JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

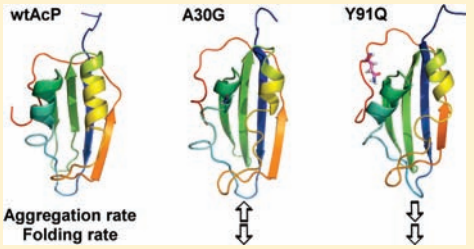
Structural and Thermodynamic Investigations on the Aggregation and Folding of Acylphosphatase by Molecular Dynamics Simulations and Solvation Free Energy Analysis

Song-Ho Chong, Chewook Lee, Guipeun Kang, Mirae Park, and Sihyun Ham*

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

Supporting Information

ABSTRACT: Protein engineering method to study the mutation effects on muscle acylphosphatase (AcP) has been actively applied to describe kinetics and thermodynamics associated with AcP aggregation as well as folding processes. Despite the extensive mutation experiments, the molecular origin and the structural motifs for aggregation and folding kinetics as well as thermodynamics of AcP have not been rationalized at the atomic resolution. To this end, we have investigated the mutation effects on the structures and thermodynamics for the aggregation and folding of AcP by using the combination of fully atomistic, explicit-water molecular dynamics simulations, and three-dimensional reference interaction site model theory. The results indicate that the A30G mutant with the fastest experimental aggregation rate displays considerably decreased α1-helical contents as well as disrupted hydrophobic core compared to the wild-type AcP. Increased solvation free energy as well as hydrophobicity upon A30G mutation is achieved due to the dehydration of hydrophilic side chains in the disrupted α1-helix region of A30G. In contrast, the Y91Q mutant with the slowest aggregation rate shows a non-native H-bonding network spanning the mutation site to hydrophobic core and α1-helix region, which rigidifies the native state protein conformation with the enhanced α1-helicity. Furthermore, Y91Q exhibits decreased solvation free energy and hydrophobicity compared to wild type due to more exposed and solvated hydrophilic side chains in the α1-region. On the other hand, the experimentally observed slower folding rates in both mutants are accompanied by decreased helicity in α2-helix upon mutation. We here provide the atomic-level structures and thermodynamic quantities of AcP mutants and rationalize the structural origin for the changes that occur upon introduction of those mutations along the AcP aggregation and folding processes.



1. INTRODUCTION

Globular proteins may convert their native conformations into non-native forms due to the intrinsic or external perturbations in the microenvironmental conditions.[1] When a protein is in its non-native state, it can further unfold or self-assemble to form amyloid aggregates, which are presumably toxic and consequently can cause various diseases. In this regard, protein aggregation is one of the most actively investigated issues in relation to the development of the therapeutics and medical applications for the cure of protein aggregation diseases.[2−5]

Amyloid fibrils are insoluble deposits of otherwise soluble protein that self-assemble into fibril. Peptide and protein structures have been shown to convert into ordered aggregates known as amyloid fibril.[3,6] Amyloid and other threadlike aggregates display the common structural motif of ordered β-pleated sheet conformation.[7] For example, amyloid-β peptide forming amyloid fibrils related to Alzheimer's disease consists of two β-strand conformations linked with loop, and their aggregated structures adopt in-register β-sheet structure.[8,9] β_2-Microglobulin also forms an amyloidogenic intermediate which comprises five β-strands at low pH.[10] The native tetrameric conformation of transthyretin dissociates and adopts partially unfolded state that is prone to form an amyloid fibril below pH 5.0.[11,12] Particularly, acylphosphatase (AcP) protein in which the overall fold is determined by α/β packing is one of the good examples to investigate the kinetics and thermodynamics of the protein aggregation because of its ability to form fibrillar assemblies of amyloid type.[13−26] The important observation is that AcP forms protein aggregates in a definite manner under proper environment and consequently is organized into amyloid structures associated with protein deposit disease.[15,18]

The protein engineering method to study the mutation effects on AcP unfolding and aggregation has been actively applied to describe amyloid fibril formation as well as protein structure−function relationships.[18,24,27−40] Dobson and co-workers have systematically reported the effects of single point mutations of AcP and found that V9A, E29D, A30G, S87T, and S92T mutations showed increased aggregation rate, whereas V20A, G34A, V39A, L89A, Y91Q, and Y98Q mutations displayed the opposite result.[14,20,41] Among those mutations, A30G accelerates and Y91Q decelerates the aggregation rate most distinctively.[13,14,41] On the other hand, both of those mutations contribute to slow down the folding rate of AcP with respect to the wild type.[14,23]

Received: December 25, 2010
Published: April 18, 2011

ACS Publications © 2011 American Chemical Society

dx.doi.org/10.1021/ja1116233 | J. Am. Chem. Soc. 2011, 133, 7075–7083

Despite the extensive mutation experiments on AcP protein, the molecular origin and the structural motifs to promote the aggregation and folding rate changes have not been rationalized at the atomic-level due to the absence of the AcP mutant structures. Flöck et al. have previously performed the molecular dynamics (MD) simulations for AcP monomer of its wild type under 25% (v/v) trifluoroethanol (TFE)/water solvent and found the enhancement of aggregation propensity mainly due to the TFE and β-propensity of the α1-helix (residues 22−33).[42] To date, there has been no report to identify the atomic-level mutant conformations of AcP in water environment.

To this end, we have investigated the structural characterization and solvation thermodynamics for the wild type AcP and its two mutants, A30G and Y91Q, by using fully atomistic, explicit-water MD simulations as well as three-dimensional reference interaction site model (3D-RISM) theory. The aims of this work are to answer the following three questions: (1) What are the structural similarities and differences between A30G and Y91Q mutant proteins with respect to the wild type AcP (mutation effect on the AcP structure)? (2) What are the molecular motifs corresponding to the experimentally observed aggregation and folding rate changes (mutation effect on the AcP aggregation/folding kinetics)? (3) What is the effect of solvation on aggregation kinetics based on the solvation thermodynamics and hydrophobicity (solvation effect on the AcP aggregation kinetics)? All-atom explicit water MD simulations reveal to what extent a single point mutation alters the protein structure in AcP as well as the molecular origin responsible for the aggregation and folding kinetics. On the basis of the 3D-RISM theory to compute solvation thermodynamic quantities, we show that the hydrophobicity changes upon mutations in AcP originate mainly from hydrophilic residues rather than from hydrophobic ones. Moreover, the structural characterization of those mutants allows us to bridge the gap between the experimentally observed aggregation/folding kinetics and the solvation thermodynamic parameters associated with the AcP monomer.

2. RESULTS AND DISCUSSION

2.1. Global Protein Topologies Are Reserved upon Introduction of Single Point Mutations in AcP. Structure for wild type AcP (wtAcP) protein with a sequence of $^{(1)}$STARPLKSVD YEVFGRVQZV CFRMYAEDEA RKIGVVGWVK NTSKG-TVTGQ VQGPEEKVNS MKSWLSKVGS PSSRIDRTNF SNEKTISKLE YSNFSVRT$^{(98)}$ was taken from the NMR determined structure of horse muscle AcP (PDB ID: 1APS[43]). Two mutants of wtAcP, A30G and Y91Q, were chosen because the changes of aggregation rate, $\ln(\nu_{mu}/\nu_{wt})$, with respect to the wild-type AcP are most dramatic, +2.0 and −3.3 for A30G and Y91Q, respectively.[14,41] The starting structures of two mutants were modeled based on the NMR structure of wtAcP. AcP protein consists of two α-helices (residues 22−33 and 55−65, denoted as α1 and α2, respectively) and five β-strands (residues 7−13, 36−42, 46−53, 77−85, and 93−97, denoted as β1, β2, β3, β4, and β5, respectively). Each secondary structure is linked to five loops (residues 14−21, 34−35, 43−45, 66−76, and 86−92, denoted as L1, L2, L3, L4, and L5, respectively) (see Figure 1).

The conformational characteristics of wild type AcP, A30G, and Y91Q were investigated by using MD simulation method at 298 K with explicit water. The distributions of the Cα root-mean-square deviation (RMSD) of wtAcP, A30G, and Y91Q with respect to their initial structures are displayed in Figure 2a. During 20 ns simulations, the Cα RMSD values for all three proteins stay fairly low with the average Cα RMSD values of 2.69 Å (standard deviation: 0.25 Å), 2.78 Å (0.15 Å), and 2.84 Å (0.21 Å) for wtAcP, A30G, and Y91Q, respectively. The distributions of the radius of gyration (R_g) as a function of time are shown in Figure 2b. In Figure 2b, the range of R_g distributions is quite restricted for all three cases, which indicates that the global structures of three proteins are relatively conserved. The representative structures of each simulation for three proteins based on the clustering analysis[44] are displayed in Figure 3.

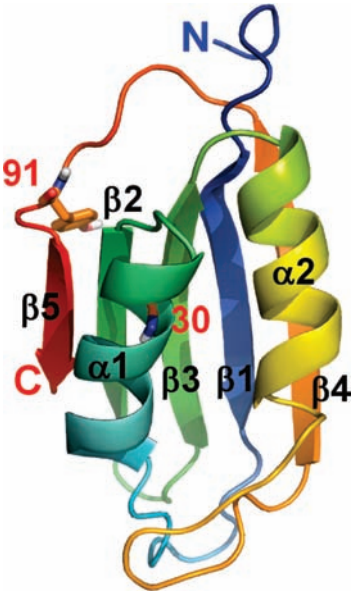


Figure 1. Ribbon representation of AcP. The residues mutated in this study are also represented in sticks (C, O, N, and H are orange, red, blue, and white, respectively) and the positions of mutated sites are written in red. The structure is color-coded according to the sequence, ranging from blue to red at N- and C-termini, respectively. The positions of five loops are not labeled for the clearness. The order of secondary structures is as follows: β1(residues 7−13)−L1(14−21)−α1(22−33)−L2(34−35)−β2(36−42)−L3(43−45)−β3(46−53)−α2(55−65) − L4(66−76)−β4(77−85)−L5(86−92)−β5(93−97).

The averaged secondary structure contents for full-length wtAcP and its two mutants are summarized in Table S1 (see Supporting Information). On average, the α-helical contents are observed to be lower in A30G mutant compared to wtAcP or Y91Q. Nevertheless, on the basis of the above structural analyses, the global protein structure and the overall topology for AcP are quite reserved by A30G and Y91Q mutations at 298 K.

2.2. The A30G Mutation Leads to Substantial Hydrophobic Core Disruption and α1-Helix Destabilization. Although the global structures are reserved between wild type and two mutants of AcP, the local structural features are quite distinctive. The heterogeneities in conformational dynamics between wtAcP and two mutants are demonstrated in the plot of the number of hydrophobic contacts in Figure 2c. The average number of hydrophobic contacts during 20 ns for A30G mutant is 21% smaller than that for wtAcP or Y91Q. Noticeably, the A30G mutation introduces substantial disruption of the hydrophobic core interactions between the methyl group of residue 30 and the surrounding hydrophobic residues of A26, I33, V35, V51, M61, L65, and F94 (denoted as the hydrophobic core). As seen in

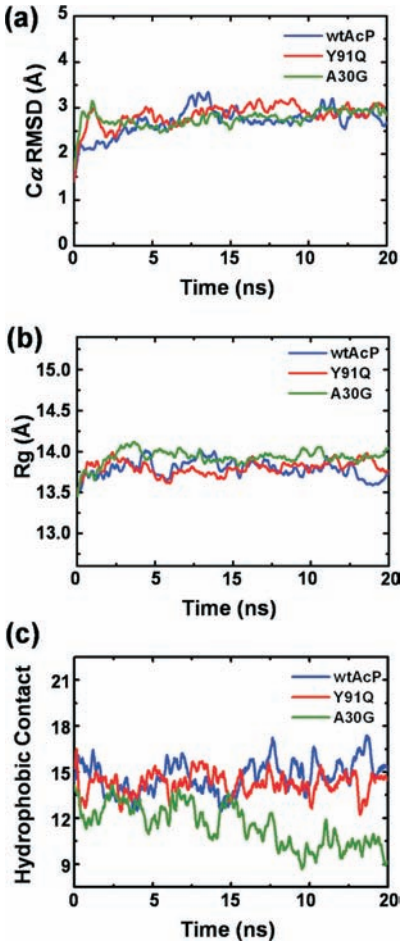


Figure 2. Structural characteristics of AcP and two mutants calculated by MD simulations at 298 K (blue line, wtAcP; green, A30G mutant; red, Y91Q mutant). (a) Cα RMSDs from the initial structure as a function of simulation time. (b) Radius of gyration as a function of simulation time. (c) The number of hydrophobic contacts as a function of simulation time.

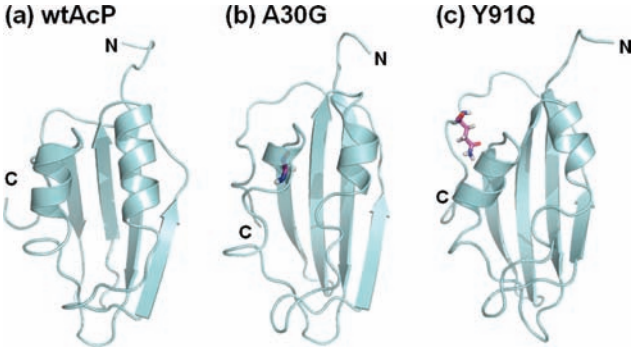


Figure 3. Representative structures of (a) wtAcP, (b) A30G, and (c) Y91Q. Mutated residues are shown in stick representation (C, O, N, and H are magenta, red, blue, and white, respectively).

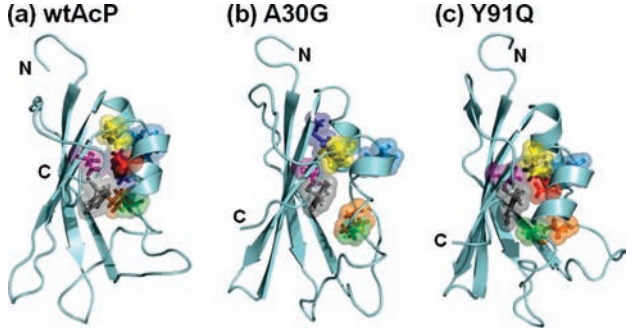


Figure 4. Hydrophobic interactions of (a) wtAcP, (b) A30G, and (c) Y91Q in their representative structures. The hydrophobic core is located at the inner space between β2, β3, α1 and α2. Only hydrophobic residues around Ala30 (or Gly30) are depicted with different colors (red, Ala30; green, Ala26; sky blue, Ile33; yellow, Val35; purple, Val51; deep blue, Met61; orange, Leu65; gray, Phe94). Only side chains, Cα, and Hα atoms are shown for clarity.

Figure 4, the hydrophobic interactions in wtAcP and Y91Q are fairly well conserved over time. However, in the absence of methyl side chain in residue 30 of A30G mutant, substantial disruptions are monitored in the hydrophobic core regions consisting of α1-, β2-, and β5-regions. This indicates that the methyl group in residue 30 plays a critical role in forming hydrophobic interactions in the core. Overall, the A30G mutation causes disruption in hydrophobic interactions around the site of mutation and decreases thereby protein stability of the native conformation.

To further characterize the mutation effect on the hydrophobic core disruption in connection with the helical contents, we analyzed the average helical structure compositions (see Table 1 and Supporting Information Table S2) and the average hydrogen bond occupancies (Supporting Information Table S3) for three proteins based on the MD trajectories. Upon A30G mutation, the α-helical contents in the α1-region were substantially decreased by 35.1% and the consequent decrease in the number of H-bonding by 24.7% was displayed with respect to the case of wtAcP. To analyze residue-specific helical propensity, we calculated the probability of each residue in helix regions to be assigned as α-helix or 3_{10}-helix, and the result is shown in Figure 5. It is noted that the 3_{10}-helical contents in A30G mutant

Table 1. Average Helical Compositions (%) of wtAcP and Two Mutants (A30G and Y91Q) Obtained from the Explicit Solvent MD Simulations at 298 K

	all	α-helix	3_{10}-helix	π-helix
Helical Contents in α1-Helix (Residues 22−33)				
wtAcP	52.92	52.40	0.34	0.18
A30G	29.75	17.27	12.47	0.01
Y91Q	59.71	53.67	5.99	0.05
Helical Contents in α2-Helix (Residues 55−65)				
wtAcP	85.33	82.98	2.35	0.00
A30G	66.09	33.24	32.85	0.00
Y91Q	66.11	49.18	16.93	0.00

is observed to be higher than the other cases. The 3_{10}-helix has been suggested as a thermodynamic intermediate in the α-helix folding pathway.[45] Figure 6 represents the occupancies of 1,4-hydrogen bonding in α1- and α2-regions and interstrand hydrogen bonding in β1−β5 beta-sheet regions. Disruption of the H-bonding network in α1-region by A30G mutation is caused by the local perturbation of the hydrophobic interaction

dx.doi.org/10.1021/ja1116233 |J. Am. Chem. Soc. 2011, 133, 7075–7083

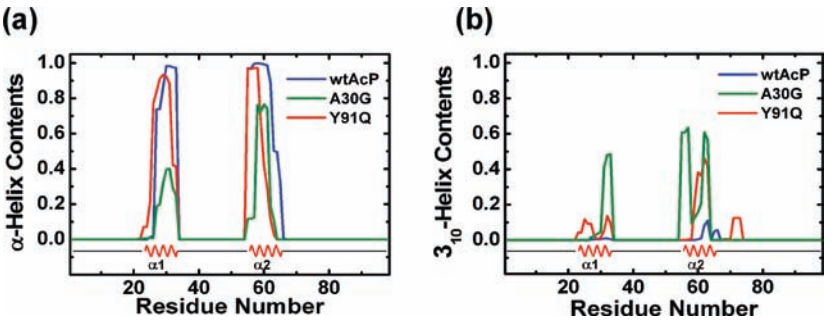


Figure 5. Average helical contents of wtAcP and its mutants during 20 ns MD simulations: (a) α-helix contents and (b) 3_{10}-helix contents.

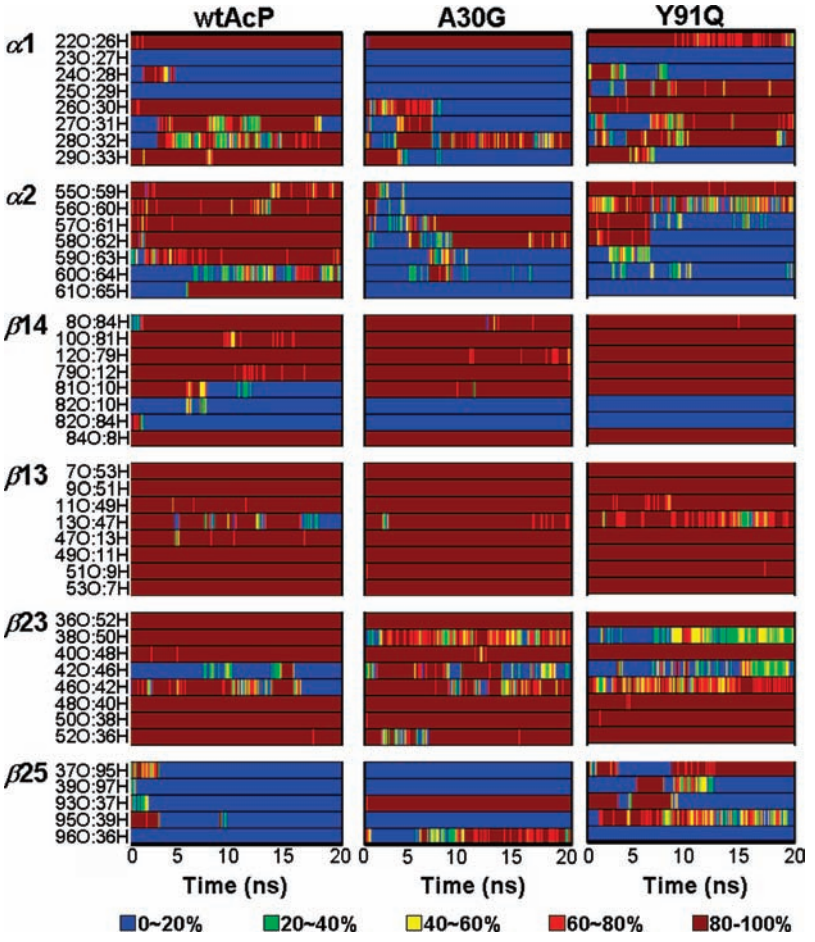


Figure 6. Hydrogen-bond occupancies of wtAcP and A30G and Y91Q mutants. In case of α1- and α2-helices (the first two rows), only α-helical $(i, i+4)$ hydrogen bonds are considered. In case of β-sheets, hydrogen bonds between adjacent two strands are considered. βMN (presented at the last four rows) refers to the interstrand hydrogen bonds formed between βM and βN.

around the methyl group of residue 30 as well as by the absence of alanine side chain, which employs the intrinsic propensity of forming α-helical structure. It is fairly notable from Table 1 as well as from the representative structures shown in Figure 3 that the A30G mutation distinctively destabilizes helical structure in α1-region, and concomitant conformational changes toward unfolding direction may be accelerated. It is in good agreement with the experimental results that A30G mutation accelerates aggregation rate[14,20,41] and that α1-helix is important for aggregation process.[23]

2.3. The Formation of Non-Native H-bonding Network Dominantly Rigidifies the α1-Helical Region in Y91Q Mutant. Y91Q mutant displays a higher degree of average helical contents in the α1-region compared to wtAcP (see Table 1 and Supporting Information Table S2). It is due to the formation of a non-native hydrogen-bonding network, i.e. G37[main chain (MC)]−Q91[side chain (SC)], Q91(SC)−N93(MC), E27-(SC)−N93(SC), and G37(MC)−S95(MC) connecting α1-, β2-, and β5-regions of Y91Q (see Figure 7). This non-native hydrogen-bonding network between main chains of β5-regions

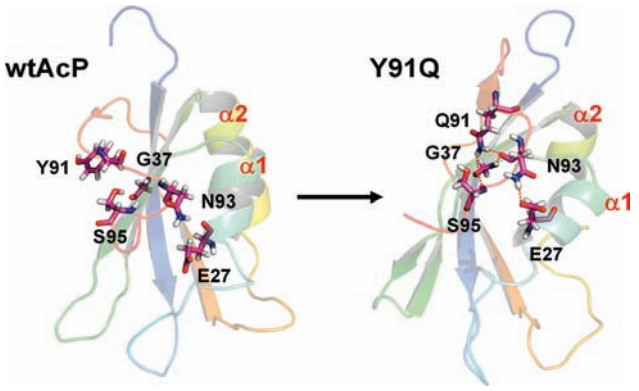


Figure 7. Formation of non-native interactions in Y91Q mutant (G37−Q91, Q91−N93, E27−N93, and G37−S95) (C, O, N, and H are magenta, red, blue, and white, respectively).

and the side chain of E27 contributes to the overall protein rigidity as well as the increased α1-helicity of Y91Q mutant. As a result, the helical contents in the α1-region of Y91Q mutant are increased compared to that of wtAcP by 6.8% and much greater than that of A30G by 30.0%. The Y91Q mutation thus stabilizes the α1-helix by developing the hydrogen-bonding network around the site of mutation, and this local non-native H-bond interaction promotes the increased helicity in the α1-region. This simulation result provides the molecular structures for AcP mutants, which are in agreement with the observed experimental aggregation kinetics.[14,20,23]

2.4. The *β*-Sheet Structures Are Fairly Well Conserved in the Wild Type AcP as Well as in Its Mutants. In contrast to the heterogeneous fluctuations in helical contents by a single point mutation, the *β*-sheet structures are fairly well conserved in all three proteins (see Supporting Information Tables S1 and S4). Thus, structural variations and dynamic heterogeneity associated with the single point mutation in AcP seem to be attributed mainly from the intrinsic α-helical propensities in α1- and α2-helix regions, which display the most distinctive variations in protein structures upon mutations.

2.5. Folding Kinetics upon A30G and Y91Q Mutations. The average α-helical contents in α2-region (residues 55−65) of both A30G and Y91Q mutants were simulated to be lower by 49.7% and 33.8% than those of wtAcP, respectively (see Table 1). This is mainly due to the non-native H-bond formation between the main chain of Trp64 in the α2-region and the side chain of Thr78 in the *β*4-region. This indicates that both A30G and Y91Q mutations contribute to destabilize the helix-forming propensity in the α2-region significantly. Previously, it was suggested that the α2-helicity plays a critical role in determining the slower folding rate of both A30G and Y91G mutants compared to that of wild-type AcP.[14,23] This implies that the α2-region in the transition state of AcP folding contains higher helical contents in wtAcP compared to those in A30G or Y91Q mutants. On the basis of our MD simulations and the experimental data, the folding transition state of wtAcP is expected to have higher degrees of helicity in the α2-region compared to those of A30G and Y91Q mutants. This result is in agreement with the experimental finding that the higher degrees of helicity in the α2-region of wtAcP compared to those of A30G and Y91Q mutants have relevance to the accelerated folding rate in wtAcP versus that of the other two mutants.[23] The decrease in helical contents in the α2-region of both A30G and Y91Q mutants in their native states

Table 2. Mutation Effects on the α-Helical Contents in AcP[a]

	α1-helical contents		α2-helical contents	
	MD	experiments	MD	experiments
A30G	↓	↓ (k_{agg} ↑)	↓	↓ (k_{fold} ↓)
Y91Q	↑	↑ (k_{agg} ↓)	↓	↓ (k_{fold} ↓)

[a] Comparison between MD simulation results and the experimental data[14,20,23] is displayed upon A30G and Y91Q mutations on the aggregation (k_{agg}) and folding (k_{fold}) rate changes.

may directly contribute to less stable protein and slower folding rate compared to those of wtAcP.

Table 2 demonstrates that the changes in α1- and α2-helical contents upon A30G and Y91Q mutations from the present MD simulations are in accord with the experimental observations[14,20,23] that (i) the aggregation rate is closely related to the helical contents in the α1-region while the folding rate is related to those of the α2-region, (ii) stabilization and destabilization of the α1-helix cause the aggregation process to be decelerated and accelerated, respectively, and (iii) destabilizing the α2-helix decelerates the folding rate.

2.6. Solvation Thermodynamics upon A30G and Y91Q Mutations. Although MD-simulated mutant structures for A30G and Y91Q exhibit the experimentally proposed structural features that can be used to speculate the origin for the experimental aggregation and folding rate changes, there is still missing information on how the conformational transitions of AcP mutants in their monomeric state would affect the corresponding aggregation kinetics. How do the protein conformational changes in AcP monomer effect the self-assembly kinetics toward the aggregation process? The aggregation kinetics would be affected by many different parameters involving not only intrinsic nature of proteins but also environmental factors. In particular, the protein stability caused by the protein conformation and the protein solubility affected by the solvent distribution around the protein would determine the propensity to aggregate from monomer. In this regard, although solvation free energy might not be the most distinctive factor to affect aggregation kinetics, it can certainly provide the molecular origin for aggregation kinetics based on the protein conformational changes by mutation.

To further understand the AcP mutation effects on the aggregation rate changes, we therefore performed the solvation free energy calculations for the wild type AcP and its two mutants by employing the 3D-RISM theory (see Methods). In addition to the solvation free energy ($\Delta\mu$), the solvation entropy (Δs) and the solvation energy ($\Delta\varepsilon$) and its decomposition into the average protein−water interaction energy ($\Delta\varepsilon^{uv}$) and the solvent reorganization energy ($\Delta\varepsilon^{vv}$) were calculated. The changes, e.g., in the solvation free energy upon mutation, were calculated via $\Delta\Delta\mu = \Delta\mu_{mu} - \Delta\mu_{wt}$ of the solvation free energies of the mutant ($\Delta\mu_{mu}$) and of the wild type ($\Delta\mu_{wt}$), and the results are summarized in Table 3. The solvation free energy for A30G is much higher than that of wtAcP by 127.2 kcal/mol, suggesting that A30G is more hydrophobic and consequently more prone to undergo self-assembly. On the other hand, the solvation free energy for Y91Q is found to be lower by 25.1 kcal/mol than that for wtAcP, indicating that Y91Q mutant is less hydrophobic and less prone to aggregate. Such hydrophobicity changes for A30G and Y91Q mutants are in accord with experimentally observed aggregation rate changes.[14,23]

dx.doi.org/10.1021/ja1116233 |*J. Am. Chem. Soc.* 2011, 133, 7075–7083

Table 3. Difference of the Thermodynamic Functions of Solvation of Two Mutants, A30G and Y91Q, with Respect to wtAcP (in kcal/mol)[a]

	$\Delta\Delta\mu$	$T\Delta\Delta s$	$\Delta\Delta\varepsilon$	$\Delta\Delta\varepsilon^{uv}$	$\Delta\Delta\varepsilon^{vv}$
A30G	+127.2	+8.6	+135.8	+277.8	−142.0
Y91Q	−25.1	+3.0	−22.1	−25.7	+3.6

[a] $\Delta\Delta\mu = \Delta\mu_{mu} - \Delta\mu_{wt}$, the solvation free energy difference; $T\Delta\Delta s = T\Delta s_{mu} - T\Delta s_{wt}$, the solvation entropy difference multiplied by the temperature $T = 298$ K; and $\Delta\Delta\varepsilon = \Delta\varepsilon_{mu} - \Delta\varepsilon_{wt}$, the solvation energy difference. The solvation energy difference $\Delta\Delta\varepsilon$ is further decomposed into the contributions from the average protein−water interaction energy ($\Delta\Delta\varepsilon^{uv}$) and the solvent reorganization energy ($\Delta\Delta\varepsilon^{vv}$).

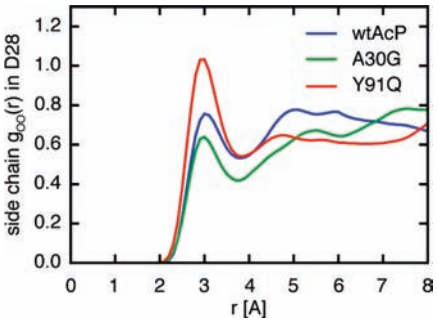


Figure 8. Radial distribution function $g_{OO}(r)$ between the side chain carboxyl oxygen in D28 and the water oxygen calculated by the 3D-RISM theory (blue line, wtAcP; green, A30G mutant; red, Y91Q mutant).

2.7. Structural Origin for Hydrophobicity Changes upon A30G and Y91Q Mutations. To investigate the structural origin for the overall hydrophobicity changes, we carried out the decomposition of the average protein−water interaction energy changes $\Delta\Delta\varepsilon^{uv}$ into the different regions of the protein, on the basis of the good correlation between the changes in $\Delta\Delta\mu$ and $\Delta\Delta\varepsilon^{uv}$ (see Table 3 and section 4.3). From this analysis, we observed that the $\Delta\Delta\varepsilon^{uv}$ change from the α1-region upon A30G mutation is 94.7 kcal/mol and that upon Y91Q mutation is −13.7 kcal/mol. Noticeably, the overall solvation energy change is found to originate mainly from the α1-region in both mutant cases. Furthermore, over 80% of these changes in α1-regions of both A30G and Y91Q arise from side chain contributions.

We therefore examined the energetics and solvation structure of side chains in each protein. The average interaction energy between D28 in the α1-helix and solvent water, which is found to exhibit the most profound change upon mutation, is increased by 71.5 kcal/mol upon A30G mutation and is decreased by 11.3 kcal/mol upon Y91Q mutation. Radial distribution functions showing associated solvation structure changes of D28 are presented in Figure 8. (See Figure S1 in Supporting Information for all the residues in α1-region.) In A30G mutant, two carbonyl oxygens in the side chain of D28 are more dehydrated than for wtAcP due to the formation of a H-bond with the nearby hydrophilic side chains of R31 and K32, which is possible due to the α1-helix deformation upon A30G mutation (see Figure 9). In Y91Q mutant, on the other hand, the rigid α1-helix due to the non-native H-bonding network (see Figure 7) does not allow such polar side chain−side chain interactions on D28, and therefore D28 is fully hydrated (Figure 9). Furthermore, we found that

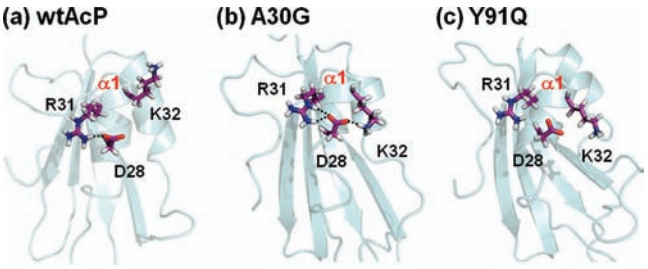


Figure 9. Polar side chain−side chain interactions indicated by dotted lines involving D28, R31, and K32 in the α1-region of (a) wtAcP, (b) A30G, and (c) Y91Q. Side chains of these selected residues are shown in stick representation (C, O, N, and H are magenta, red, blue, and white, respectively).

the main contributors to such hydrophobicity changes upon AcP mutation are hydrophilic residues rather than hydrophobic ones, which is reasonable, since the energetic contribution is larger in the former than the latter. Moreover, those changes mainly originate from side chains rather than main chains.

Overall, we observed that the solvation free energy changes upon the introduction of point mutation in AcP are greatly influenced by the hydrophilic residues rather than by hydrophobic residues. Upon A30G mutation, the solvation free energy is greatly increased by the decrease in hydration of hydrophilic side chains due to the formation of side chain−side chain electrostatic interactions, which is caused by the partial helix unfolding in the α1-region of A30G. With emphasis, it is noted that the faster aggregation rate in A30G presumably associated with the increased hydrophobicity in A30G monomer comes from the less hydrophilic, thus more dehydrated, side chains located in the disrupted α1-helix region of A30G. Similarly, the slower aggregation rate in Y91Q is a result from the more hydrophilic, thereby more solvated, side chains, which are caused by more rigid helix formation in the monomeric state.

3. CONCLUSIONS

By performing fully hydrated, all-atom MD simulations along with 3D-RISM theory calculations, we have investigated the atomic-level structural variations and solvation thermodynamics for AcP and its two mutants of experimental interests (A30G and Y91Q) to understand the molecular origin for the aggregation and folding kinetics as well as thermodynamics. Although the overall protein topologies are fairly conserved between wtAcP and two mutants, the heterogeneous variations in the secondary structures and in local fluctuations are observed upon mutations. The observed local structural changes upon single point mutations are directly projected onto the hydrophobicity changes in AcP. Moreover, the decomposition analysis of the average protein−water interaction energy highlights the fact that the changes in solvation free energy mainly originate from hydrophilic residues rather than from the hydrophobic residues.

The A30G mutation causes an overall decrease in the protein stability, mainly attributed by the hydrophobic core disruption due to the deformation in α1-helicity by 23.2%. A30G with disrupted α1-helix region allows the side chain interactions of hydrophilic residues in the α1-helix region to form salt bridges. The hydrophilic side chain interactions induce hydrophilic dehydration in the α1-region, which dominantly contributes to the overall solvation free energy increase as well as the

hydrophobicity increase in A30G. A30G with increased hydrophobicity compared to the wild type AcP would be much more prone to aggregate each other in aqueous environment. In contrast to A30G mutation, Y91Q causes the formation of a non-native hydrogen-bonding network connecting α1-, β2-, and β5-regions, which contributes to the overall protein rigidity as well as the increased α1-helicity of Y91Q mutant. The increased helicity in the α1-region reinforces the polar side chains to be fully hydrated by solvent water. This hydrophilic hydration contributes to the decreased hydrophobicity and decreased solvation free energy in Y91Q. Consequently, more hydrated Y91Q monomer would have to pay more dehydration energy penalty to be aggregated.

On the other hand, both mutations contribute to the lower degrees of α2-helicity due to the non-native H-bond formation between the main chain of Trp64 in the α2-region and the side chain of Thr78 in the β4-region. Decreased α2-helicity in both mutants observed by MD simulation is in accord with the experimental observations that folding rate in A30G and Y91Q is slower than in wtAcP and that α2-helicity plays a critical role in determining folding rate. Nonetheless, the effect on the conformational changes upon single point mutations in AcP can occur not only around the mutation site but also long-range through the propagation of favorable interactions. The MD simulation results along with the 3D-RISM theory render the structural motif and molecular origin to rationalize the experimentally observed rate changes of aggregation and folding by the single point mutation.

4. METHODS

4.1. MD Simulation Methods. All MD simulations were performed by using the SANDER module of the AMBER 9 simulation package[46] with force field ff99.[47] Two mutants (A30G and Y91Q) of AcP were generated by Swiss PDB Viewer.[48] Simulations of the wtAcP and its two mutants were performed for 20 ns at 298 K under pH range of 5−7, where Lys and Arg residues are positively charged and Glu and Asp are negatively charged.

The starting protein system was explicitly solvated with 6882 TIP3P water molecules[49] in the rectangular box where the distance to the edge of the solvent box from protein was chosen to be 12 Å and periodic boundary condition was applied. To neutralize the system, six chloride ions were added. The particle mesh Ewald method[50] was applied for treating long-range electrostatic interactions, and a 8.0 Å force-shifted cutoff was used for short-range nonbonded interactions. The hydrogen atoms were constrained to the equilibrium bond length using the SHAKE algorithm.[51] In order to remove unfavorable van der Waals contacts, these systems were subjected to 1000 steps of steepest decent minimization followed by 4000 steps of conjugate gradient minimization while protein was constrained by 500 kcal(mol Å^2)$^{-1}$ harmonic potential. Then, the whole systems (protein with water) were minimized using 10 000 steps of steepest decent minimization without harmonic restraints. The systems were subsequently subjected to 20 ps equilibrium process in which the temperature was gradually raised from 0 to 298 K under atomic restraints. After the equilibration step, the production runs were carried out for 20 ns with 1 fs time step and with *NPT* ensemble, i.e., a constant number of particle (N), pressure (P), and temperature (T). Temperature and pressure were controlled by a Berendsen thermostat and barostat with coupling constant of 1.0 and 2.0 ps, respectively.[52] All trajectories were recorded every 1 ps. For each protein, three independent MD simulation trajectories were performed with different initial velocities; in total, 60 ns trajectories were performed for each protein. The statistics of secondary structural analysis data based on all MD trajectories are summarized in Supporting Information Tables S1 and S2. In the main text, we presented the results from the representative trajectory for each protein system.

4.2. MD Trajectory Analysis. The trajectories were analyzed using several order parameters. The root-mean-square deviation (RMSD) value is a useful tool for quantifying the conformational changes between two structures of the same protein. In this study, Cα RMSDs for three proteins from the initial structure were calculated. The size of protein was roughly estimated by radius of gyration (R_g), which is defined as the mass-weighted positional mean of the distances of atoms from the center of mass. To analyze the fluctuation and statistics of the secondary structure of AcP and two mutants, the dictionary of secondary structure of proteins (DSSP) program developed by Kabsch and Sander[53] was used. All secondary structures were characterized by DSSP and were classified into three helix types (α-helix, 3_{10}-helix, and π-helix), two sheet types (antiparallel and parallel β-sheet, but in this study both were denoted as only "β-sheet"), and others (turn and coil). The hydrophobic contact for a pair of the hydrophobic side chains of residues (Ala, Cys, Phe, Leu, Met, Ile, and Val) was considered to be formed when a minimal distance between the center of mass of any pair of residues is less than 5.4 Å. The definition of hydrophobic residue is based on the method defined by Kyte and Doolittle.[54] To evaluate the local structure fluctuation of helices and β-sheets, the hydrogen-bond populations were calculated. The hydrogen bond is defined if the distance between the donor (D) and the acceptor (A) is shorter than 3.6 Å and the angle (A···H−D) is larger than 135.0°. For α-helical hydrogen bonds, only (i, i + 4) backbone hydrogen bonds were considered, whereas for the β-sheet, hydrogen bonds between adjacent two strands were considered. The clustering analysis was employed to generate the pools of conformation for each MD trajectory. The clustering method is based on the root-mean-square deviation (RMSD) as a measure of the distance between any two given conformations within the trajectory. By using this clustering method, we were able to figure out each cluster by one centroid structure, since all structures in each cluster are within the cutoff value in RMSD. Depending on the cutoff value of RMSD, the total number of clusters and their centroid structures were determined. The cutoff value for neighbor counting is set to 1.0 Å. We made use of the MMTSB toolset[44] for this analysis. Also various distance analyses between two specific atoms were directly measured from MD trajectories. The *ptraj* toolset available in the AMBER program package and in-house analysis programs were used to analyze the resulting MD trajectories. The VMD[55] and PyMol[56] software were used for the preparation of structural figures and visualization for various analyses in this study.

4.3. Solvation Free Energy Calculation Based on the 3D-RISM Theory. Thermodynamic functions of solvation were calculated by applying the 3D-RISM theory[57] to the wild type AcP and its two mutant conformations generated by the MD simulations. The 3D-RISM theory is an integral equation theory based on statistical mechanics, which enables us to obtain the three-dimensional distribution function $g_\gamma(\mathbf{r})$ of the site γ, oxygen or hydrogen, of water molecules around a given protein conformation. Thermodynamic properties such as the solvation free energy ($\Delta\mu$), the solvation entropy (Δs), and the solvation energy ($\Delta\varepsilon$) can be calculated from $g_\gamma(\mathbf{r})$.[58,59] The solvation energy $\Delta\varepsilon$ can further be decomposed into the average protein−water interaction energy ($\Delta\varepsilon^{uv}$) and the solvent reorganization energy ($\Delta\varepsilon^{vv}$).[60,61] Since the hydrophobicity can be defined in terms of the transfer free energy $\Delta\mu - \Delta\mu^{HC}$ of a solute from a nonpolar reference solvent such as a liquid hydrocarbon to an aqueous medium,[62] where $\Delta\mu^{HC}$ refers to the solvation free energy in the reference solvent, the increase/decrease in $\Delta\mu$ has a direct consequence to the increase/decrease in the hydrophobicity. The change in the solvation free energy and hydrophobicity upon mutation were studied via $\Delta\Delta\mu = \Delta\mu_{mu} - \Delta\mu_{wt}$ of the solvation free energies of the mutant ($\Delta\mu_{mu}$) and of the wild

type ($\Delta\mu_{wt}$). Details on the 3D-RISM calculation are provided in the Supporting Information.

The average protein−water interaction energy can be decomposed into contribution from each constituent atom of protein, $\Delta\varepsilon^{uv} = \Sigma_\alpha \Delta\varepsilon^{uv}_\alpha$ with $\Delta\varepsilon^{uv}_\alpha = 4\pi\rho\Sigma_\gamma \int r^2\, dr\, u_{\alpha\gamma}(r)\, g_{\alpha\gamma}(r)$. Here, ρ is the average number density of water, and $u_{\alpha\gamma}(r)$ and $g_{\alpha\gamma}(r)$ respectively refer to the interaction potential and the radial distribution between the atom α in protein and the atom γ in water. $g_{\alpha\gamma}(r)$ can be obtained from $g_\gamma(r)$ by performing the orientational average about the position of atom α[57] and is more convenient in quantifying the water distribution around a specific atom α in protein. The decomposition of $\Delta\varepsilon^{uv}$ enables us to investigate the relative importance of regions or residues in protein in determining the strength of the hydration. From the above expression for $\Delta\varepsilon^{uv}_\alpha$, it is clear that the magnitude of $\Delta\varepsilon^{uv}_\alpha$ is correlated with the first peak height in $g_{\alpha\gamma}(r)$, and this feature was exploited when we analyzed the hydrophobicity change upon mutation in terms of the solvation structure change.

We decomposed the value of the protein−water interaction energy change upon mutation, ($\Delta\Delta\varepsilon^{uv}$) instead of the solvation free energy change ($\Delta\Delta\mu$) because of the following reasons. First, the decomposition of $\Delta\Delta\mu$ is not straightforward.[63] Second, the solvation entropy change ($\Delta\Delta s$) is found to be relatively small, and $\Delta\Delta\mu$ is dominated by the energetic change ($\Delta\Delta\varepsilon$). Lastly, $\Delta\Delta\varepsilon^{uv}$ contributes to $\Delta\Delta\varepsilon$ much more than the solvent reorganization energy change ($\Delta\Delta\varepsilon^{vv}$). As a result, there is a good correlation between the changes in $\Delta\Delta\mu$ and $\Delta\Delta\varepsilon^{uv}$ (see Table 3).

The 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.[57,58] However, it is known that relative values of the solvation free energies are reasonably accurate,[57] and it is expected that the conclusion we draw on the solvation free energy change upon mutation does not significantly suffer from the limitation of the 3D-RISM theory. In fact, the hydrophobicity change upon mutation predicted by the 3D-RISM theory was found to be in accord with the experimental observation.

■ ASSOCIATED CONTENT

Supporting Information. Details of the 3D-RISM calculation, additional tables (secondary structure compositions for whole proteins, average helical compositions, average hydrogen bond occupancies, and average β-sheet structure compositions), and a figure (protein−water radial distribution functions for side-chain atoms). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
sihyun@sookmyung.ac.kr

■ ACKNOWLEDGMENT

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2008-313-C00404) and by the SRC Research Center for Women's Diseases of Sookmyung Women's University. We also would like to acknowledge the support from the KISTI Supercomputing Center.

■ REFERENCES

(1) Monti, M.; di Bard, B. L. G.; Calloni, G.; Chiti, F.; Amoresano, A.; Ramponi, G.; Pucci, P. *J. Mol. Biol.* **2004**, *336*, 253.
(2) Bemporad, F.; Calloni, G.; Campioni, S.; Plakoutsi, G.; Taddei, N.; Chiti, F. *Acc. Chem. Res.* **2006**, *39*, 620.
(3) Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2006**, *75*, 333.
(4) Gregersen, N.; Bross, P.; Vang, S.; Christensen, J. H. *Annu. Rev. Genomics Hum. Genet.* **2006**, *7*, 103.
(5) Muñoz, V. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 395.
(6) Jahn, T. R.; Radford, S. E. *Arch. Biochem. Biophys.* **2008**, *469*, 100.
(7) Dobson, C. M. *Nature* **2003**, *426*, 884.
(8) Petkova, A. T.; Yau, W. M.; Tycko, R. *Biochemistry* **2006**, *45*, 498.
(9) Petkova, A. T.; Ishii, Y.; Balbach, J. J.; Antzutkin, O. N.; Leapman, R. D.; Delaglio, F.; Tycko, R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16742.
(10) McParland, V. J.; Kalverda, A. P.; Homans, S. W.; Radford, S. E. *Nat. Struct. Biol.* **2002**, *9*, 326.
(11) Lai, Z.; Colón, W.; Kelly, J. W. *Biochemistry* **1996**, *35*, 6470.
(12) McCutchen, S. L.; Lai, Z.; Miroy, G. J.; Kelly, J. W.; Colon, W. *Biochemistry* **1995**, *34*, 13527.
(13) Chiti, F.; Calamai, M.; Taddei, N.; Stefani, M.; Ramponi, G.; Dobson, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16419.
(14) Chiti, F.; Taddei, N.; Baroni, F.; Capanni, C.; Stefani, M.; Ramponi, G.; Dobson, C. M. *Nat. Struct. Biol.* **2002**, *9*, 137.
(15) Chiti, F.; Webster, P.; Taddei, N.; Clark, A.; Stefani, M.; Ramponi, G.; Dobson, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3590.
(16) Ramazzotti, M.; Parrini, C.; Stefani, M.; Manao, G.; Degl'Innocenti, D. *FEBS Lett.* **2006**, *280*, 6763.
(17) Chiti, F.; Stefani, M.; Taddei, N.; Ramponi, G.; Dobson, C. M. *Nature* **2003**, *424*, 805.
(18) Chiti, F.; Taddei, N.; Bucciantini, M.; White, P.; Ramponi, G.; Dobson, C. M. *EMBO J.* **2000**, *19*, 1441.
(19) Chiti, F.; Taddei, N.; Stefani, M.; Dobson, C. M.; Ramponi, G. *Protein Sci.* **2001**, *10*, 879.
(20) Chiti, F.; Taddei, N.; White, P. M.; Bucciantini, M.; Magherini, F.; Stefani, M.; Dobson, C. M. *Nat. Struct. Biol.* **1999**, *6*, 1005.
(21) Chiti, F.; Nuland, N. A. J. v.; Taddei, N.; Magherini, F.; Stefani, M.; Ramponi, G.; Dobson, C. M. *Biochemistry* **1998**, *37*, 1447.
(22) Krantz, B. A.; Srivastava, A. K.; Nauli, S.; Baker, D.; Sauer, R. T.; Sosnik, T. R. *Nat. Struct. Biol.* **2002**, *9*, 458.
(23) Taddei, N.; Capanni, C.; Chiti, F.; Stefani, M.; Dobson, C. M.; Ramponi, G. *J. Biol. Chem.* **2001**, *276*, 37149.
(24) Taddei, N.; Chiti, F.; Fiaschi, T.; Bucciantini, M.; Capanni, C.; Stefani, M.; Serrano, L.; Dobson, C. M.; Ramponi, G. *J. Mol. Biol.* **2000**, *300*, 633.
(25) Bemporad, F.; Capanni, C.; Calamai, M.; Tutino, M. L.; Stefani, M.; Chiti, F. *Biochemistry* **2004**, *43*, 9116.
(26) Taddei, N.; Chiti, F.; Paoli, P.; Fiaschi, T.; Bucciantini, M.; Stefani, M.; Dobson, C. M.; Ramponi, G. *Biochemistry* **1999**, *38*, 2135.
(27) Calamai, M.; Chiti, F.; Dobson, C. M. *Biophys. J.* **2005**, *89*, 4201.
(28) Chiti, F.; Taddei, N.; Giannoni, E.; Nuland, N. A. J. v.; Ramponi, G.; Dobson, C. M. *J. Biol. Chem.* **1999**, *274*, 20151.
(29) Plakoutsi, G.; Bemporad, F.; Monti, M.; Pagnozzi, D.; Pucci, P.; Chiti, F. *Structure* **2006**, *14*, 993.
(30) Taddei, N.; Chiti, F.; Magherini, F.; Stefani, M.; Thunnissen, M. M. G. M.; Nordlund, P.; Ramponi, G. *Biochemistry* **1997**, *36*, 7217.
(31) Taddei, N.; Modesti, A.; Bucciantini, M.; Stefani, M.; Magherini, F.; Vecchi, M.; Raugei, G.; Ramponi, G. *FEBS Lett.* **1995**, *362*, 175.
(32) Calloni, G.; Taddei, N.; Plaxco, K. W.; Ramponi, G.; Stefani, M.; Chiti, F. *J. Mol. Biol.* **2003**, *330*, 577.
(33) Chiti, F.; Magherini, F.; Taddei, N.; Ilardi, C.; Stefani, M.; Bucciantini, M.; Dobson, C. M.; Ramponi, G. *Protein Eng.* **1998**, *11*, 557.
(34) Paoli, P.; Taddei, N.; Fiaschi, T.; Veggi, D.; Camici, G.; Manao, G.; Raugei, G.; Chiti, F.; Ramponi, G. *Arch. Biochem. Biophys.* **1999**, *363*, 349.
(35) Soldi, G.; Plakoutsi, G.; Taddei, N.; Chiti, F. *J. Med. Chem.* **2006**, *49*, 6057.
(36) Taddei, N.; Stefani, M.; Magherini, F.; Chiti, F.; Modesti, A.; Raugei, G.; Ramponi, G. *Biochemistry* **1996**, *35*, 7077.
(37) Bemporad, F.; Taddei, N.; Stefani, M.; Chiti, F. *Protein Sci.* **2006**, *15*, 862.
(38) Calamai, M.; Taddei, N.; Stefani, M.; Ramponi, G.; Chiti, F. *Biochemistry* **2003**, *42*, 15078.

(39) Parrini, C.; Taddei, N.; Ramazzotti, M.; Degl'Innocenti, D.; Ramponi, G.; Dobson, C. M.; Chiti, F. *Structure* **2005**, *13*, 1143.
(40) Soldi, G.; Bemporad, F.; Chiti, F. *J. Am. Chem. Soc.* **2008**, *130*, 4295.
(41) Calamai, M.; Tartaglia, G. G.; Vendruscolo, M.; Chiti, F.; Dobson, C. M. *J. Mol. Biol.* **2009**, *387*, 965.
(42) Flock, D.; Daidone, I.; Nola, A. D. *Biopolymers* **2004**, *75*, 491.
(43) Pastore, A.; Saudek, V.; Ramponi, G.; Williams, R. J. P. *J. Mol. Biol.* **1992**, *224*, 427.
(44) Feig, M.; Karanicolas, J.; Brooks, C. L. *J. Mol. Graph. Model* **2004**, *22*, 377.
(45) Millhauser, G. L. *Biochemistry* **1995**, *34*, 3873.
(46) Case, D. A. *AMBER 9*; University of California: San Francisco, 2006.
(47) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
(48) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714.
(49) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(50) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(51) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(52) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(53) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577.
(54) Kyte, J.; Doolittle, R. F. *J. Mol. Biol.* **1982**, *157*, 105.
(55) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.
(56) DeLano, W. L. *The PyMOL Molecular Graphics System and Users Manual*; DeLano Scientific: San Carlos, CA, 2002.
(57) Hirata, F. *Molecular Theory of Solvation*; Kluwer: Dordrecht, 2003.
(58) Imai, T.; Harano, Y.; Kinoshita, M.; Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2006**, *125*, 024911.
(59) Imai, T.; Harano, Y.; Kinoshita, M.; Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2007**, *126*, 225102.
(60) Yu, H.; Karplus, M. *J. Chem. Phys.* **1988**, *89*, 2366.
(61) Yu, H.; Roux, B.; Karplus, M. *J. Chem. Phys.* **1990**, *92*, 5020.
(62) Ben-Naim, A. *Hydrophobic Interactions*; Plenum Press: New York, 1980.
(63) Mark, A. E.; van Gunsteren, W. F. *J. Mol. Biol.* **1994**, *240*, 167.