THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

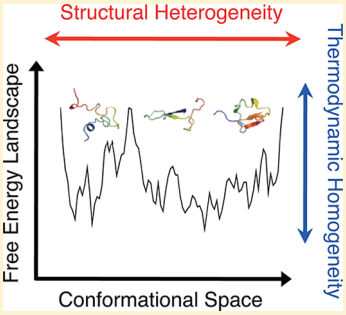
Conformational Entropy of Intrinsically Disordered Protein

Song-Ho Chong and Sihyun Ham*

Department of Chemistry, Sookmyung Women's University, Cheongpa-ro 47-gil 100, Yongsan-Ku, Seoul 140-742, Korea

Supporting Information

ABSTRACT: Intrinsically disordered proteins (IDPs), though lacking stable tertiary structures, are known to possess a certain amount of residual structure. Conformational disorder plays a crucial role through the conformational entropy in regulating protein−protein and protein−ligand interactions involved in signaling and regulation, and also modulates protein aggregation and amyloidogenesis associated with a number of human diseases. However, a direct and quantitative connection between the residual structure and the conformational entropy remains to be established. Here we show using a novel computational approach that the conformational entropy of amyloid-beta protein, an IDP whose aggregation is associated with Alzheimer's disease, is significantly correlated with the contents of the residual helical structure, β-sheet structure, and salt-bridge network. Identification of the thermodynamically significant residual structure is of fundamental importance for a comprehensive understanding of the relationship between the functional conformational disorder and the protein activity regulation, and will also serve the thermodynamic basis of the amyloid polymorphism.



INTRODUCTION

It is now widely acknowledged that a wide range of proteins involved in gene regulation and signal transduction are intrinsically disordered or comprise inherently unstructured regions.[1−3] Intrinsically disordered proteins (IDPs) are also associated with a number of human diseases such as cancer, neurodegenerative disorders, and diabetes.[4,5] It has been quite a challenge to detect and characterize conformational disorders in IDPs, since the secondary structure and long-range contacts in IDPs are inherently transient and comprise only a small fraction of the total protein sequence. The past several years, however, have witnessed significant improvements in sensitivity and resolution of spectroscopic methods, and the application of those advanced experimental techniques has revealed that IDPs are far from homogeneous statistical random coil polymers and instead exhibit a rich variety of local and even long-range residual structure.[6,7]

Intrinsic disorder is a unique feature that enables IDPs to bind with multiple partners.[8−10] While the binding affinity with different partners obeys general thermodynamic principles, a potentially important role of the conformational entropy is expected to play in determining the binding free energy of IDPs. Historically, the protein conformational entropy has remained one of the most elusive thermodynamic parameters to estimate. However, a recent NMR analysis method that employs conformational dynamics as a proxy for conformational entropy is breaking new ground.[11] In fact, this method has not only uncovered the significance of the conformational entropy in the protein−ligand binding free energy but also revealed how the binding affinity with different partners is tuned through the changes in the conformational entropy.[11,12] However, a quantitative connection between the protein structural order/disorder and the conformational entropy remains elusive. Clarifying those structural factors that influence the protein conformational entropy will open the possibility of a direct interpretation of the conformational entropy in structural terms.

Here, we investigate a quantitative connection between the residual structure and the conformational entropy of amyloid-beta (Aβ) protein, an IDP whose aggregation is linked to Alzheimer's disease,[13] by using extensive explicit-water molecular dynamics (MD) simulations and the molecular theory of solvation. The full-length, 42-residue form of the wild-type Aβ protein (wt-Aβ42) as well as its pathogenic familial mutants[14−16] such as Flemish (A21G), Arctic (E22G), Dutch (E22Q), Italian (E22K), and Iowa (D23N) were studied. We also analyzed synthetic mutants termed RR (I41R, A42R) and DQ (I41D, A42Q) whose aggregation propensity was investigated *in vivo*.[17] MD simulations were used for the characterization of heterogeneous conformational ensembles of these IDPs. On the other hand, the estimation of the protein conformational entropy from simulations still remains a considerable challenge.[18] The applicability of the well-adopted quasi-harmonic approximation[19−21] is questionable, since it assumes a Gaussian distribution of protein conformations about an average structure, which does not hold for IDPs lacking stable three-dimensional structures. In this respect, we recently developed an energy-based approach,[22] implemented by combining MD simulations with the liquid integral-equation theory, that does not require the Gaussian distribution of protein conformations and thus is applicable to IDPs. Through the characterization of the conformational entropy in terms of residual structural properties,

Received: January 30, 2013
Revised: March 25, 2013
Published: March 26, 2013

   dx.doi.org/10.1021/jp401049h | J. Phys. Chem. B 2013, 117, 5503−5509

Table 1. Structural and Thermodynamics Characteristics of the Wild-Type (wt) Aβ42 and Its Mutants Studied in the Present Work

system	helix[a]	β-sheet[b]	SB[c]	SB network[d]	ΔΔμ[e]	TS_{conf}[f]	agg.[g]
wt-Aβ42	4.9 ± 6.1	9.1 ± 7.1	5.7 ± 3.8	0.6 ± 2.1	0.0	382.3	
Flemish (A21G)	6.8 ± 7.2	4.1 ± 5.7	4.7 ± 3.4	0.5 ± 1.8	−5.5	404.2	↓
Arctic (E22G)	4.8 ± 5.7	7.7 ± 6.3	5.5 ± 3.7	1.4 ± 3.2	142.0	372.0	↑
Dutch (E22Q)	6.9 ± 6.3	3.2 ± 3.3	4.4 ± 3.2	0.7 ± 2.1	116.1	392.7	↑
Italian (E22K)	5.0 ± 7.3	7.8 ± 6.9	6.0 ± 4.2	0.5 ± 1.8	164.7	376.2	↑
Iowa (D23N)	6.5 ± 6.1	9.3 ± 8.3	5.2 ± 3.4	1.3 ± 2.8	112.1	361.3	↑
RR (I41R, A42R)	4.7 ± 7.4	5.3 ± 6.8	7.4 ± 4.3	0.9 ± 2.5	−14.1	405.6	↓
DQ (I41D, A42Q)	8.0 ± 8.4	2.9 ± 3.5	5.6 ± 3.9	1.1 ± 2.7	−166.8	376.6	↓

[a]Average population (%) and root-mean-squared fluctuation of the helical structure. [b]Average population (%) and root-mean-squared fluctuation of the β-sheet structure. [c]Average population (%) and root-mean-squared fluctuation of the salt-bridge (SB). [d]Average population (%) and root-mean-squared fluctuation of the salt-bridge (SB) network. [e]Change in the solvation free energy upon mutation, $\Delta\Delta\mu = \Delta\mu(\text{mutant}) - \Delta\mu(\text{wt})$, in kcal/mol. [f]Protein conformational entropy S_{conf} multiplied by the temperature T in kcal/mol. [g]Experimental aggregation propensity change (decrease, ↓; increase, ↑) upon mutation.

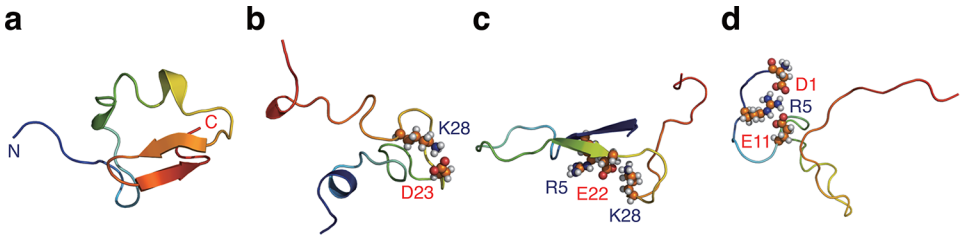


Figure 1. Representative structures of the wild-type Aβ42 protein. Each structure is color-coded according to the sequence, ranging from blue to red at the N- and C-termini, respectively. Side chains of charged residues forming salt-bridges are drawn with sphere representation (C, N, O, H: orange, blue, red, white). Structures in panels c and d exhibit the salt-bridge network (R5−E22−K28 and D1−R5−E11, respectively) consisting of two salt-bridges that include a common charged residue (E22 and R5, respectively).

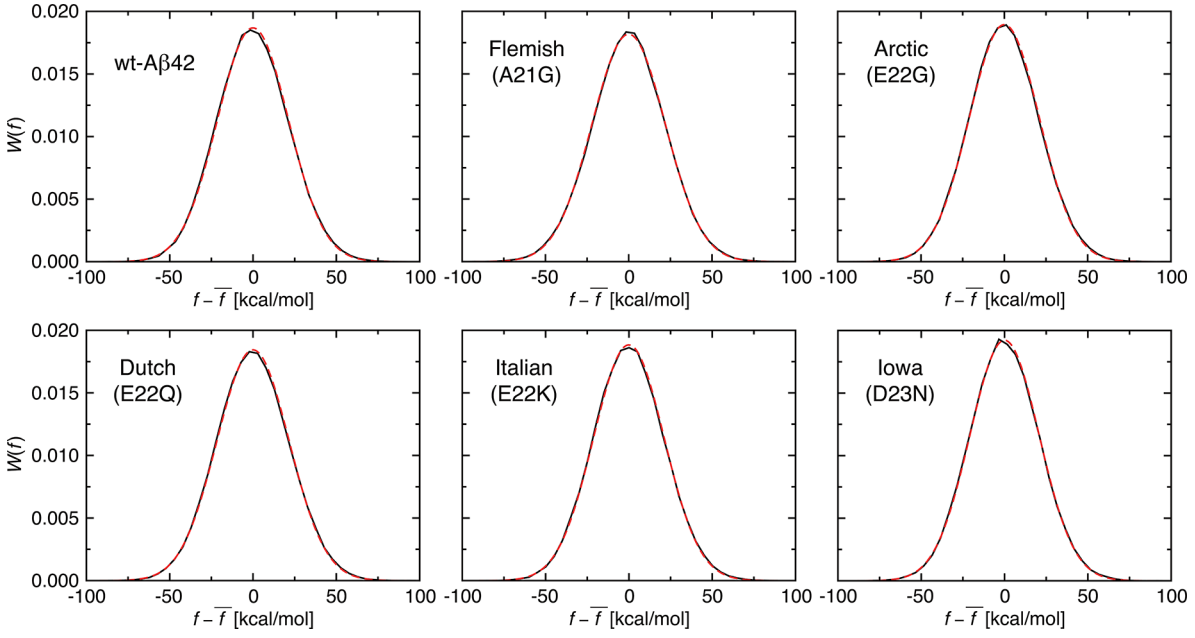


Figure 2. Probability distribution function $W(f)$ (solid curves) of the solvent-averaged protein potential energy f of the wild-type (wt) Aβ42 and its pathogenic familial mutants indicated in each panel as a function of the deviation $f - \bar{f}$ from the mean value $\bar{f}$. Dashed red curves denote the fit by the Gaussian function. Corresponding results for RR (I41R, A42R) and DQ (I41D, A42Q) mutants are presented in Figure S1 of the Supporting Information.

we explore a quantitative link between the conformational disorder and entropy in IDPs.

RESULTS

Structurally Heterogeneous Ensemble of Aβ Protein. Extensive explicit-water MD simulations were carried out at 300 K and 1 bar for the wt-Aβ42 protein and each of its mutants listed in Table 1 to investigate their structural characteristics (see the Methods section for details on the simulation procedures). While no stable tertiary structure was found from our simulations reflecting the intrinsically disordered nature in aqueous environments, Aβ protein does not exhibit completely random

dx.doi.org/10.1021/jp401049h | J. Phys. Chem. B 2013, 117, 5503−5509

conformations. In fact, certain residual structure can be recognized from Figure 1, where some "representative" conformations of the simulated wt-Aβ42 protein are displayed for illustrative purposes. By "representative" conformations, we do not mean that those structures occupy a significant portion of the conformational ensemble. Indeed, no particular structures were found from the clustering analysis of simulated conformations that are dominantly populated relative to the others. It is therefore inappropriate to describe Aβ protein using a limited number of representative conformations, and their structural characterization in terms of ensemble averages over the full conformational ensembles is essential.

While the protein conformations appear to be dominated by turn and coil, as can be inferred from Figure 1, the presence of certain fractions of regular secondary structures as well as salt-bridges and their networks is discernible. (By regular secondary structures, we mean helical and β-sheet structures. The former consists of α-helix, 3_{10}-helix, and π-helix, whereas the latter comprises parallel and antiparallel β-sheets. The salt-bridge network is considered formed if two or more salt-bridges contain a common residue as exemplified in Figure 1c and d.) The average populations of the residual helical structure, β-sheet structure, salt-bridges, and salt-bridge networks are summarized in Table 1 for the systems studied. The disordered nature of Aβ protein is reflected in the small values of the average populations of those residual structures, whereas the heterogeneous character of the conformational ensembles is manifested in the large values of the root-mean-squared fluctuations relative to the average values.

Thermodynamically Homogeneous Ensemble of Aβ Protein. To investigate thermodynamic properties of Aβ protein, we computed the protein internal energy E_u directly from the simulations and calculated the solvation free energy $\Delta\mu$ by applying the molecular integral-equation theory[23,24] to simulated protein conformations (see the Methods section for details). These two quantities were combined to define the solvent-averaged protein potential energy $f = E_u + \Delta\mu$.[22] The function f is also simply referred to as the effective energy, and defines a hypersurface in the conformational space which is often called the free energy landscape.[25] We constructed the probability distribution function $W(f)$ for f based on the conformational ensemble of each Aβ protein (see Figure 2 for the wt-Aβ42 and its familial mutants and Figure S1 in the Supporting Information for the rest). Remarkably, we find that the *heterogeneous* conformational ensemble of Aβ protein is associated with a *homogeneous*, Gaussian distribution of the effective energy. This striking observation regarding the structural heterogeneity and the thermodynamic homogeneity of the Aβ conformational ensemble illuminates that the free energy landscape of the intrinsically disordered Aβ protein comprises numerous local free energy minima, which may be separated by high free energy barriers, of similar well depth.

We also note here an interesting correlation between the solvation free energy $\Delta\mu$ and the aggregation propensity of Aβ protein. According to *in vivo* and *in vitro* mutational studies,[15−17] the aggregation propensity of Arctic, Dutch, Italian, and Iowa mutants is higher than that of wt-Aβ42, whereas the opposite trend is observed for Flemish, RR, and DQ mutants. (We have selected the particular RR and DQ mutants from numerous mutants studied in ref 17, since we wanted to include such mutants that exhibit a lower aggregation propensity than wt-Aβ42. The aggregation propensities of RR and DQ mutants are moderately and significantly lower, respectively, than that of wt-Aβ42.[17]) Inspection of the experimental aggregation propensity and the solvation free energy data summarized in Table 1 indicates that the higher (lower) aggregation propensity is associated with the increase (decrease) in the solvation free energy upon mutation. Such a correlation was demonstrated to be present also for mutants of acylphosphatase.[26]

Correlation between Conformational Entropy and Residual Structure of Aβ Protein. The Gaussian nature of the distribution function $W(f)$ allows us to estimate the protein conformational entropy S_{conf} via $TS_{conf} = \overline{\delta f^2}/(2k_BT)$ in terms of the mean-squared fluctuations of $\delta f = f - \overline{f}$;[22] i.e., S_{conf} is expressed as the extent the system explores the free energy landscape. Here, $\overline{X}$ denotes an average over the protein conformational ensemble, and k_B and T refer to Bolzmann's constant and temperature, respectively. Numerical results for TS_{conf} are summarized in Table 1. We observe that the values of S_{conf} vary significantly depending on the mutation. We also note that, unlike the solvation free energy $\Delta\mu$, there is apparently no clear correlation between S_{conf} and the aggregation propensity.

In order to explore a connection between the conformational disorder and entropy, we characterized S_{conf} in terms of the residual helical-structure content (f_{helix}), β-sheet content (f_{sheet}), and salt-bridge network content ($f_{SB\ network}$). Assuming that the influences of these residual structural contents are independent of each other and additive, we used a simple combined function to model S_{conf} (see Figure 3):

$$TS_{conf} = Af_{helix} + Bf_{sheet} + Cf_{SB\ network} + D \quad (1)$$

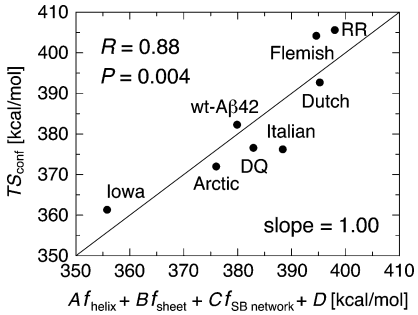


Figure 3. Characterization of the protein conformational entropy in terms of the residual structure. Protein conformational entropy TS_{conf} (multiplied by the temperature T so that it has units of energy) is modeled by $Af_{helix} + Bf_{sheet} + Cf_{SB\ network} + D$, where f_{helix} denotes the helical content, f_{sheet} the β-sheet content, and $f_{SB\ network}$ the salt-bridge (SB) network content. The Pearson correlation coefficient (R), P value, and slope of the fitted curve are also displayed.

The parameters determined from the least-squares fit are $A = -7.7$, $B = -5.5$, $C = -15.2$, and $D = 476.7$ in units of kcal/mol. Since S_{conf} is a measure of the conformational disorder, it is natural that S_{conf} is a decreasing function of the populations of the residual orders. We used the salt-bridge-network content rather than the salt-bridge content, since the latter was found to hardly affect the protein conformational entropy (see Figure S2 in the Supporting Information). The correlation between the modeled and calculated conformational entropy is highly significant ($R = 0.88$, $P = 0.004$), and the value of the slope is unity (Figure 3), indicating a direct connection between the conformational entropy and the residual structure of Aβ protein. It is remarkable that the conformational entropy is quite sensitive to subtle variations in the residual structure caused by the mutation (see Table 1).

dx.doi.org/10.1021/jp401049h | J. Phys. Chem. B 2013, 117, 5503−5509

DISCUSSION

Inherent heterogeneity in the structural characteristics of IDPs hinders the application of the conventional quasi-harmonic approximation[19−21] to compute their conformational entropy. Indeed, it has been argued that this conventional method does not provide an accurate estimate of the conformational entropy for systems that occupy multiple potential energy wells.[27] On the other hand, the absence of a stable tertiary structure implies that the free energy landscape of IDPs is rather homogeneous comprising numerous local free energy minima of similar well depth. Here, we observe that the distribution of the effective energy for Aβ protein is well approximated by the Gaussian distribution. This indicates that the conformational entropy can be evaluated by the extent of fluctuations in the effective energy.[22] In this regard, our energy-based approach provides an unprecedented way for computing the conformational entropy of structurally disordered systems.

On the basis of such a novel computational method, we present here the first quantitative attempt that connects the conformational entropy of IDPs with their residual structure. We demonstrate that the conformational entropy of IDPs is significantly correlated with the populations of the residual helical structure, β-sheet structure, and salt-bridge network. While it is reasonable that the protein conformational entropy is associated with the amount of order present in the protein, it is remarkable that the significant level of correlation is captured just with those representative residual structures. Interestingly, we find the more relevance of the salt-bridge networks than isolated salt-bridges as a determinant of the protein conformational entropy. Salt-bridges incorporated in networks have also been argued to give rise to a cooperative enhancement of interactions between charged residues, which is considered to be a major enthalpic contributor to the stability of thermophilic proteins.[28,29] The reduction in the protein conformational entropy and enthalpy by the formation of salt-bridge networks may constitute another instance of the entropy−enthalpy compensation.[30] Our characterization of the thermodynamically significant residual structure in disordered proteins will open a possibility of an interpretation of molecular recognition[11] and protein activity[12] regulated by the conformational entropy in structural terms. Our computational approach and the implications therefrom may also contribute to rational drug design in which the role of the conformational entropy is rarely addressed.[31]

Structural heterogeneity and residual structure in IDPs have also been implicated to modulate protein aggregation and amyloid-fibril formation associated with a number of human diseases. While amyloid fibrils commonly exhibit a characteristic cross-β architecture,[32] the diversity in the fibril structures, called amyloid polymorphism, has been identified depending on the growth conditions as well as on the mutation.[33] Electron microscopy studies combined with FTIR/CD and NMR measurements have revealed that such variations in morphology are linked to heterogeneity in underlying molecular structures.[34,35] Indeed, changes in the residual structure upon mutation were shown to lead to different amyloid formation.[36] Furthermore, it has been demonstrated that a single point mutation can alter the cross-β fibril morphology from commonly observed parallel to predominantly antiparallel β sheet.[37] These experimental observations highlighting the sensitivity of fibril morphologies to structural heterogeneity and to subtle changes in the protein sequence imply the relevance of the protein conformational entropy in the amyloidogenesis. In fact, we observe significant variations in the protein conformational entropy in response to minor changes in the residual structure caused by a single-point mutation.

Thermodynamic analysis for amyloidogenic Aβ proteins presented in this study also suggests distinct roles played by the conformational entropy and the solvation free energy in the protein aggregation and amyloid-fibril formation. We find that the solvation free energy change upon mutation is correlated with the change in the experimental aggregation propensity, whereas no sensible correlation is detected concerning the protein conformational entropy. This is in accord with the experimental observation in ref 36 that the extent of disruption in the residual structure caused by mutation, which would significantly alter the conformational entropy, did not show a clear correlation with the change in the aggregation rate. While the kinetic factors[38,39] were suggested to play a crucial role in the amyloid fibrillogenesis, a kinetic selection would be based on those protein conformations dictated by thermodynamics under given conditions. Our characterization of the free energy landscape of IDPs and its extension to aggregation free energy landscape will therefore contribute to elucidating the mechanisms of protein aggregation and amyloid formation.

CONCLUSIONS

Identification of thermodynamically significant residual structure in disordered systems opens a possibility of a direct interpretation of their conformational entropy in structural terms. We demonstrate in the present work the residual helical contents, β-sheet contents, and salt-bridge-network contents as major determinants of the conformational entropy of intrinsically disordered proteins. The novel computational approach employed here based on molecular dynamics simulations and the molecular theory of solvation that does not require the presence of a stable three-dimensional structure enables the calculation of thermodynamic quantities for conformationally disordered biomolecular systems, and will find a wide range of applications in protein sciences and in the structure-based drug design.

METHODS

Simulation Procedures. We carried out all-atom, explicit-water MD simulations at the temperature T = 300 K and the pressure P = 1 bar for each of the Aβ proteins listed in Table 1 using the AMBER11 simulation package.[40] In total, 10 independent simulations of 100 ns length (i.e., an accumulative simulation time of 1 μs) were conducted for each Aβ protein. We employed the ff99SB force field[41] for protein and the TIP4P-Ew model[42] for water. The simulations were performed under neutral pH, where Lys and Arg residues are positively charged and Glu and Asp are negatively charged. The particle mesh Ewald method[43] was applied for treating long-range electrostatic interactions, while a 10 Å cutoff was used for the short-range nonbonded interactions. The hydrogen atoms were constrained to the equilibrium bond length using the SHAKE algorithm,[44] which enables simulations with a 2 fs time step. Temperature and pressure were controlled by Berendsen's thermostat and barostat with coupling constants of 1.0 and 2.0 ps, respectively.[45]

Since Aβ protein in the monomeric state is intrinsically disordered and no atomic-level structures in aqueous environments are available experimentally, the initial protein systems for our simulations were carefully prepared as follows. The starting conformation for the wild-type Aβ42 was taken from the NMR

dx.doi.org/10.1021/jp401049h | J. Phys. Chem. B 2013, 117, 5503−5509

structure determined in an apolar solvent (PDB code: 1IYT[46]). The starting structures for the Aβ42 mutants were generated by Swiss PDB Viewer[47] based on the 1IYT structure for the wild-type Aβ42. Each system was then solvated by ~14 000 water molecules and neutralized by counter Na^+ ions in a cubic periodic box of side length ~75 Å. The rest of the simulation procedures are common to all the protein systems. The starting system was initially subjected to 500 steps of steepest descent minimization followed by 500 steps of conjugate gradient minimization, while Aβ protein was constrained by 500 kcal/(mol Å^2) harmonic potential. Then, the system was minimized using 1000 steps of steepest descent minimization followed by 1500 steps of conjugate gradient minimization without harmonic restraints. The system was subsequently subjected to a 20 ps equilibration process in which the temperature was gradually raised from $T = 0$ to 300 K with a constant volume. This is followed by 200 ps constant-pressure (NPT) ensemble simulation of P = 1 bar to achieve the proper simulation box size (V). We then heated the system to 600 K and carried out a 20 ns canonical (NVT) ensemble simulation. From this 600 K simulation, we took 10 protein structures with a 2 ns time interval. The Cα RMSD (root-mean-squared deviation) between any pair of these 10 structures was ~10 Å, and they were considered as independent conformations. Each of the 10 structures was then gradually annealed to 300 K: the temperature was decreased with 50 K intervals, and a 1 ns NVT simulation was performed at each of the intervening temperatures between 600 and 300 K. Subsequently, an additional 5 ns equilibration NPT ensemble simulation was performed at 300 K and 1 bar. The final structures obtained from the annealing simulations starting from the 10 different 600 K structures were used as the initial structures for the 10 independent production runs of 100 ns at 300 K and 1 bar.

For each Aβ system, 200 000 protein conformations were selected and analyzed from the respective production runs of 1 μs (10 × 100 ns) length with a 5 ps time interval. The secondary-structure contents were calculated by using the DSSP program[48] implemented in the *ptraj* module of the AMBER11 simulation package. Salt-bridges were considered formed if the distance between charged atoms was less than 3.5 Å. The RMSD-based clustering analysis was performed to explore representative protein conformations from the simulated trajectories. We made use of the MMTSB toolset[49] for this analysis with a cutoff value of 3.0 Å.

Solvation Free Energy Calculation. For each protein conformation generated by the MD simulations, we applied the three-dimensional reference interaction site model (3D-RISM) theory[23,24] to calculate the solvation free energy. 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 water site γ at position $\mathbf{r}$ around a molecular solute such as protein. For a solute−solvent system at infinite dilution, the 3D-RISM equation is given by

$$h_\gamma(\mathbf{r}) = \sum_{\gamma'} c_{\gamma'}(\mathbf{r}) * [w^{vv}_{\gamma'\gamma}(r) + \rho h^{vv}_{\gamma'\gamma}(r)] \tag{2}$$

Here $h_\gamma(\mathbf{r}) = g_\gamma(\mathbf{r}) - 1$ 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; $w^{vv}_{\gamma'\gamma}$ and $h^{vv}_{\gamma'\gamma}$ are the site−site intramolecular and total correlation functions of water; 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 Hirata[23,24]

$$h_\gamma(\mathbf{r}) = \begin{cases} \exp[d_\gamma(\mathbf{r})] - 1 & \text{for } d_\gamma(\mathbf{r}) \le 0 \\ d_\gamma(\mathbf{r}) & \text{for } d_\gamma(\mathbf{r}) > 0 \end{cases} \tag{3}$$

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 protein. We used the same numerical procedure as described in ref 24 with the fast algorithm developed in ref 50 to solve eqs 2 and 3 self-consistently. The solvation free energy $\Delta\mu$ can then be calculated from the following analytical expression which is valid under the use of the Kovalenko−Hirata closure:[23,24]

$$\Delta\mu = \rho k_B T \sum_\gamma \int d\mathbf{r} \left[\frac{1}{2} h_\gamma(\mathbf{r})^2 \Theta(-h_\gamma(\mathbf{r})) - c_\gamma(\mathbf{r}) - \frac{1}{2} h_\gamma(\mathbf{r}) c_\gamma(\mathbf{r}) \right] \tag{4}$$

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

The main limitation of the 3D-RISM theory lies in the use of an approximate closure relation, which is inherent in all the integral-equation theories. In particular, the absolute value of the solvation free energy depends on the closure relation used.[23,24] However, it is known that relative values of the solvation free energies are reasonably accurate.[23] We note in this connection that only the relative values of the solvation free energies matter in the thermodynamic quantities reported in Table 1. It is therefore expected that our results do not significantly suffer from the limitation of the integral-equation theory.

Protein Conformational Entropy. 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 $f = E_u + \Delta\mu$.[22] 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 systems studied here (see Figure 2 and Figure S1 of the Supporting Information), the protein conformational entropy is given by $TS_{conf} = \overline{\delta f^2}/(2k_B T)$ with $\delta f = f - \bar{f}$.[22] Here, $\bar{X}$ denotes an average over the set of protein conformations used in defining $W(f)$.

■ ASSOCIATED CONTENT

Supporting Information

Additional figures on the probability distribution function $W(f)$ and on the examination of the effect of the salt-bridge content on the protein conformational energy. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

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

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Mr. J. Yim and Dr. Y. Maruyama for technical support. This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF)

funded by the Ministry of Education, Science and Technology (Grant Nos. 2012-0003068 and 2012R1A2A01004687).

REFERENCES

(1) Dyson, H. J.; Wright, P. E. Intrinsically Unstructured Proteins and Their Functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197−208.

(2) Tompa, P. *Structure and Function of Intrinsically Disordered Proteins*; CRC Press: Boca Raton, FL, 2010.

(3) Dyson, H. J. Expanding the Proteome: Disordered and Alternatively Folded Proteins. *Q. Rev. Biophys.* **2011**, *44*, 467−518.

(4) Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **2006**, *75*, 333−366.

(5) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. Intrinsically Disordered Proteins in Human Diseases: Introducing the D^2 Concept. *Annu. Rev. Biophys.* **2008**, *37*, 215−246.

(6) Eliezer, D. Biophysical Characterization of Intrinsically Disordered Proteins. *Curr. Opin. Struct. Biol.* **2009**, *19*, 23−30.

(7) Bowler, B. E. Residual Structure in Unfolded Proteins. *Curr. Opin. Struct. Biol.* **2012**, *22*, 4−13.

(8) Kalodimos, C. G.; Biris, N.; Bonvin, A. M. J. J.; Levandoski, M. M.; Guennuegues, M.; Boelens, R.; Kaptein, R. Structure and Flexibility Adaptation in Nonspecific and Specific Protein-DNA Complexes. *Science* **2004**, *305*, 386−389.

(9) Henzler-Wildman, K.; Kern, D. Dynamic Personalities of Proteins. *Nature* **2007**, *450*, 964−972.

(10) Boehr, D. D.; Nussinov, R.; Wright, P. E. The Role of Dynamic Conformational Ensembles in Biomolecular Recognition. *Nat. Chem. Biol.* **2009**, *5*, 789−796.

(11) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. Conformational Entropy in Molecular Recognition by Proteins. *Nature* **2007**, *448*, 325−329.

(12) Tzeng, S.-R.; Kalodimos, C. G. Protein Activity Regulation by Conformational Entropy. *Nature* **2012**, *488*, 236−240.

(13) Hardy, J. A.; Higgins, G. A. Alzheimer's Disease: The Amyloid Cascade Hypothesis. *Science* **1992**, *256*, 184−185.

(14) Selkoe, D. J.; Podlisny, M. B. Deciphering the Genetic Basis of Alzheimer's Disease. *Annu. Rev. Genomics Hum. Genet.* **2002**, *3*, 67−99.

(15) Nilsberth, C.; Westlind-Danielsson, A.; Eckman, C. B.; Condron, M. M.; Axelman, K.; Forsell, C.; Stenh, C.; Luthman, J.; Teplow, D. B.; Younkin, S. G.; et al. Neurotoxicity and Physicochemical Properties of Aβ Mutant Peptides from Cerebral Amyloid Angiopathy. *Nat. Neurosci.* **2001**, *4*, 887−893.

(16) Murakami, K.; Irie, K.; Morimoto, A.; Ohigashi, H.; Shindo, M.; Nagao, M.; Shimizu, T.; Shirasawa, T. Neurotoxicity and Physicochemical Properties of Aβ Mutant Peptides from Cerebral Amyloid Angiopathy. *J. Biol. Chem.* **2003**, *278*, 46179−46187.

(17) Kim, W.; Hecht, M. H. Sequence Determinants of Enhanced Amyloidogenicity of Alzheimer Aβ42 Peptide relative to Aβ40. *J. Biol. Chem.* **2005**, *280*, 35069−35076.

(18) Meirovitch, H. Recent Developments in Methodologies for Calculating the Entropy and Free Energy of Biological Systems by Computer Simulation. *Curr. Opin. Struct. Biol.* **2007**, *17*, 181−186.

(19) Karplus, M.; Kushick, J. N. Method for Estimating the Configurational Entropy of Macromolecules. *Macromolecules* **1981**, *14*, 325−332.

(20) Schlitter, J. Estimation of Absolute and Relative Entropies of Macromolecules Using the Covariance Matrix. *Chem. Phys. Lett.* **2003**, *215*, 617−621.

(21) Schäfer, H.; Mark, A. E.; van Gunsteren, W. F. Absolute Entropies from Molecular Dynamics Simulation Trajectories. *J. Chem. Phys.* **2000**, *113*, 7809−7813.

(22) Chong, S.-H.; Ham, S. Configurational Entropy of Protein: A Combined Approach based on Molecular Simulation and Integral-Equation Theory of Liquids. *Chem. Phys. Lett.* **2011**, *504*, 225−229.

(23) Kovalenko, A. In *Molecular Theory of Solvation*; Hirata, F., Ed.; Kluwer Academic: Dordrecht, The Netherlands, 2003; p 169.

(24) Imai, T.; Harano, Y.; Kinoshita, M.; Kovalenko, A.; Hirata, F. A Theoretical Analysis on Hydration Thermodynamics of Proteins. *J. Chem. Phys.* **2006**, *125*, 024911.

(25) Lazaridis, T.; Karplus, M. Thermodynamics of Protein Folding: A Microscopic View. *Biophys. Chem.* **2003**, *100*, 367−395.

(26) Chong, S.-H.; Lee, C.; Kang, G.; Park, M.; Ham, S. Structural and Thermodynamic Investigations on the Aggregation and Folding of Acylphosphatase by Molecular Dynamics Simulations and Solvation Free Energy Analysis. *J. Am. Chem. Soc.* **2011**, *133*, 7075−7083.

(27) Chang, C. E.; Chen, W.; Gilson, M. K. Evaluating the Accuracy of the Quasiharmonic Approximation. *J. Chem. Theory Comput.* **2005**, *1*, 1017−1028.

(28) Kumar, S.; Nussinov, R. How Do Thermophilic Proteins Deal with Heat? *Cell. Mol. Life Sci.* **2001**, *58*, 1216−1233.

(29) Karshikoff, A.; Ladenstein, R. Ion Pairs and the Thermotolerance of Proteins from Hyperthermophiles: A 'Traffic Rule' for Hot Roads. *Trends Biochem. Sci.* **2001**, *26*, 550−556.

(30) Sharp, K. Entropy-Enthalpy Compensation: Fact or Artifact? *Protein Sci.* **2001**, *10*, 661−667.

(31) Diehl, C.; Engström, O.; Delaine, T.; Håkansson, M.; Genheden, S.; Modig, K.; Leffler, H.; Ryde, U.; Nilsson, U. J.; Akke, M. Protein Flexibility and Conformational Entropy in Ligand Design Targeting the Carbohydrate Recognition Domain of Galectin-3. *J. Am. Chem. Soc.* **2010**, *132*, 14577−14589.

(32) Tycko, R. Molecular Structure of Amyloid Fibrils: Insights from Solid-State NMR. *Q. Rev. Biophys.* **2006**, *39*, 1−55.

(33) Eichner, T.; Radford, R. E. A Diversity of Assembly Mechanisms of a Generic Amyloid Fold. *Mol. Cell* **2011**, *43*, 8−18.

(34) Petkova, A. T.; Leapman, R. D.; Guo, Z.; Yau, W.-M.; Mattson, M. P.; Tycko, R. Self-Propagating, Molecular-Level Polymorphism in Alzheimer's β-Amyloid Fibrils. *Science* **2005**, *307*, 262−265.

(35) Pedersen, J. S.; Dikov, D.; Flink, J. L.; Hjuler, H. A.; Christiansen, G.; Otzen, D. E. The Changing Face of Glucagon Fibrillation: Structural Polymorphism and Conformational Imprinting. *J. Mol. Biol.* **2006**, *355*, 501−523.

(36) Mishima, T.; Ohkuri, T.; Monji, A.; Imoto, T.; Ueda, T. Amyloid Formation in Denatured Single-Mutant Lysozymes where Residual Structures are Modulated. *Protein Sci.* **2006**, *15*, 2448−2452.

(37) Qiang, W.; Yau, W.-M.; Luo, Y.; Mattson, M. P.; Tycko, R. Antiparallel β-Sheet Architecture in Iowa-Mutant β-Amyloid Fibrils. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 4443−4448.

(38) Hwang, W.; Zhang, S.; Kamm, R. D.; Karplus, M. Kinetic Control of Dimer Structure Formation in Amyloid Fibrillogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12916−12921.

(39) Morriss-Andrews, A.; Bellesia, G.; Shea, J.-E. β-Sheet Propensity Controls the Kinetic Pathways and Morphologies of Seeded Peptide Aggregation. *J. Chem. Phys.* **2012**, *137*, 145104.

(40) Case, D. A.; et al. *AMBER11*; University of California: San Francisco, CA, 2010.

(41) Hornak, V.; Okur, R. A. 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.

(42) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665−9678.

(43) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An *N*-log(*N*) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(44) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *n*-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.

(45) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(46) Crescenzi, O.; Tomaselli, S.; Guerrini, R.; Salvadori, S.; D'Ursi, A. M.; Temussi, P. A.; Picone, D. Solution Structure of the Alzheimer Amyloid β-Peptide (1−42) in an Apolar Microenvironment. *Eur. J. Biochem.* **2002**, *269*, 5642−5648.

(47) Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: An Environment for Comparative Protein Modeling. *Electrophoresis* **1997**, *18*, 2714−2723.
(48) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22*, 2257−2637.
(49) Feig, M.; Karanicolas, J.; Brooks, C. L., III. MMTSB Tool Set: Enhanced Sampling and Multiscale Modeling Methods for Applications in Structural Biology. *J. Mol. Graphics Modell.* **2004**, *22*, 377−395.
(50) Maruyama, Y.; Hirata, F. Modified Anderson Method for Accelerating 3D-RISM Calculations using Graphics Processing Unit. *J. Chem. Theory Comput.* **2012**, *8*, 3015−3021.