where the Aβ42 aggregation takes place.[9] The relevance of the aqueous-phase structures of the Aβ42 monomer generated from our MD simulations is validated through the comparison of the collision cross-section measured by ion-mobility mass spectrometry.[58] Protein internal energy was directly computed during the simulations, while the integral-equation theory for molecular systems[55,56] was used to yield solvation thermodynamic quantities. These results were combined to obtain protein configurational entropy,[59] allowing one to determine the Gibbs free energy change and all of its constituents upon the Aβ42 conformational transitions in water. The atomic decomposition analysis of the solvation free energy[60] was further conducted to elucidate how Aβ42 protein acquires an aggregation-prone nature during the conformational transitions and to reveal the regions of the Aβ42 protein that significantly affect its aggregation propensity in water.

RESULTS

Aqueous-Phase Aβ42 Monomer Structures from Simulations Are in Accord with Experimental Data. We performed MD simulations to study the conformational transitions of Aβ42 monomer in water starting from the helical structure determined in a membrane-mimicking organic solvent [see Figure 1a].[46] In total, 10 independent simulations of 50 ns were performed at a temperature T = 300 K and a pressure P = 1 bar under neutral pH (see Methods). Representative aqueous-phase Aβ42 monomer structures are displayed in Figure 1a, which were determined from the clustering analysis of each of the 10 independent simulated trajectories. The initial event of conformational transitions for Aβ42 simulation in water is found to be the unfolding of the C-terminal region,[46] which is in accord with the experimental observation that the C-terminal unfolding was detected when Aβ42 was exposed to more water from the organic solvent.[41] Our simulated structure also agrees with the NMR dynamic study in that the C terminus of Aβ42 monomer in water is rather structured and buried.[61] The J-coupling constants from our simulations are also found to be in agreement with the experimental NMR data.[42,46]

For further corroboration, we compared the collision cross-section of Aβ42 monomer structures from our simulations with that determined from the ion-mobility mass spectrometry.[15,58] In order to compare with the solvent-free experiments on samples electrosprayed from the solution phase, each protein structure from our simulations was instantaneously dehydrated through energy minimization in a vacuum. The collision cross-section of the dehydrated structure was then calculated using the trajectory method implemented in the MOBCAL software.[62] We obtained 738 ± 25 Å^2 from the 10

representative structures in water shown in Figure 1a. The average value from our simulations is about 5% larger than the experimental value 700 Å^2.[15,58] The agreement within ~5%, which is slightly better than ~8% agreement (760 Å^2) obtained from the replica-exchange MD simulations,[15,58] is satisfactory considering the size and complexity of the Aβ42 protein. (See the discussion in ref 58 on possible sources of disagreement from the experimental value.)

Constituents of Gibbs Free Energy Uncover the Thermodynamic Driving Forces for the Conformational Transitions of Aβ42 Protein in Water. Gibbs free energy G and its constituents—protein internal energy E_u, solvation free energy $\Delta\mu$ along with solvation enthalpy (Δh) and solvation entropy (Δs) components, protein configurational entropy S_{conf}, and pressure-volume term PV_u with protein partial molar volume V_u—were calculated by applying the integral-equation theory to protein structures generated from the MD simulations (see Methods). To discuss the Gibbs free energy change upon the Aβ42 conformational transitions in water, resembling the *in vivo* conformational transitions of Aβ42 from a helical membrane-mimicking-phase structure to extracellular aqueous-phase structure, we divided the time regimes of the protein configurations in the simulation trajectories as follows. To represent the initial helical structures of Aβ42 in water, we took configurations in the short-time regime $t \leq 400$ ps from each of the 10 simulation trajectories with a 2 ps time interval. To represent aqueous-phase structures, we took configurations in the longer time regime $t > 400$ ps with a 50 ps time interval from each of the 10 trajectories. In total, 2000 configurations were chosen to represent the initial helical structure and 9920 configurations were taken for the aqueous-phase structure. These two sets of protein configurations were used to construct the probability distribution functions $W(f)$ of the solvent-averaged protein potential energy[59] $f = E_u + \Delta\mu$ shown in Figure 1b. Our choice of the two time regimes was made on the basis of the observation that the system explored different portions of the solvent-averaged potential energy surface for $t \leq 400$ ps and $t > 400$ ps, as can be inferred from Figure 1b: the distribution function $W(f)$ for $t \leq 400$ ps is centered around $\overline{f} = -329.1$ kcal/mol, whereas for $t > 400$ ps the system explores deeper portions of $W(f)$ centered around $\overline{f} = -347.9$ kcal/mol, and both of the distribution functions can be well approximated by the Gaussian function.

Now, we can discuss the Gibbs free energy change $\Delta G = \Delta\overline{E_u} + \Delta\overline{\Delta\mu} - T\Delta S_{conf} + P\Delta\overline{V_u}$, where the bar represents the average over the set of protein configurations used in defining $W(f)$, and ΔX ($X = \overline{E_u}$, $\overline{\Delta\mu}$, S_{conf}, or $\overline{V_u}$) denotes the difference between the two sets of protein configurations. We obtained $\Delta G = -39.6$ kcal/mol consisting of $\Delta\overline{E_u} = -91.0$ kcal/mol, $\Delta\overline{\Delta\mu} = +72.2$ kcal/mol, $T\Delta S_{conf} = +20.9$ kcal/mol, and $P\Delta\overline{V_u} = 1.24 \times 10^{-4}$ kcal/mol (Table 1). The $P\Delta\overline{V_u}$ term shall not be considered from here on since its magnitude is negligibly small. The solvation free energy change $\Delta\overline{\Delta\mu}$ comprises the solvation enthalpy change $\Delta\overline{\Delta h} = +81.5$ kcal/mol and the solvation entropy change $T\Delta\overline{\Delta s} = +9.3$ kcal/mol. Grouping the enthalpic and entropic contributions yields $\Delta G = \Delta H - T\Delta S$ with $\Delta H = \Delta\overline{E_u} + \Delta\overline{\Delta h} = -9.5$ kcal/mol and $T\Delta S = T\Delta S_{conf} + T\Delta\overline{\Delta s} = +30.1$ kcal/mol. Thus, the conformational transition of the Aβ42 protein in water starting from a helical structure is a *spontaneous* process driven mainly by *entropic* factors. On the other hand, this is accompanied by a large unfavorable *increase* in the solvation free energy ($\Delta\overline{\Delta\mu}$) dominated by the solvation enthalpy component ($\Delta\overline{\Delta h}$), indicating that the A$\beta$42 protein becomes much more hydrophobic and, hence, prone to aggregate upon the conformational transitions in water.

Table 1. Gibbs Free Energy Change and Its Constituents in kcal/mol

	protein	solvation	total
enthalpy	$\Delta\overline{E_u} = -91.0$	$\Delta\overline{\Delta h} = +81.5$	−9.5
entropy	$T\Delta S_{conf} = +20.9$	$T\Delta\overline{\Delta s} = +9.3$	+30.1
free energy			−39.6

Protein Internal Energy Decreases Due to the Electrostatic Hydrophilic Side-Chain Interactions and the Formation of the Nonlocal Backbone Contacts. To understand how the thermodynamic changes are associated with the changes in protein conformation and hydration structure in a time sequence, we shall study a representative trajectory from ref 46 in more detail. As displayed in Figure 2a–c, the thermodynamic functions (ΔE_u, $\Delta\Delta\mu$, $\Delta\Delta s$) in this trajectory exhibit the same trend as the averaged ones ($\Delta\overline{E_u}$, $\Delta\overline{\Delta\mu}$, $\Delta\overline{\Delta s}$); i.e., the protein internal energy decreases, the solvation free energy increases, and the solvation entropy increases upon the conformational transitions in water from the initial helical structure. Here, $\Delta X = X - X_{ini}$ ($X = E_u$, $\Delta\mu$, or Δs) for single trajectory denotes the change from the value X_{ini} for the initial helical structure.

We start from analyzing how the protein internal energy change ΔE_u is associated with the protein conformational changes. For this purpose, we divided the sequence of Aβ42 into four regions: regions I (residues 1–15), II (16–24), III (25–31), and IV (32–42) (see the inset of Figure 3). As displayed in Figure 3, the components in ΔE_u that exhibit the most drastic change are the one from region I and the interaction energy between regions II and IV. These changes in the protein internal energy reflect the structural evolution of Aβ42 protein in the representative trajectory where three distinct time regimes have been found.[46] Conformational features in the first time regime located at 4–18 ns are the disruption of the initial helical structure and the formation of a hydrophilic loop (residues 25–29) [see structure A in Figure 2d]. The large energetic change in region I reflects considerable conformational fluctuations as the helical structure is disrupted in the N-terminal region. Since region I mainly consists of hydrophilic residues, conformational fluctuations such as the formation of salt-bridges among nearby hydrophilic side chains, which becomes possible due to the release of helical backbone restraints, lead to significant energy variations. In the second time regime (18–43 ns), nonlocal backbone contacts are starting to be developed between the C-terminal end and the central region (residues 16–18). In particular, a strong electrostatic interaction is formed between the NH_3^+ group of Lys16 and the COO^- group of the C-terminal end [structure B in Figure 2d], which is reflected in the sudden drop around 18 ns of the interaction energy between regions II and IV. In the third time regime (after 43 ns), the two regions between the C-terminal (residues 39–42) end and the central region (16–18) are distinctively close to each other, and nonlocal backbone H bonds between these regions are formed [structure C in Figure 2d], leading to the further decrease in the interaction energy between regions II and IV after 43 ns. Another characteristic feature in the third time regime

dx.doi.org/10.1021/ct200757a | J. Chem. TheoryComput. 2012, 8, 724–734

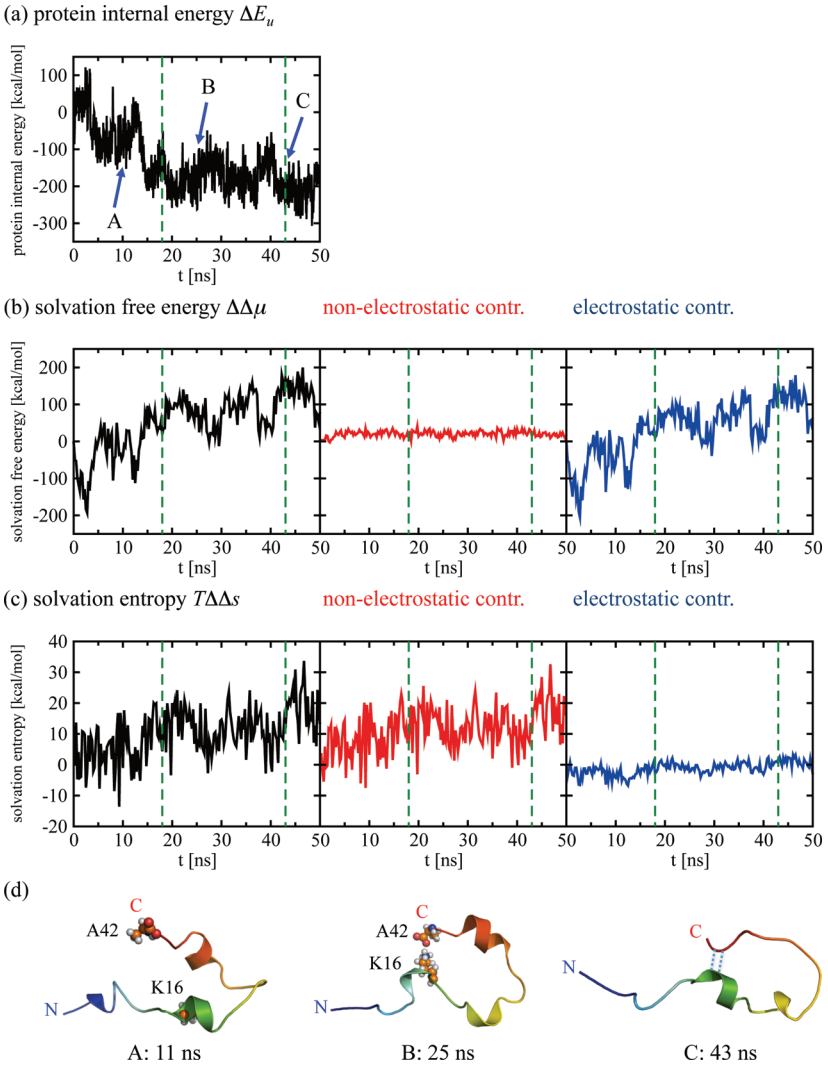

Figure 2. (a) Protein internal energy change ΔE_u as a function of time. In this and the following panels, vertical dashed lines refer to 18 and 43 ns separating the three time regimes discussed in the text. The arrows labeled A to C indicate the times 11, 25, and 43 ns, respectively. (b) Solvation free energy change $\Delta\Delta\mu$ (black curve) as a function of time and its decomposition into nonelectrostatic (red) and electrostatic (blue) contributions. (c) Solvation entropy change $T\Delta\Delta s$ (black curve) as a function of time and its decomposition into nonelectrostatic (red) and electrostatic (blue) contributions. (d) Unfolded structures in water at 11, 25, and 43 ns from the representative MD simulation trajectory. Each structure is color-coded according to sequence, ranging from blue to red at the N and C termini, respectively. Side chains of the residues K16 and A42 as well as the COO^- group of the C-terminal end in the 11 and 25 ns structures are drawn with sphere representation to indicate the positions of these residues (C, N, O, H: orange, blue, red, white). The nonlocal backbone contacts are shown with dotted lines in the 43 ns structure.

is the exposure of hydrophobic side chains,[46] which, however, does not contribute significantly to the protein internal energy change.

Aβ42 Protein Is Dehydrated upon the Conformational Transitions in Water Leading to an Increase in the Solvation Entropy. The conformational transitions of the Aβ42 protein also alter the hydration structure, through which changes in the solvation thermodynamic properties are determined. Let us first note a contrasting behavior between the solvation free energy change $\Delta\Delta\mu$ and the solvation entropy change $\Delta\Delta s$, which is based on their decomposition into nonelectrostatic and electrostatic contributions (see Methods) presented in Figure 2b and c. (Here, we do not separately display the solvation enthalpy $\Delta\Delta h$ since the change in the solvation free energy is dominated by the solvation enthalpy component, as can be inferred from the large ordinate-scale difference in Figure 2b and c, and hence, the change in the solvation free energy is almost identical to that of the solvation enthalpy.) While the change in the solvation free energy is dominated by the electrostatic contribution, the nonelectrostatic contribution is predominant in the solvation entropy change. The former feature simply reflects the fact that the electrostatic interactions provide larger energetic contributions than the nonelectrostatic (i.e., van der Waals) interactions. The latter observation can be understood as follows. When a protein is immersed in water, a volume from which water molecules are excluded is generated, and this leads to the loss in the translational entropy of water. If the orientational motion of water becomes restricted due to a strong protein−water electrostatic interaction, the orientational entropy is reduced. Thus, the water translational and orientational entropies are reflected in the nonelectrostatic and electrostatic components in the solvation entropy, respectively. The dominance of the nonelectrostatic contribution indicates that it is the translational entropy of water that governs the solvation

dx.doi.org/10.1021/ct200757a | J. Chem. TheoryComput. 2012, 8, 724−734

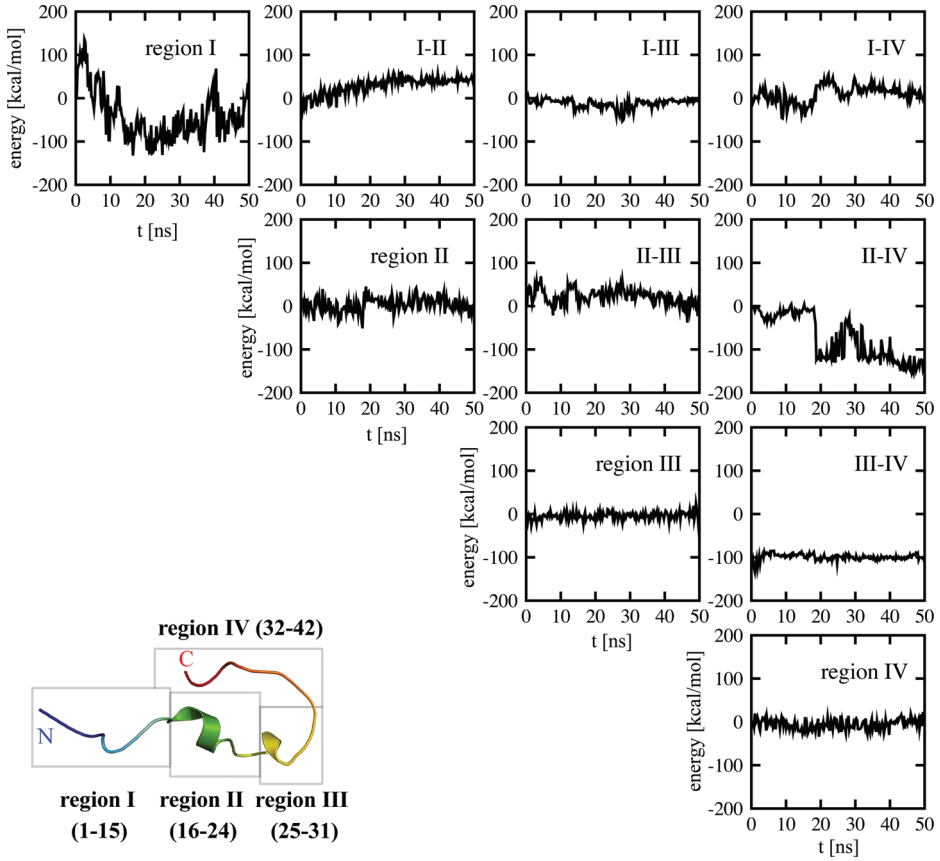

Figure 3. Decomposition of the protein internal energy change ΔE_u as a function of time based on the four regions I–IV of the protein sequence as defined in the inset. Diagonal panels refer to the protein internal energy within a region and off-diagonal ones to the interaction energy between the two regions.

entropy change upon the conformational transitions of Aβ42 protein in water.

The changes in the solvation thermodynamic quantities $\Delta\Delta\mu$ and $\Delta\Delta s$ can be traced back to the hydration structure changes, which can be quantified by the radial distribution function $g_{\alpha\gamma}(r)$ between the protein atom α and the water atom γ. The salt-bridge formation due to the disruption of the helical structure in the N-terminal region results in the dehydration of those hydrophilic side chains involved, as exemplified in Figure 4a for Arg5. The formation of the above-mentioned strong electrostatic interaction between the NH_3^+ group of Lys16 and the COO^- group of the C-terminal end also leads to the dehydration of these charged groups, as demonstrated in Figure 4a. Upon the formation of the nonlocal backbone H bonds between the C-terminal end and the central region (residues 16–18), the backbones in these residues become inaccessible to water, as shown in Figure 4b. The exposure of the hydrophobic side chains is also reflected in the radial distribution function, as exemplified in Figure 4c for Phe20.

The dehydration of hydrophilic side chains and polar backbones leads to a positive change in $T\Delta\Delta s$. This is because, when water molecules bound to the hydrophilic side chains and backbones via hydrogen-bonding are released to the bulk upon dehydration, those water molecules gain translational entropy and because $T\Delta\Delta s$ is dominated by the translational entropy as we discussed in connection with Figure 2c. On the other hand, the exposure of the hydrophobic side chains would yield an opposite contribution to $T\Delta\Delta s$. The overall increase in $T\Delta\Delta s$ shown in Figure 2c indicates that the dehydration dominates the solvation entropy change upon the conformational transitions of Aβ42 protein in water.

Atomic-Decomposition Analysis of the Solvation Free Energy Reveals the Regions in Aβ42 Protein That Significantly Affect Its Aggregation Propensity in Water. To understand the molecular origin for the solvation free energy increase upon the conformational transitions of the Aβ42 protein in water, we examined how each residue contributes to $\Delta\Delta\mu$ on the basis of the atomic decomposition analysis (see Methods). Contributions from the selected residues Lys16, Val40, and Phe20 to the solvation free energy change $\Delta\Delta\mu$ are presented in Figure 5. Also added there is the time dependence of the peak heights of the backbone and side chain radial distribution functions shown in Figure 4. The dehydration of the charged NH_3^+ group in the side chain of Lys16, which occurs starting from 18 ns, leads to a large loss in the peptide–water interaction energy and, hence, to a significant positive change in $\Delta\Delta\mu$ around that time. The backbone dehydration of Lys16 and Val40, i.e., the replacement of backbone–water H-bonding with backbone–backbone H-bonding that occurs near 43 ns, also leads to a positive change in $\Delta\Delta\mu$ around that time. These dehydration phenomena basically explain the changes in $\Delta\Delta\mu$ for Lys16 and Val40. On the other hand, only a tiny positive change in $\Delta\Delta\mu$ is observed near 40 ns when the side chain of Phe20 is exposed to water. Thus, the solvation free energy increase is caused by (i) hydrophilic side-chain dehydration, (ii) polar backbone dehydration, and (iii) hydrophobic side-chain exposure, when ordered according to the significance of the contribution.

dx.doi.org/10.1021/ct200757a | J. Chem. TheoryComput. 2012, 8, 724–734

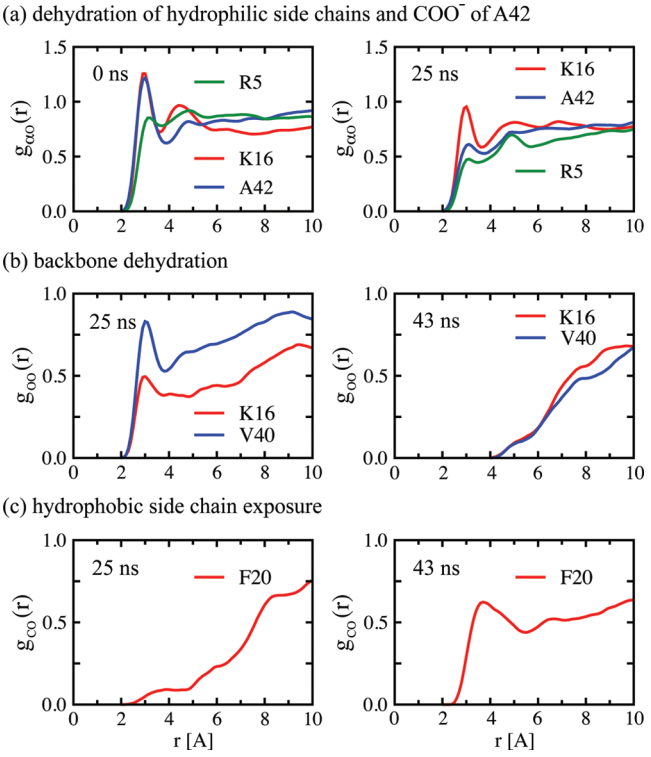

Figure 4. (a) Radial distribution functions $g_{\alpha O}(r)$ between the atom α in the protein and the water oxygen for the residues R5 (green curves), K16 (red curves), and A42 (blue curves) of the initial organic-phase helical structure and the unfolded structure in water of Aβ42 at 25 ns. For R5, nitrogen in the guanidinium group is chosen for the atom α, and $g_{\alpha O}(r)$ is obtained by averaging over three nitrogens. Nitrogen in the NH_3^+ group is chosen for the atom α for K16. Oxygen in the COO^- group is chosen for the atom α for the C-terminal A42. (b) Radial distribution functions $g_{OO}(r)$ between the main-chain carbonyl oxygen and the water oxygen for K16 (red curves) and V40 (blue curves) of the unfolded structures in water at 25 and 43 ns. (c) Radial distribution functions $g_{CO}(r)$ between the side-chain carbon atom and the water oxygen for the residue F20 (red curves) of the unfolded structures in water at 25 and 43 ns.

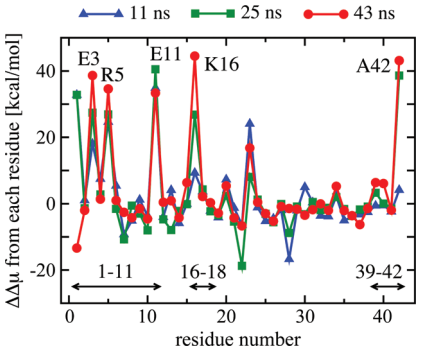

Figure 6. Each residue's contribution to the solvation free energy change $\Delta\Delta\mu$ between the initial organic-phase helical structure and the unfolded structures in water at 11 ns (blue triangles), 25 ns (green squares), and 43 ns (red circles). Large contributions to $\Delta\Delta\mu$ come from hydrophilic as well as terminal residues E3, R5, E11, K16, and A42.

By performing the atomic-decomposition analysis for all of the residues constituting the Aβ42 protein, we show in Figure 6 how each residue contributes to the solvation free energy change $\Delta\Delta\mu$ upon the conformational transitions of Aβ42 protein in water. It is seen that large contributions to $\Delta\Delta\mu$ come from hydrophilic residues in the N-terminal region as well as from those residues that participate in the nonlocal backbone contacts between the C-terminal end and the central region (residues 16−18), demonstrating the importance of charged and polar groups over nonpolar ones in determining the protein solvation free energy.[63] Since the increase in the solvation free energy has a direct consequence on the decrease in the solubility as well as the increase in the hydrophobicity, Figure 6 elucidates that the N-terminal (residues 1−11) and C-terminal (39−42) regions as well as the central region (16−18) are the regions that significantly affect the aggregation propensity of Aβ42 protein in water.

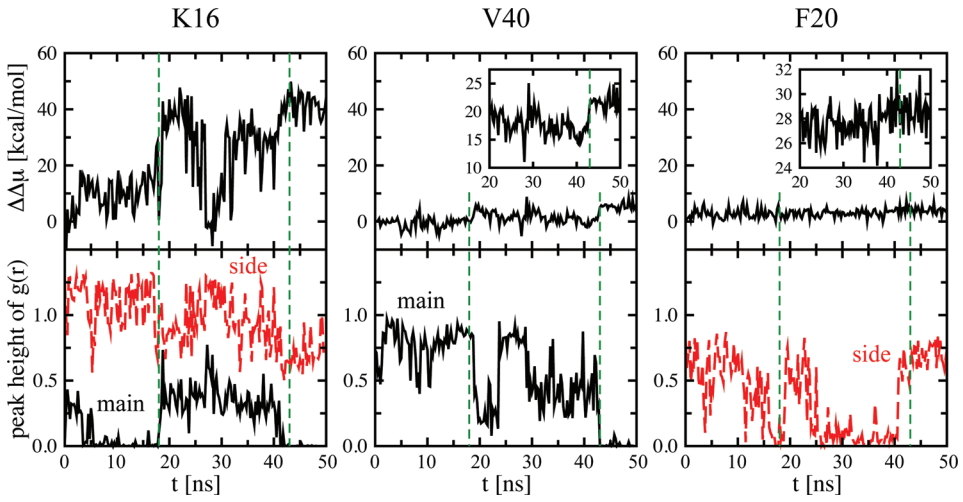

Figure 5. (Top panels) Contribution from the residues K16, V40, and F20 specified in each panel to the solvation free energy change $\Delta\Delta\mu$ as a function of time. Results for V40 and F20 are enlarged in the insets for clarity. Vertical dashed lines refer to 18 and 43 ns. (Bottom panels) Solid curves denote the peak height of the radial distribution function between the main-chain carbonyl oxygen and the water oxygen [cf. Figure 4b] as a function of time. Dashed curves refer to the corresponding results between the side-chain atom and the water oxygen [cf. Figure 4a and c]: the side-chain nitrogen is chosen for K16, while the side-chain carbon atom is chosen for F20.

dx.doi.org/10.1021/ct200757a | J. Chem. TheoryComput. 2012, 8, 724−734

DISCUSSION

The conformational transitions of the Aβ monomer that occur in water are crucial for its subsequent aggregation since, in this process, the Aβ protein acquires an aggregation-prone structure. The detailed atomic-level descriptions on the aqueous-phase Aβ monomer structure derived from experimental methods have been limited due to the strong tendency of Aβ proteins to aggregate in water. In this regard, MD simulation studies coupled with NMR experiments[42,48,50] and ion-mobility mass spectrometry measurements[15,58] are breaking new ground to provide structural characterizations of the Aβ monomers in water. The agreement between the experiments and our simulations on the *J*-coupling constant as well as on the collision cross-section suggests that our simulations yield aqueous-phase Aβ42 monomer structures relevant to the experiment.

By further performing the structure-based solvation free energy analysis, we here address key structural and thermodynamic features of the Aβ42 monomer that may seed the subsequent oligomerization process. It has been suggested from the comparative NMR dynamic study on Aβ42 and Aβ40 monomers differing at the C-terminal ends that the C terminus of Aβ42 may serve as an internal seed for aggregation.[61,64,65] This conjecture is based on the observation that Aβ42 has a more rigid C-terminal tail than the Aβ40 monomer. Computational studies on the Aβ42 protein in water confirmed the existence of the structured C-terminus in this protein.[42,46] Here, we provide direct evidence on how the rigid C-terminus is associated with the aggregation propensity of Aβ42 protein in water. The rigidity of the C terminus is attributed to the formation of nonlocal backbone contacts involving the C-terminal (residues 39−42) and central (16−18) residues, and the formation of nonlocal backbone contacts and the concomitant dehydration is shown to have a direct consequence on the increase in the solvation free energy. Furthermore, the hydrophobic side chains in the C-terminal and central regions are exposed to the surface of the protein due to the formation of nonlocal backbone H-bonding.[46] On the basis of these structural and solvation thermodynamic characterizations, the C-terminal region (residues 39−42) as well as the central region (16−18) are the "aggregation-prone" regions with exposed hydrophobic side chains that significantly affect the solubility/hydrophobicity, and hence, the aggregation propensity of the Aβ42 protein in water (Figure 7).

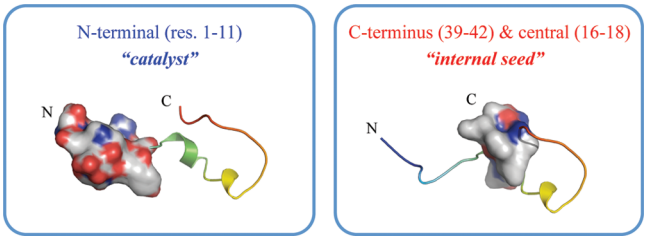

Figure 7. Surface-representation of the regions in the Aβ42 protein that significantly affect its aggregation propensity in water. The negatively charged part is colored in red, the positively charged part in blue, and the rest in gray.

Recently, a possible critical role of the N-terminal segment in Aβ aggregation has been suggested on the basis of the experimental observation that an antibody or ligand that specifically binds with the N-terminal region (mainly residues 1−10) influences the propensity of Aβ proteins to form aggregates.[66,67] MD simulation studies on the wild-type Aβ42 and two of its mutants also demonstrated that the rigidity of the N-terminal region is in good correlation with the experimentally observed aggregation-rate change upon mutation.[49] Furthermore, it has been demonstrated from an extensive coarse-grained simulation study that the N-terminal region plays an important role in Aβ_{1-28} aggregation.[39] Since the N-terminal region is not a part of the fibril core,[68] it was conjectured that the N-terminal segment might act as a "catalyst" of Aβ aggregation.[49] We here demonstrate that the N-terminal region (residues 1−11), mainly consisting of hydrophilic residues, plays an important role in determining the solvation free energy, i.e. the solubility of Aβ42 protein in water, which is consistent with the previous experimental and computational implications. In the initial helical structure determined in the organic phase, the side chains of these hydrophilic residues are exposed to water due to the helical backbone restraints. When the helical structure in the N-terminal region of Aβ42 is disrupted upon its unfolding in water, the side chains of those hydrophilic residues become dehydrated by constructing the electrostatic interactions with nearby hydrophilic side chains as well as polar backbone atoms, affecting the solubility of the Aβ42 protein in water. The polar and flexible segment of the N-terminal region that significantly contributes to determining the intrinsic nature of solubility of the Aβ42 protein in water may thus act as a "catalyst" in inducing the subsequent Aβ42 self-assembly (Figure 7). Thus, our work takes a step forward toward the identification of structural and thermodynamic features of the Aβ monomer that may seed the subsequent aggregation process in water.

The increase in the solvation free energy upon the conformational transitions of the Aβ42 monomer in water is thermodynamically an unfavorable process. We find that such an unfavorable increase in the solvation free energy is surmounted by favorable changes in the other thermodynamic quantities, and overall the conformational transition of the Aβ42 protein in water is a spontaneous process which is driven mainly by entropic factors (Table 1). The unfavorable increase in the solvation free energy ($\Delta\overline{\Delta\mu}$), dominated by its enthalpy component ($\Delta\overline{\Delta h}$), is largely compensated by the favorable decrease in the protein internal energy ($\Delta\overline{E_u}$). Such a large cancellation basically reflects the competition of the protein−water H-bond formation with the intraprotein H-bond and salt-bridge formation. As a result, the thermodynamic driving force for the Aβ42 conformational transitions in water is largely entropic in origin, arising from the increase in the protein configurational entropy and in the solvation entropy. The increase in the protein configurational entropy reflects the presence of a more heterogeneous distribution of conformational ensembles of the Aβ42 monomer in the aqueous phase as compared to the helical organic-phase structure. The solvation entropy increase is shown to be almost exclusively ascribed to the translational entropy gain of water upon dehydration.

The thermodynamic investigation on the Aβ42 monomer presented here is also a prerequisite to discussing that of a subsequent aggregation process since the free energy of dimerization $G_d - 2G_m$ is expressed in terms of the Gibbs free energies for dimer (G_d) and monomer (G_m) species and since G_m enters similarly into the free energy of monomer addition to oligomers. Our theoretical work on the protein configurational entropy may also serve to promote an understanding of another possible role of the rigid C terminus of the

Aβ42 protein in explaining the stark difference in biological behavior between Aβ42 and Aβ40. This structural difference of the C terminus would also result in the difference in the protein configurational entropy, and such an entropic difference between the Aβ42 and Aβ40 monomers might account for why Aβ42 aggregates faster and with higher toxicity than Aβ40.[61] In this regard, performing a similar thermodynamic analysis on Aβ40 and comparing results with those for Aβ42 will be illuminating.

CONCLUSION

We report a detailed thermodynamic investigation on the Aβ42 protein in water based on the explicit-water MD simulations as well as on the integral-equation theory of liquids. We perform structure-based thermodynamic analysis on the solvation free energy and on the Gibbs free energy and its constituents—protein internal energy, protein configurational entropy, solvation enthalpy, and solvation entropy—to probe key structural and thermodynamic features that seed aggregation as well as thermodynamic driving forces for the conformational transitions of the Aβ42 protein in water. We show that the conformational transition of the Aβ42 protein in water starting from the organic-phase structure is a spontaneous process mainly driven by entropic factors, which accompanies a large increase in the solvation free energy (decrease in solubility), thus affecting the aggregation propensity of the Aβ42 protein in water. By further performing the atomic-decomposition analysis of the solvation free energy, we elucidate how Aβ42 proteins acquire an aggregation-prone nature through the conformational transitions in water. We identify the C terminus (residues 39−42) and the central region (16−18) as an "internal seed" and the N-terminal (1−11) region as a "catalytic" region that significantly influence the intrinsic nature of the hydrophobicity/solubility of Aβ42 in water relevant for initiating subsequent hydrophobicity-induced self-assembly. Our work will contribute to understanding driving factors and mechanical information of the misfolding and aggregation of Aβ proteins which are crucial in developing therapeutics for the cure of Alzheimer's disease.

METHODS

MD Simulations. The initial structure for the wild-type Aβ42 with a sequence of $^{(1)}$DAEFR HDSGY EVHHQ KLVFF AEDVG SNKGA IIGLM VGGVV IA$^{(42)}$ was selected from the NMR determined structure in organic solvent at neutral pH (PDB ID: 1IYT).[57] Under neutral pH, Lys and Arg residues are positively charged, and Glu and Asp are negatively charged. The total net charge of the Aβ42 monomer is −3. MD simulations for the Aβ42 monomer in explicit water, reported in ref 46, were performed with the SANDER module of the AMBER9 program package[69] using the ff99 force field.[70] The starting Aβ42 monomer was solvated in a rectangular box containing 5500 TIP3P waters molecules,[71] where the distance to the edge of the box from Aβ42 was chosen to be 10 Å. Three Na^+ counterions were added to neutralize the system. Each simulation trajectory was run for 50 ns with a 2.0 fs time step at the temperature T = 300 K and the pressure P = 1 bar under neutral pH. Temperature and pressure were controlled with a Berendsen's thermostat and barostat with coupling constants of 1.0 and 2.0 ps, respectively.[72] In total, 10 independent simulations with different initial velocities were performed. These MD simulation trajectories were used to generate the time evolution of the protein structure in water as well as the protein internal energy E_w which can be directly computed from the force field used in the simulations.

Solvation Thermodynamics Based on the Integral-Equation Theory. For each protein configuration generated by the MD simulations, we applied the three-dimensional reference interaction site model (3D-RISM) theory to calculate solvation thermodynamic properties. The 3D-RISM theory is an integral equation theory based on statistical mechanics for obtaining the 3D distribution function $g_\gamma(\mathbf{r})$ of the site γ, oxygen or hydrogen, of water at position $\mathbf{r}$ around a molecular solute such as protein.[55,56] For a solute−solvent system at infinite dilution, the 3D-RISM equation is given by

$$h_\gamma(\mathbf{r}) = \sum_{\gamma'} c'(\mathbf{r}) * [w^{vv}_{\gamma'\gamma}(r) + \rho h^{vv}_{\gamma'\gamma}(r)]$$

Here, $h_\gamma(\mathbf{r})$ and $c_\gamma(\mathbf{r})$ refer to the 3D total and direct correlation functions of the water site γ, respectively, the asterisk denotes a convolution integral, $w^{vv}_{\gamma'\gamma}(r)$ and $h^{vv}_{\gamma'\gamma}(r)$ are the site−site intramolecular and total correlation functions of the water, respectively, and ρ represents the average number density of water. This equation is to be supplemented by an approximate closure relation, and in the present study we adopted the one suggested by Kovalenko and co-workers:[55,56]

$$h_\gamma(\mathbf{r}) = \begin{cases} \exp[d_\gamma(\mathbf{r})] - 1 & \text{for} \quad d_\gamma(\mathbf{r}) \le 0 \\ d_\gamma(\mathbf{r}) & \text{for} \quad d_\gamma(\mathbf{r}) > 0 \end{cases}$$

in which $d_\gamma(\mathbf{r}) = -u_\gamma(\mathbf{r})/(k_B T) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})$ with k_B denoting Boltzmann's constant. $u_\gamma(\mathbf{r})$ refers to the interaction potential acting on the water site γ which is generated by atoms in the protein and is represented by a sum of Lennard-Jones (LJ) and Coulomb electrostatic terms which are centered at the position r_α of protein atom α, $u_\gamma(\mathbf{r}) = \sum_\alpha [u^{(LJ)}_{\alpha\gamma}(|\mathbf{r} - \mathbf{r}_\alpha|) + u^{(elec)}_{\alpha\gamma}(|\mathbf{r} - \mathbf{r}_\alpha|)]$. Here, $u^{(LJ)}_{\alpha\gamma}(r) = 4\varepsilon_{\alpha\gamma}[(\sigma_{\alpha\gamma}/r)^{12} - (\sigma_{\alpha\gamma}/r)^6]$ and $u^{(elec)}_{\alpha\gamma}(r) = q_\alpha q_\gamma / r$ with $\varepsilon_{\alpha\gamma}$, $\sigma_{\alpha\gamma}$, q_α, and q_γ being the LJ parameters and atomic charges, respectively.

The 3D-RISM calculation for $g_\gamma(r)$ around the Aβ42 protein was performed as follows. For each protein configuration generated by MD simulations, one can determine the interaction potential $u_\gamma(\mathbf{r})$. On the basis of the knowledge of $u_\gamma(\mathbf{r})$, the two unknown functions $h_\gamma(\mathbf{r})$ and $c_\gamma(\mathbf{r})$ can be determined by solving the 3D-RISM equation and the closure relation self-consistently, and the 3D water distribution function is obtained via $g_\gamma(r) = h_\gamma(\mathbf{r}) + 1$. This calculation was performed on a grid of 128 × 128 × 128 points in a cubic supercell with a side length of 64 Å. We used the dielectrically consistent RISM theory[73] for the site−site correlation functions $w^{vv}_{\gamma'\gamma}(r)$ and $h^{vv}_{\gamma'\gamma}(r)$ of water determined under the same thermodynamic conditions (T = 300 K and ρ = 1 g/cm^3) as the MD simulations and with the dielectric constant of 78.4. Technical details concerning the 3D-RISM calculation can be found in ref 55.

The radial distribution function $g_{\alpha\gamma}(r)$ between the atom α in protein and the atom γ in water rather than the 3D distribution function $g_\gamma(\mathbf{r})$ is sometimes more convenient in quantifying the water distribution around a specific atom α. The former can be obtained from the latter by shifting the origin of the coordinate to the position r_α followed by the orientational average, i.e., $g_{\alpha\gamma}(r) = (1/4\pi) \int d\hat{\mathbf{r}}\, g_\gamma(\mathbf{r}_\alpha + \mathbf{r})$ with $\hat{\mathbf{r}} = \mathbf{r}/r$ and $r = |\mathbf{r}|$.

Thermodynamic functions of solvation can be obtained on the basis of the water distribution function. For the solvation

free energy $\Delta\mu$, the following analytical expression is available under the use of the Kovalenko−Hirata closure given above:[55,56]

$$\Delta\mu = \rho k_B T \sum_\gamma \int d\mathbf{r}\left[\frac{1}{2}h_\gamma(\mathbf{r})^2\,\Theta(-h_\gamma(\mathbf{r})) - c_\gamma(\mathbf{r}) - \frac{1}{2}h_\gamma(\mathbf{r})\,c_\gamma(\mathbf{r})\right]$$

Here, $\Theta(x)$ is the Heaviside step function.

While the solvation free energy does not depend on whether the solute insertion is done under the isochoric (constant volume) or isobaric (constant pressure) condition, its energetic and entropic components depend on such a condition.[74] Under the isochoric condition, the solvation free energy $\Delta\mu$ comprises the solvation energy $\Delta\varepsilon_V$ and isochoric solvation entropy Δs_V. The latter is given by the temperature derivative of $\Delta\mu$ at constant density:

$$\Delta s_V = -\left(\frac{\partial\Delta\mu}{\partial T}\right)_\rho$$

In the present work, the temperature derivative was calculated numerically using the first order finite difference with ΔT = 2 K. Under the isobaric condition, which is more relevant to the present study since the MD simulations were performed at constant pressure, the solvation free energy is decomposed into the solvation enthalpy Δh and isobaric solvation entropy Δs. The relations between the quantities under isochoric and isobaric conditions are given by[74]

$$\Delta h = \Delta\varepsilon_V + \frac{T\alpha_P}{\kappa_T}V_u,\; T\Delta s = T\Delta s_V + \frac{T\alpha_P}{\kappa_T}V_u$$

in terms of the isobaric thermal expansion coefficient α_P and the isothermal compressibility κ_T of the solvent and the partial molar volume V_u of the solute. Using the experimental values for α_P and κ_T for water at T = 300 K and P = 1 bar, the right-hand side of this equation can be estimated as $0.041 \times V_u$ kcal/mol when V_u is measured in cm^3/mol. The partial molar volume can be obtained in terms of the 3D direct correlation function via $V_u = k_B T\kappa_T\{1 - \rho\sum_\gamma \int d\mathbf{r}\, c_\gamma(\mathbf{r})\}$.[75]

The main limitation of the 3D-RISM theory lies in the use of an approximate closure relation, which is inherent in all of the integral-equation theories. In particular, the absolute value of the solvation free energy depends on the closure relation used.[55,56] However, it is known that relative values of the solvation free energies are reasonably accurate.[55] We note in this connection that only the relative values of the thermodynamic functions matter in the present study since we are interested in the changes in thermodynamic properties upon the conformational transitions of protein. It is therefore expected that our results in the main text do not significantly suffer from the limitation of the integral-equation theory.

Decomposition Method of the Solvation Thermodynamic Functions. The nonelectrostatic and electrostatic contributions to the solvation free energy $\Delta\mu$ and solvation entropy Δs, plotted in Figure 2b and c, can be obtained as follows. First, we calculate $\Delta\mu$ and Δs with the full protein−water interaction. Next, we repeat this calculation with the electrostatic protein−water interaction turned off, $u_{\alpha\gamma}^{(\mathrm{elec})}(r) = 0$, which yields the nonelectrostatic contributions $\Delta\mu^{(\mathrm{LJ})}$ and $\Delta s^{(\mathrm{LJ})}$. The electrostatic contributions can then be obtained by subtraction, $\Delta\mu^{(\mathrm{elec})} = \Delta\mu - \Delta\mu^{(\mathrm{LJ})}$ and $\Delta s^{(\mathrm{elec})} = \Delta s - \Delta s^{(\mathrm{LJ})}$.

Further decomposition of the solvation thermodynamic quantities into atomic contributions can be carried out using the exact partitioning method developed in ref 60. The atomic-decomposition of the solvation free energy $\Delta\mu$ into contribution $\Delta\mu_\alpha$ from atom α in protein is given by[60]

$$\Delta\mu = \sum_\alpha \Delta\mu_\alpha \text{ with } \Delta\mu_\alpha = \Delta\mu_\alpha^{(\mathrm{LJ})} + \Delta\mu_\alpha^{(\mathrm{elec})}$$

in which

$$\Delta\mu_\alpha^{(\mathrm{LJ})} = 4\pi\rho\sum_\gamma \int_0^1 d\lambda_1 \int r^2\, dr\, \frac{\partial u_{\alpha\gamma}^{(\mathrm{LJ})}(r;\lambda_1)}{\partial\lambda_1} \times g_{\alpha\gamma}(r;\lambda_1,\lambda_2 = 0)$$

$$\Delta\mu_\alpha^{(\mathrm{elec})} = 4\pi\rho\sum_\gamma \int_0^1 d\lambda_2 \int r^2\, dr\, \frac{\partial u_{\alpha\gamma}^{(\mathrm{elec})}(r;\lambda_2)}{\partial\lambda_2} \times g_{\alpha\gamma}(r;\lambda_1 = 1,\lambda_2)$$

Here, λ_1 and λ_2 are the parameters for scaling the LJ parameter ($\lambda_1\sigma_{\alpha\gamma}$) and the atomic charge ($\lambda_2 q_\alpha$) of the protein, respectively, and the resulting interaction potentials are denoted as $u_{\alpha\gamma}^{(\mathrm{LJ})}(r;\lambda_1)$ and $u_{\alpha\gamma}^{(\mathrm{elec})}(r;\lambda_2)$. $g_{\alpha\gamma}(r;\lambda_1,\lambda_2)$ refers to the radial distribution function when the protein−water interaction potential is given by $u_\gamma(\mathbf{r};\lambda_1,\lambda_2) = \sum_\alpha[u_{\alpha\gamma}^{(\mathrm{LJ})}(|\mathbf{r} - \mathbf{r}_\alpha|;\lambda_1) + u_{\alpha\gamma}^{(\mathrm{elec})}(|\mathbf{r} - \mathbf{r}_\alpha|;\lambda_2)]$, which can also be obtained from the 3D-RISM calculation with the aforementioned procedure.

Protein Configurational Entropy and Gibbs Free Energy. The protein internal energy E_u from the MD simulation and the solvation free energy $\Delta\mu$ from the integral-equation theory were combined to define the solvent-averaged protein potential energy $f = E_u + \Delta\mu$.[59] For a given set of protein conformations, one can construct the probability distribution function $W(f)$ for f. When $W(f)$ is well approximated by Gaussian, which is the case for the system studied here [see Figure 1b], the Helmholtz free energy F is given by $F = \overline{E_u} + \overline{\Delta\mu} - TS_{\mathrm{conf}}$ where the protein configurational entropy is given by $TS_{\mathrm{conf}} = \overline{\delta f^2}/(2k_BT)$ with $\delta f = f - \overline{f}$.[59] Here, $\overline{X}$ denotes an average over the set of protein conformations used in defining $W(f)$. The Gibbs free energy G can then be obtained by adding the pressure−volume term to F. A major advantage of our method for the protein configurational entropy over the conventional quasi-harmonic approximation[76−78] is that it does not assume a Gaussian distribution of protein *conformations*, and it can be applied to unfolded proteins such as Aβ42 in which an average protein structure is not well-defined.

AUTHOR INFORMATION

Corresponding Author

*Phone: +82-2-710-9410. Fax: +82-2-2077-7321. E-mail: sihyun@sookmyung.ac.kr.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (KRF-2008-313-C00404 and No. 20090065791). The authors would like to acknowledge the support from KISTI supercomputing center.

■ REFERENCES

(1) Ross, C. A.; Poirier, M. A. *Nat. Med.* **2004**, *10*, S10−S17.
(2) Dobson, C. M. *Nature* **2003**, *426*, 884−890.
(3) Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2006**, *75*, 333−366.
(4) Bemporad, F.; Calloni, G.; Campioni, S.; Plakoutsi, G.; Taddei, N.; Chiti, F. *Acc. Chem. Res.* **2006**, *39*, 620−627.
(5) Muñoz, V. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 395−412.
(6) Kayed, R.; Head, E.; Thompson, J. L.; McIntire, T. M.; Milton, S. C.; Cotman, C. W.; Glabe, C. G. *Science* **2003**, *300*, 486−489.
(7) Walsh, D. M.; Selkoe, D. J. *J. Neurochem.* **2007**, *101*, 1172−1183.
(8) Bernstein, S. L.; Dupuis, N. F.; Lazo, N. D.; Wyttenbach, T.; Condron, M. M.; Bitan, G.; Teplow, D. B.; Shea, J.-E.; Ruotolo, B. T.; Robinson, C. V.; Bowers, M. T. *Nat. Chem.* **2009**, *1*, 326−331.
(9) Hardy, J. A.; Higgins, G. A. *Science* **1992**, *256*, 184−185.
(10) Teplow, D. B. *Methods Enzymol.* **2006**, *413*, 20−33.
(11) Massi, F.; Peng, J. W.; Lee, J. P.; Straub, J. E. *Biophys. J.* **2001**, *80*, 31−44.
(12) Fawzi, N. L.; Phillips, A. H.; Ruscio, J. Z.; Doucleff, M.; Wemmer, D. E.; Head-Gordon, T. *J. Am. Chem. Soc.* **2008**, *130*, 6145−6158.
(13) Borreguero, J. M.; Urbanc, B.; Lazo, N. D.; Buldyrev, S. V.; Teplow, D. B.; Stanley, H. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6015−6020.
(14) Lam, A.; Teplow, D.; Stanley, H. E.; Urbanc, B. *J. Am. Chem. Soc.* **2008**, *130*, 17413−17422.
(15) Baumketner, A.; Bernstein, S. L.; Wyttenbach, T.; Bitan, G.; Teplow, D. B.; Bowers, M. T.; Shea, J.-E. *Protein Sci.* **2006**, *15*, 420−428.
(16) Khandogin, J.; Brooks, C. L. III. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 16880−16885.
(17) Li, W.; Zhang, J.; Su, Y.; Wang, J.; Qin, M.; Wang, W. *J. Phys. Chem. B* **2007**, *111*, 13814−13821.
(18) Yang, M.; Teplow, D. B. *J. Mol. Biol.* **2008**, *384*, 450−464.
(19) Vitalis, A.; Caflisch, A. *J. Mol. Biol.* **2010**, *403*, 148−165.
(20) Cruz, L.; Urbanc, B.; Borreguero, J. M.; Lazo, N. D.; Teplow, D. B.; Stanley, H. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 18258−18263.
(21) Baumketner, A.; Bernstein, S. L.; Wyttenbach, T.; Lazo, N. D.; Teplow, D. B.; Bowers, M. T.; Shea, J.-E. *Protein Sci.* **2006**, *15*, 1239−1247.
(22) Krone, M. G.; Baumketner, A.; Bernstein, S. L.; Wyttenbach, T.; Lazo, N. D.; Teplow, D. B.; Bowers, M. T.; Shea, J.-E. *J. Mol. Biol.* **2008**, *381*, 221−228.
(23) Wei, G.; Shea, J.-E. *Biophys. J.* **2006**, *91*, 1638−1647.
(24) Ito, M.; Johansson, J.; Stromberg, R.; Nilsson, L. *PLoS One* **2011**, *6*, e17587.
(25) Baumketner, A.; Krone, M. G.; Shea, J.-E. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6027−6032.
(26) Daidone, I.; Simona, F.; Roccatano, D.; Broglia, R. A.; Tiana, G.; Colombo, G.; Nola, A. D. *Proteins* **2004**, *57*, 198−204.
(27) Baumketner, A.; Shea, J.-E. *J. Mol. Biol.* **2006**, *362*, 567−579.
(28) Kamiya, N.; Mitomo, D.; Shea, J.-E.; Higo, J. *J. Phys. Chem. B* **2007**, *111*, 5351−5356.
(29) Ikebe, J.; Kamiya, N.; Ito, J. I.; Shindo, H.; Higo, J. *Protein Sci.* **2007**, *16*, 1596−1608.
(30) Massi, F.; Straub, J. E. *Biophys. J.* **2001**, *81*, 697−709.
(31) Massi, F.; Klimov, D.; Thirumalai, D.; Straub, J. E. *Protein Sci.* **2002**, *11*, 1639−1647.
(32) Massi, F.; Straub, J. E. *J. Comput. Chem.* **2003**, *24*, 143−153.
(33) Han, W.; Wu, Y.-D. *J. Am. Chem. Soc.* **2005**, *127*, 15408−15416.
(34) Tarus, B.; Straub, J. E.; Thirumalai, D. *J. Am. Chem. Soc.* **2006**, *128*, 16159−16168.
(35) Baumketner, A.; Shea, J.-E. *J. Mol. Biol.* **2007**, *366*, 275−285.
(36) Reddy, G.; Straub, J. E.; Thirumalai, D. *J. Phys. Chem. B* **2009**, *113*, 1162−1172.
(37) Wu, C.; Murray, M. M.; Bernstein, S. L.; Condron, M. M.; Bitan, G.; Shea, J.-E.; Bowers, M. T. *J. Mol. Biol.* **2009**, *387*, 492−501.
(38) Rojas, A.; Liwo, A.; Browne, D.; Scheraga, H. A. *J. Mol. Biol.* **2010**, *404*, 537−552.
(39) Rojas, A. V.; Liwo, A.; Scheraga, H. A. *J. Phys. Chem. B* **2011**, *115*, 12978−12983.
(40) Flöck, D.; Colacino, S.; Colombo, B.; Nola, A. D. *Proteins* **2006**, *62*, 183−192.
(41) Tomaselli, S.; Esposito, V.; Vangone, P.; van Nuland, N. A. J.; Bonvin, A. M. J. J.; Guerrini, R.; Tancredi, T.; Temussi, P. A.; Picone, D. *ChemBioChem* **2006**, *7*, 257−267.
(42) Sgourakis, N. G.; Yan, Y. L.; McCallum, S. A.; Wang, C. Y.; Garcia, A. E. *J. Mol. Biol.* **2007**, *368*, 1448−1457.
(43) Raffa, D. F.; Rauk, A. *J. Phys. Chem. B* **2007**, *111*, 3789−3799.
(44) Triguero, L.; Singh, R.; Prabhakar, R. *J. Phys. Chem. B* **2008**, *112*, 2159−2167.
(45) Yang, C.; Li, J.; Li, Y.; Zhu, X. *THEOCHEM* **2009**, *895*, 1−8.
(46) Lee, C.; Ham, S. *J. Comput. Chem.* **2011**, *32*, 349−355.
(47) Wood, G. P. F; Rothlisberger, U. *J. Chem. Theory Comput.* **2011**, *7*, 1552−1563.
(48) Sgourakis, N. G.; Merced-Serrano, M.; Boutsidis, C.; Drineas, P.; Du, Z.; Wang, C.; Garcia, A. E. *J. Mol. Biol.* **2011**, *405*, 570−583.
(49) Velez-Vega, C.; Escobedo, F. A. *J. Phys. Chem. B* **2011**, *115*, 4900−4910.
(50) Ball, K. A.; Phillips, A. H.; Nerenberg, P. S.; Fawzi, N. L.; Wemmer, D. E.; Head-Gordon, T. *Biochemistry* **2011**, *50*, 7612−7628.
(51) Chiti, F.; Taddei, N.; Baroni, F.; Capanni, C.; Stefani, M.; Ramponi, G.; Dobson, C. M. *Nat. Struct. Biol.* **2002**, *9*, 137−143.
(52) Wurth, C.; Guimard, N. K.; Hecht, M. H. *J. Mol. Biol.* **2002**, *319*, 1279−1290.
(53) Ben-Naim, A. *Hydrophobic Interactions*; Plenum Press: New York, 1980.
(54) Wang, H.; Ben-Naim, A. *J. Phys. Chem. B* **1997**, *101*, 1077−1086.
(55) Kovalenko, A. In *Molecular Theory of Solvation*; Hirata, F., Ed; Kluwer: Dordrecht, The Netherlands, 1980; p 169.
(56) Imai, T.; Harano, Y.; Kinoshita, M.; Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2006**, *125*, 024911.
(57) Crescenzi, O.; Tomaselli, S.; Guerrini, R.; Salvadori, S.; D'Ursi, A. M.; Temussi, P. A.; Picone, D. *Eur. J. Biochem.* **2002**, *269*, 5642−5648.
(58) Bernstein, S. L.; Wyttenbach, T.; Baumketner, A.; Shea, J.-E.; Bitan, G.; Teplow, D. B.; Bower, M. T. *J. Am. Chem. Soc.* **2005**, *127*, 2075−2084.
(59) Chong, S.-H.; Ham, S. *Chem. Phys. Lett.* **2011**, *504*, 225−229.
(60) Chong, S.-H.; Ham, S. *J. Chem. Phys.* **2011**, *135*, 034506.
(61) Yan, Y.; Wang, C. *J. Mol. Biol.* **2006**, *364*, 853−862.
(62) Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. *J. Phys. Chem.* **1996**, *100*, 16082−16086.
(63) Chong, S.-H.; Lee, C.; Kang, G.; Park, M.; Ham, S. *J. Am. Chem. Soc.* **2011**, *133*, 7075−7083.
(64) Morimoto, A.; Irie, K.; Murakami, K.; Masuda, Y.; Ohigashi, H.; Nagao, M.; Fukuda, H.; Shimizu, T.; Shirasawa, T. *J. Biol. Chem.* **2004**, *279*, 52781−52788.
(65) Lim, K. H.; Henderson, G. L.; Jha., A.; Louhivuori, M. *ChemBioChem* **2007**, *8*, 1251−1254.
(66) Lee, M.; Bard, F.; Johnson-Wood, K.; Lee, C.; Hu, K.; Griffith, S. B.; Black, R. S.; Schenk, D.; Seubert, P. *Ann. Neurol.* **2005**, *58*, 430−435.
(67) Orner, B. P.; Liu, L.; Murphy, R. M.; Kiessling, L. L. *J. Am. Chem. Soc.* **2006**, *128*, 11882−11889.
(68) Lührs, T.; Ritter, C.; Adrian, M.; Riek-Loher, D.; Bohrmann, B.; Dobeli, H.; Schubert, D.; Riek, R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17342−17347.

(69) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. *AMBER 9*; University of California: San Francisco, CA, 2006.
(70) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(71) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(72) Berendsen, H. J. C.; Postma, J. P. M; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(73) Perkyns, J.; Pettitt, B. M. *J. Chem. Phys.* **1992**, *97*, 7656−7666.
(74) Cann, N. M.; Patey, G. N. *J. Chem. Phys.* **1997**, *106*, 8165−8195.
(75) Harano, Y.; Imai, T.; Kovalenko, A.; Kinoshita, M.; Hirata, F. *J. Chem. Phys.* **2011**, *114*, 9506−9511.
(76) Karplus, M.; Kushick, J. N. *Macromolecules* **1981**, *14*, 325−332.
(77) Schlitter, J. *Chem. Phys. Lett.* **1993**, *215*, 617−621.
(78) Schäfer, H.; Mark, A. E.; van Gunsteren, W. F. *J. Chem. Phys.* **2000**, *113*, 7809−7817.