

Interaction of Arginine with Proteins and the Mechanism by Which It Inhibits Aggregation

Diwakar Shukla and Bernhardt L. Trout*

Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States

Received: September 2, 2010

Aqueous arginine solutions are used extensively for inhibiting protein aggregation. There are several theories proposed to explain the effect of arginine on protein stability, but the exact mechanism is still not clear. To understand the mechanism of protein cosolvent interaction, the intraprotein, protein–solvent, and intrasolvent interactions have to be understood. Molecular dynamics simulations of aqueous arginine solutions were carried out for experimentally accessible concentrations and temperature ranges to study the structure of the solution and its energetic properties and obtain insight into the mechanism by which arginine inhibits protein aggregation. Simulations of proteins (α -chymotrypsinogen A and melittin) were performed. Structurally, the most striking feature of the aqueous arginine solutions is the self-association of arginine molecules. Arginine shows a marked tendency to form clusters with head to tail hydrogen bonding. Due to the presence of the three charged groups, there are several possible configurations in which arginine molecules interact. At relatively high concentrations, these arginine clusters associate with other clusters and monomeric arginine molecules to form large clusters. The hydrogen bonds between arginine molecules were found to be stronger than those between arginine and water, which makes the process of self-association enthalpically favorable. From the simulation of the proteins in aqueous arginine solution, arginine is found to interact with the aromatic and charged side chains of surface residues. A probable mechanism of the effect of arginine on protein stability consistent with our findings is proposed. In particular, arginine interacts with aromatic and charged residues due to cation– π interaction and salt-bridge formation, respectively, to stabilize the partially unfolded intermediates. The self-interaction of arginine leads to the formation of clusters which, due to their size, crowd out the protein–protein interaction. The mechanisms proposed in the literature are analyzed on the basis of the simulation results reported in this paper and recent experimental data.

Introduction

Protein solutions degrade via a number of routes, including aggregation. Solution additives or cosolvents can be used to modify the solution behavior of the proteins and affect their solubility and stability. Arginine in particular is widely used to suppress protein aggregation.^{2–4} Various experimental observations have been made about the effects of arginine on protein–protein association reactions. It has been shown that arginine reduces attractive protein–protein interactions, as indicated by a shift from a negative to positive osmotic second virial coefficient measured by light scattering experiments,⁵ increases the solubility of unfolded species of hen egg white lysozyme,⁴ and decreases the rate of association of unfolded and partially folded intermediates on the folding pathway during refolding as measured by native protein activity and size-exclusion chromatography.⁶ Attempts have been made to develop cosolvents that are similar to arginine, but more effective at inhibiting aggregation.^{7,8} Therefore, a clear picture of the mechanism by which arginine inhibits protein aggregation is desirable.

Many theories have been proposed to explain the effect of cosolvents on proteins.^{1,9–14} In 1888, Hofmeister⁹ ordered cations and anions according to their ability to stabilize protein solutions. The origin of this series has been attributed to the structural changes that ions cause in the water network. However, this rationale only applies to the simple monatomic ions. For

complex molecular cosolvents such as arginine, several functional groups in the molecule produce complex solvent structuring patterns that depend on the nature of the functional groups and their relative positions. Arginine is an aggregation suppressor, but it increases the surface tension of water on addition, which is similar to the behavior of the protein denaturants, such as guanidinium hydrochloride (GdmHCl).¹¹ Experimental studies to determine the solubility of amino acids in aqueous arginine reveal that arginine, like GdmHCl, interacts favorably with all the amino acid side chains, with both compounds showing strong interaction with aromatic residues.¹¹ However, the interaction of arginine with the protein surface is limited due to its large size as compared to Gdm. It is speculated that the limited binding of arginine plays a major role in its ability to suppress aggregation. Preferential interaction is a measure of the excess number of cosolvent molecules around the protein surface as compared to the bulk.¹² Preferential interaction measurements carried out by dialysis/densimetry or vapor pressure osmometry highlight the importance of protein–cosolvent interactions as compared to solvent interactions. Recently, Schneider and Trout¹ reported an interesting trend in the interaction of arginine with proteins as a function of concentration and protein size. They observed that, as the concentration increases, arginine becomes increasingly excluded from the protein surface. They suggested that the possible reason for this nonlinear exclusion of arginine from the protein surface is that the protein surface becomes saturated with arginine as the concentration is increased.

* To whom correspondence should be addressed. E-mail: trout@mit.edu.
Phone: (617) 258-5021. Fax: (617) 253-2272.

The current understanding of the mechanism by which arginine inhibits aggregation is limited. There are three proposed hypotheses for the effect of arginine on the stability of protein solutions.

(1) Tsumoto et al.⁵ suggested that interactions between the guanidine group of arginine and tryptophan side chains on the protein surface may be responsible for suppression of protein aggregation. The solubility of tryptophan is significantly increased in GdmHCl solutions due to the cation–π interactions.¹¹

(2) Baynes and Trout¹³ proposed the “gap effect” theory,¹³ in which, depending on the cosolvent size and preferential interaction with the protein surface as compared to water, cosolvents may accelerate, decelerate, or have no effect on the rate of aggregation. During protein association reactions, a complex of two or more protein molecules is formed, and the solvation shell at the interface must be removed. The cosolvent and water molecules can solvate the protein equally in the associated (P_2) and dissociated ($P + P$) states. When the separation between the protein molecules is such that large cosolvent molecules are excluded for steric reasons, but the water is still allowed to solvate the gap, this situation leads to an increase in the free energy of protein–protein encounter complexes and an increase in the barrier to association. In its pure sense, the gap effect is a purely kinetic effect because it only affects the free energy of the encounter complexes. The gap effect theory differs from the Asakura–Oosawa (AO) theory,^{15,16} although both the theories describe the effect of the osmotic pressure of the medium on aggregating objects. The AO theory applies only to noninteracting and completely excluded cosolvents, whereas the gap effect theory is not restricted to cosolvents which are completely excluded. The AO theory predicts a net attraction between aggregating molecules because the excluded volume is larger in the dissociated state than in the associated state for completely excluded crowders. Thus, the AO theory describes an equilibrium effect. This is in stark contrast to the gap effect theory, which is a kinetic effect. On the basis of the gap effect theory, the concept of “neutral crowders” was proposed. Neutral crowders do not affect the free energy of unfolding and are, hence, “neutral”, but due to their larger size as compared to water molecules, they “crowd out” the protein–protein interactions. The authors proposed that arginine can be a neutral crowder as the magnitude of the observed aggregation suppression matches the theoretical prediction for that of a neutral crowder of a size of arginine.

(3) Arginine molecules stack in a head-to-tail fashion, exposing their methylene groups as a hydrophobic column along one crystallographic axis.¹⁷ Das et al.¹⁴ proposed that the arginine clusters in solution also display a hydrophobic surface by a similar alignment of arginine’s three methylene groups. This hydrophobic surface can interact with the hydrophobic residues on the protein surface, which could inhibit protein aggregation. They showed that arginine increases the solubility of pyrene in water and modulates the hydrophobic interaction of Alzheimer’s amyloid β by binding to its surface.

On the basis of the above survey, it can be seen that there is no agreed upon mechanistic picture of the arginine-induced aggregation suppression. Recently, aqueous urea¹⁸ and Gdm¹⁹ solutions have been investigated to obtain a comprehensive picture of the interaction of water and cosolvent molecules in solution, which has implications for the mechanism by which these cosolvents denature proteins. In view of the importance of these interactions in aqueous cosolvent solutions, it is of interest to examine more closely the structure of and interactions in aqueous arginine solutions, the interactions between the

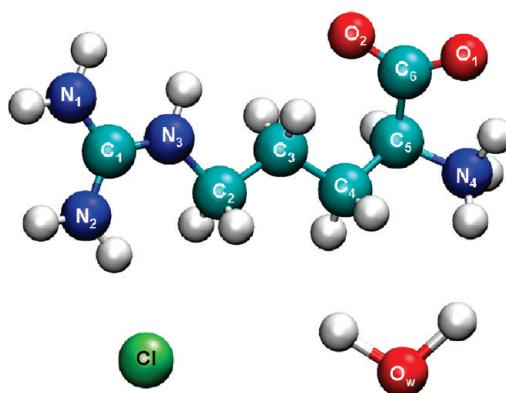


Figure 1. Labeling scheme for the atoms in arginine. Also shown are the water molecule and chloride ion. In the subsequent figures, all nitrogen atoms are shown in blue, oxygen atoms in red, carbon atoms in cyan, chloride atoms in green, and hydrogen atoms in white.

arginine and the protein surface residues, and the implications of these interactions for the mechanism by which arginine inhibits aggregation. There have been no reported simulation studies of aqueous arginine solutions to the best of our knowledge. Here, a molecular dynamics study of this binary system is described. The main focus is on the structural and energetic (hydrogen-bonding) properties of aqueous arginine solutions. The temperature and arginine concentration dependence is covered here for the full experimentally accessible range. Simulations of proteins in aqueous arginine solutions were also performed to understand the interaction of arginine with the protein surface.

Methods

Simulation Setup. All simulations were performed using the NAMD²⁰ package with the CHARMM22²¹ force field. The TIP3P²² water model was used. The pK_a values for the C-terminal, the N-terminal, and the side chain in an arginine molecule are 1.8, 9.0, and 12.5, respectively.²³ The N-terminal and the side chain are protonated whereas the C-terminal is deprotonated in the pH range 1.8–9.0. This pH range is of interest as proteins are observed to be highly unstable at low and high pH.^{24,25} Therefore, force field parameters for arginine were taken from the CHARMM force field with the protonated N-terminal, the side chain, and the deprotonated C-terminal. The structure of an arginine molecule and the atom labels used in the paper are shown in Figure 1. The parameters for the N- and C-terminals were taken from the CTER and NTER parameters available in CHARMM. Mass densities were compared to the experimental density data,¹ and the error was found to be ~1%. All simulations were performed in the NpT ensemble with periodic boundary conditions, and full electrostatics were computed using the particle mesh Ewald (PME) method,²⁶ with a grid spacing of 1 Å or less. The pressure was maintained at 1 atm using the Langevin piston method,²⁷ with a piston period of 200 fs, a damping time constant of 100 fs, and a piston temperature of 298 K. An integration step of 1 fs was used. The initial size of the periodic rectangular box was set to 50 Å³ in all of the simulations. To set up the simulation systems for various arginine hydrochloride concentrations ranging from 0.25 to 2.75 m , arginine and chloride ions were randomly placed within the simulation box (while assuring no overlap), and subsequently overlapping water molecules were removed. The system was then equilibrated for 1 ns at constant pressure and temperature. All 11 systems described in Table 1

TABLE 1: Setup of the Simulation System^a

| number of water molecules | number of ArgHCl molecules | molality |
|---------------------------|----------------------------|----------|
| 4000 | 18 | 0.25 |
| 3778 | 34 | 0.50 |
| 3556 | 48 | 0.75 |
| 3333 | 60 | 1.00 |
| 3111 | 70 | 1.25 |
| 2963 | 80 | 1.50 |
| 2825 | 89 | 1.75 |
| 2694 | 97 | 2.00 |
| 2568 | 104 | 2.25 |
| 2444 | 110 | 2.50 |
| 2343 | 116 | 2.75 |

^a The numbers of ArgHCl and water molecules in the system and molal concentrations are listed. The saturation limit at 298 K is at a molal concentration of 2.81.¹

were simulated for 10 ns each. Simulations with 2.50 *m* concentration were also performed at five different temperatures ranging from 278 to 358 K for 10 ns. We have also analyzed the interaction of arginine (1 *m* concentration, 298 K, 1 atm, pH 4.75) with the surface groups of the protein α -chymotrypsinogen A. A 50 ns simulation of α -chymotrypsinogen A was performed in a 75 Å box containing 188 arginine and 10411 water molecules. To study the interaction of arginine with hydrophobic residues which are not normally accessible in a folded protein, a 25 ns simulation of a helical peptide, melittin, in a 2.75 *m* ArgHCl solution was performed. Melittin is a 26 amino acid peptide with 15 hydrophobic residues and has a +6 charge at pH 7.0.^{28,29} The simulation box with sides of 60 Å contained 1 melittin, 204 arginines, 210 chloride ions, and 4121 water molecules.

Calculation of the Preferential Interaction Parameter. MD simulations of protein in mixed solvents are used to measure the preferential binding parameter. The preferential binding parameter, Γ_{23} , describes how the cosolvent concentration changes when protein is added to the solution to keep the chemical potential of the cosolvent constant. The method of calculating preferential interaction parameters, based on a statistical mechanical method applied to the all-atom model with no adjustable parameter, has been described in the literature.^{30,31} This approach is used to calculate the number of molecules bound to the protein without a priori information about any binding sites on the protein and yields a detailed description of interactions between proteins and cosolvents. The variation of the concentration as a function of the distance from the protein surface is used to calculate Γ_{23} as a function of the distance from the protein until it approaches a constant value. Preferential interaction coefficients (Γ_{23}) are calculated using the following equation:^{30–34}

$$\Gamma_{23} = \rho_3(\infty) \int_0^\infty (g_3 - g_1) dV \quad (1)$$

where subscripts 1, 2, and 3 stand for water, protein, and cosolvent, respectively, g_i is the radial distribution of component *i*, and $\rho_3(\infty)$ denotes the density of the cosolvent in the bulk region away from the protein surface. The integral extends from $r = 0$ (protein surface) to $r = \infty$ (bulk region). The procedure described above can be used if more than one species is present in the solution to calculate the excess number of molecules of a particular species within the local domain of the protein.

Clustering. Self-aggregation or clustering of arginine molecules was quantified in terms of the reduction of the total

solvent-accessible surface area of arginine with respect to a monomeric arginine molecule. The solvent-accessible area was estimated using standard CHARMM commands with a probe sphere of 1.4 Å radius. The solvent-accessible surface area was used as a measure, since minimization of the exposed surface area is one of the main driving forces for clustering. Furthermore, it has been reported that this measure is more sensitive than other measures such as Kirkwood–Buff integrals.¹⁸

Effect of Cosolvent on the Protein Association Reaction.

To quantify the effect of cosolvents on protein–protein association reactions, we study a model system for the association of proteins. The reaction of two parallel, planar plates and the reaction of two spheres are used as the extreme cases of the geometry of two associating proteins. The two spheres, each 20 Å in radius, are used as a model for the association reaction involving two spherical proteins. The distance between the center of the proteins is defined as the reaction coordinate, *x*. For planes, the distance between the faces of the planes acts as the reaction coordinate. The surface area of the plates (400 Å²) is selected to make the change in the protein solvent-accessible area of the reaction the same as for the case of two spheres. The thermodynamic properties of these plates are obtained by calculating the property per unit surface area of a pair of infinite plates and then multiplying by the area. The free energy of this protein complex, $\mu_{2,0}(x)$, can be modeled as described previously:¹³

$$\text{spheres: } \mu_{2,0}(x) = \left(\frac{15}{x}\right)^6 - 8.21e^{-[(x - 20)/10]^2} + 2.12e^{-[(x - 40)/10]^2} \quad (2)$$

$$\text{planes: } \mu_{2,0}(x) = \left(\frac{1}{x}\right)^6 - 8.51e^{-(x - 1.5)^2} + 2.02e^{-[(x - 4)/2]^2} \quad (3)$$

The above equations place a 2 kcal/mol free energy barrier for the association between the two proteins. The dimer state is chosen to be 8 kcal/mol more stable than the monomer state. The free energy in the presence of cosolvent is computed by adding the free energy in the absence of additive with the transfer free energy. The transfer free energy, $\Delta\mu_2^T$, is computed via^{6,13}

$$\Delta\mu_2^T = - \int_0^{m_3} \left(\frac{\partial\mu_3}{\partial m_3} \right)_{T,P,m_2} \Gamma_{23} dm_3 \quad (4)$$

The expression for Γ_{23} from eq 1 is substituted into the above equation:

$$\Delta\mu_2^T = - \int_0^{m_3} c_3 \left(\frac{\partial\mu_3}{\partial m_3} \right)_{T,P,m_2} \left(\int (g_3 - g_1) dV \right) dm_3 \quad (5)$$

The above equation can be simplified using the following assumptions: (1) The water cosolvent interactions are ideal, which makes the derivative of μ_3 with respect to m_3 equal to RT/m_3 . (2) The concentration of cosolvent is low so that molal and molar concentrations are equal. (3) The radial distribution function (RDF) of the cosolvent and water with respect to the protein is represented using a three-parameter Exp-6 potential. This function was fitted to the radial distribution functions

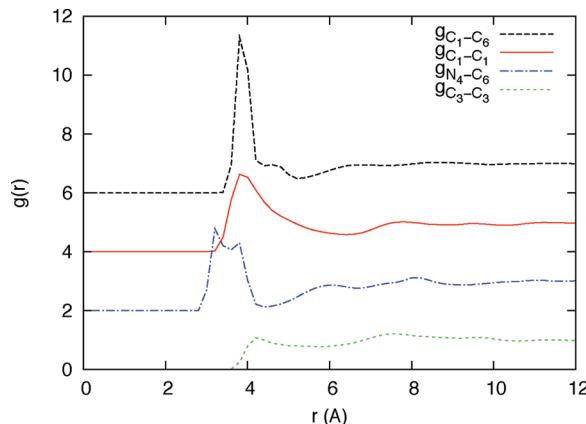


Figure 2. Radial distribution functions between Gdm and carboxylate carbon ($g_{C_1-C_6}$), the Gdm carbon atoms ($g_{C_1-C_1}$), and the N-terminal and the C-terminal ($g_{N_4-C_6}$) of the arginine molecule at a concentration of 2.75 m.

obtained using molecular dynamics simulations.^{6,13} Applying the above assumptions, eq 5 is simplified to

$$\Delta\mu_2^{\text{tr}} = -RTc_3 \int (e^{-\langle U_{23} \rangle / RT} - e^{-\langle U_{21} \rangle / RT}) dV(x) \quad (6)$$

where R is the gas constant, T is the absolute temperature, c_3 is the cosolvent concentration, $\langle U_{23} \rangle$ is the protein–cosolvent potential of mean force, and $\langle U_{21} \rangle$ is the protein–water potential of mean force. The integral is over the system volume, which is a function of the reaction coordinate. The relative change in the association rate can be calculated using

$$\frac{k_a}{k_0} = e^{-\Delta\Delta\mu_2^{\text{tr}}/k_b T} \quad (7)$$

where k_a is the rate constant in the presence of cosolvent, k_0 is the rate constant in the absence of cosolvent, $\Delta\Delta\mu_2^{\text{tr}}$ is the change in the activation free energy, k_b is Boltzmann's constant, and T is the absolute temperature.

Results and Discussion

Structural Properties. Spatial density distributions, including the translational and the rotational degrees of freedom, would capture details of the solution order. Such an analysis is possible for planar and rigid molecules such as urea and Gdm.^{18,19} A conformational analysis of arginine in the gas phase shows a large number of local minima with comparable energies, due to the many easily rotatable single bonds.^{35,36} Due to the flexibility of arginine, spatial density distributions were found to be smeared in the vicinity of the molecule. Therefore, site–site RDFs were used to characterize the geometry and short-range order of arginine–water solutions.

Figure 2 shows the RDF between the middle carbon atoms ($g_{C_3-C_3}$) of the arginine molecules at 2.75 m concentration. The first distinct peak is at 4.2 Å, and the second peak is at 7.6 Å, with a shoulder at 9.6 Å. The first peak corresponds to the two adjacent arginine molecules bonded together by two or more hydrogen bonds in a head-to-tail fashion due to the strong interaction between the Gdm and carboxylate groups as shown in the left panel of Figure 3. The second peak corresponds to adjacent arginine molecules with their N-terminal and carboxylate groups hydrogen bonded (middle panel of Figure 3) and

Gdm groups stacked on top of each other as shown in the right panel of Figure 3. Gdm ion stacking in aqueous solutions has been reported by Mason et al.¹⁹ on the basis of the MD simulation and neutron diffraction experiment of 3 m GdmHCl solution. Vondrášek et al.³⁷ have reported that the cavitation energy, dispersion interactions, and reduction in electrostatic repulsion due to the flat geometry and nonhomogeneous charge distribution are the main factors responsible for the favorable association of like-charged Gdm–Gdm pairs. The distance between Gdm groups is observed to be around 4 Å. This distance is larger than the van der Waals contact distance, but it is not sufficiently large for water molecules to occupy the gap. The Gdm–Gdm pair can also be stabilized by the tendency of the chloride ions to occupy equatorial positions in the Gdm plane. The presence of the chloride ion diminishes the inter-ion repulsion due to like-charged Gdm groups.³⁸ The distance between Gdm carbons in this case is observed to be around 7.5 Å, which is the same as the value reported for hetero-ion complexes in GdmHCl solutions.¹⁹ Some arginine–arginine dimers were found to be in fully extended conformations in which the carboxylate and N-terminal of the adjacent arginine molecules were held together by two hydrogen bonds as shown in Figure 3. The shoulder in the second peak corresponds to such dimers. The dimers formed due to hydrogen bonding between Gdm and carboxylate groups of adjacent arginine molecules also contribute to the same peak. It was observed that such tail-to-tail or head-to-tail hydrogen bonding increases with an increase in concentration as arginine dimers can interact with other dimers or free arginine molecules. The guanidinium carbon–carbon ($g_{C_1-C_1}$) and the Gdm carbon–carboxylate carbon ($g_{C_1-C_6}$) RDFs are shown in Figure 2. The C_1-C_1 RDF shows two peaks around 4 and 7.5 Å. These peaks correspond to the two different types of Gdm–Gdm stacking discussed above. The C_1-C_6 RDF shows a peak at 3.8 Å, with a shoulder at 4.2 Å. The maximum corresponds to the hydrogen bonding between Gdm and carboxylate groups of adjacent arginine molecules. The shoulder corresponds to the carboxylate group hydrogen bonded to the N-terminal group of the adjacent arginine molecule. These RDFs also give a measure of the population of two different types of dimers shown in Figure 3. It can be seen from the figures and the coordination number obtained after integration of these curves (not shown here) that the population of the first dimer which has two arginine molecules bonded with two to four hydrogen bonds is higher than that of the other dimers. The N_4-C_6 RDF exhibits a peak around 3.5 Å, which corresponds to the hydrogen bond between the N-terminal and C-terminal of adjacent arginine molecules. The C-terminal carbon within the same molecule is not included in the calculation of the RDF.

The radial distribution functions for the N-terminal nitrogen and the chloride ion are shown in Figure 4a. The prominent first peak in this function is at 3.2 Å, with a significant shoulder at 5.4 Å. The maximum corresponds to the chloride ion occupying the position between hydrogen atoms of the N-terminal nitrogen (see Figure 5). The shoulder corresponds to the chloride ions coordinated to a water molecule hydrogen bonded to the N-terminal nitrogen. There is a competition among the chloride, water, and arginine atoms to hydrogen bond with the N-terminal nitrogen. The peak height decreases with increasing concentration due to the hydrogen-bonding interaction between the N-terminal and the carboxylate group of arginine. The chloride–chloride radial distribution function (Figure 4b) shows three peaks. The first maximum is at 4.8 Å, which corresponds to the two chloride ions mutually coordinated to

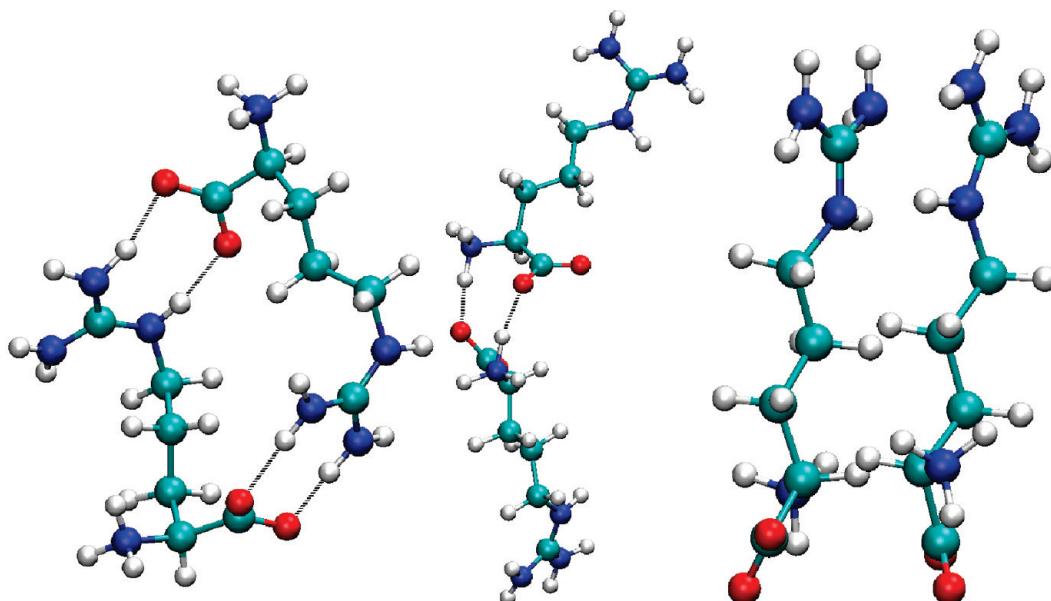


Figure 3. Snapshots of arginine molecules from the MD simulation illustrating possible configurations in which the two arginine molecules can interact to form a dimer: a dimer formed due to hydrogen bonding between the Gdm and carboxylate groups (left); a dimer formed due to hydrogen bonding between the N-terminal and the C-terminal of adjacent arginine molecules (center); a dimer with the stacked Gdm groups (right).

an intervening water molecule. The second peak falls at 7.4 Å and corresponds to the two chloride ions occupying adjacent positions in the first hydration shell around the Gdm group and the N-terminal (see Figure 5). The third peak at 9.8 Å corresponds to the two chloride ions each coordinated with the Gdm group and the N-terminal. In agreement with the previous MD simulations of chloride salts,^{19,39} no significant direct chloride ion pairing was observed under these conditions. The Gdm carbon–chloride RDF (Figure 4c) shows two prominent peaks. The first peak is at 3.8 Å, with a shoulder at 4.5 Å. The first peak corresponds to the chloride ion occupying the position in the first hydration shell between the hydrogen atoms of the Gdm, hydrogen bonding to each at an angle of 140°. The shoulder corresponds to the chloride ions making a single linear hydrogen bond to a single hydrogen of the Gdm (see Figure 5). The second peak in the RDF is at a distance of 6.6 Å and corresponds to the chloride hydrogen bonded to hydrogen atoms of the Gdm with a water molecule acting as a bridge. The first peak decreases in height and the second peak height increases with an increase in concentration due to the hydrogen-bonding interaction between arginine molecules which restricts the direct interaction of chloride ion.

RDFs between water–oxygen and arginine atoms are shown in Figure 6. Water molecules can be seen to be strongly coordinated with the three charged groups in the molecule as shown by the prominent peaks in the RDFs (N_4-O_w , C_6-O_w , C_1-O_w). However, the middle carbon atom shows no prominent peak due to the hydrophobic nature of the three methylene groups in the center of the molecule. The water structure around arginine is not significantly perturbed due to the self-interaction of arginine. The water molecules form linear hydrogen bonds with the Gdm group and are, therefore, constrained to remain in the plane of the Gdm group. When two arginine molecules form a cluster with Gdm–Gdm stacking, the equatorial positions are still unoccupied. Furthermore, these *n*-mers are held together by bonding with only a part of the group. For example, in a dimer with head-to-tail bonding, only one of the nitrogen atoms in the Gdm group is involved. The water–water radial distribution functions (not shown here) also support this observation. It is seen that arginine only slightly perturbs the water structure.

The population of the first solvation shell, which can be used as a measure of the short-range order in the water structure does not change significantly with an increase in the ArgHCl concentration. This observation indicates that the indirect effect of arginine on proteins through its effect on the structure of water is not significant.

Hydrogen Bonds. The strength and the number of hydrogen bonds in the aqueous arginine solutions give further insight into the nature of the interactions. Figure 7 shows the number of hydrogen bonds per molecule for different arginine concentrations. The number of arginine–arginine hydrogen bonds per arginine molecule increases with increasing concentration. The number of arginine–water bonds decreases with increasing concentration. The total number of bonds per arginine molecule decreases slightly as fewer arginine–arginine bonds are formed than arginine–water bonds are lost. For water, a similar behavior is observed in terms of the total number of hydrogen bonds per water molecule. This marginal decrease in the number of hydrogen bonds per arginine or water molecule again confirms the observation that arginine perturbs the water structure only slightly.

Figure 8 shows the hydrogen bond energies for all donor–acceptor combinations. The strongest hydrogen bond is formed between the carboxylate oxygen (acceptor) and the N-terminal nitrogen (donor) pair with an average energy of 33 kJ/mol (298 K). The strongest hydrogen bonds between arginine and water were those formed between the water oxygen (acceptor) and the carboxylate oxygen (donor) with a mean energy of ~30 kJ/mol (298 K). The water–water hydrogen bonds (20.5 kJ/mol) are weak as compared to arginine–arginine or arginine–water hydrogen bonds. The self-aggregation of arginine is therefore enthalpically favorable. From the entropic viewpoint, a single arginine molecule would free about 10 water molecules. The number of water molecules replaced by an arginine is calculated by placing an arginine molecule in a water box and counting the number of water molecules within 2.2 Å of the arginine molecule. Therefore, the self-interaction is promoted as the water molecules solvating the individual arginine molecules are released, which increases the translational and rotational entropy. The lifetime of individual arginine–arginine

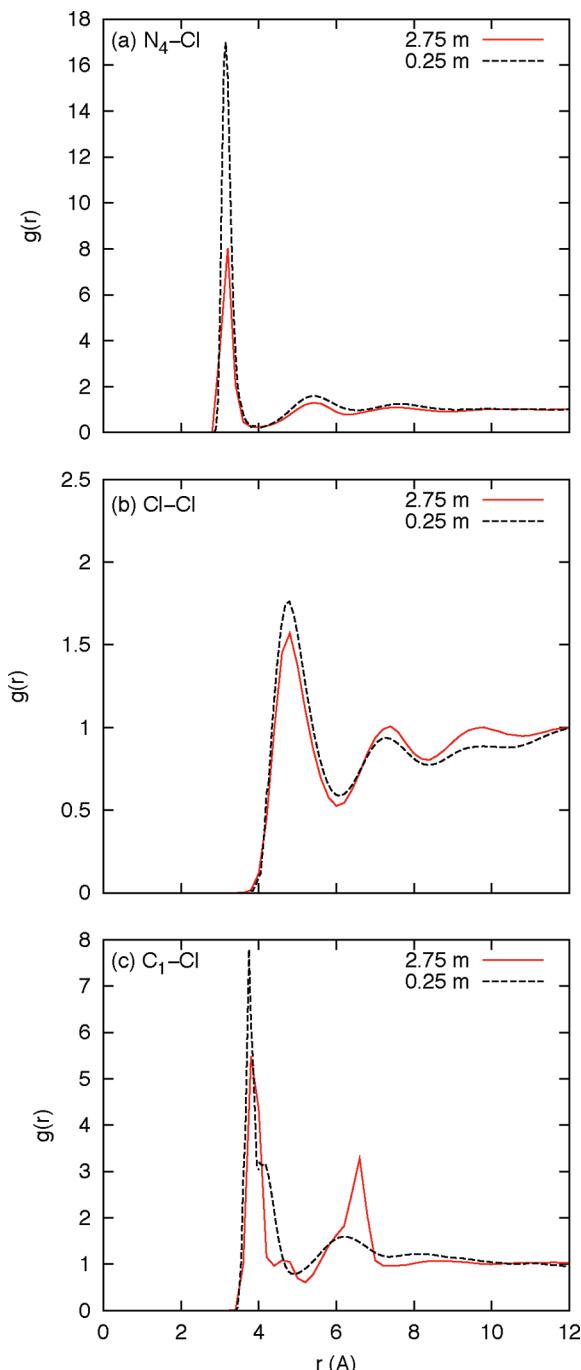


Figure 4. Radial distribution functions between chloride ions and (a) the N-terminal nitrogen, (b) the chloride ion, and (c) the Gdm carbon.

hydrogen bonds is calculated using the procedure of Rapaport.⁴⁰ The average bond lifetime is found to vary between 7 and 9 ps, depending on the solution concentration. However, the time for which two arginine molecules in a cluster stay connected is found to lie between 90 and 110 ps. The clusters are held together by multiple hydrogen bonds, the individual hydrogen bonds breaking after every 7–9 ps, but the cluster breaks after a much longer time.

Clustering. Figure 9 shows the mean surface area per arginine molecule and the percentage loss in surface area compared to that of a single arginine molecule in a water box as a function of concentration. The maximum percentage loss in area is 50%. The solubility of ArgHCl at 298 K is 2.81 m . For the highest concentration (2.75 m) studied here, the percentage loss in

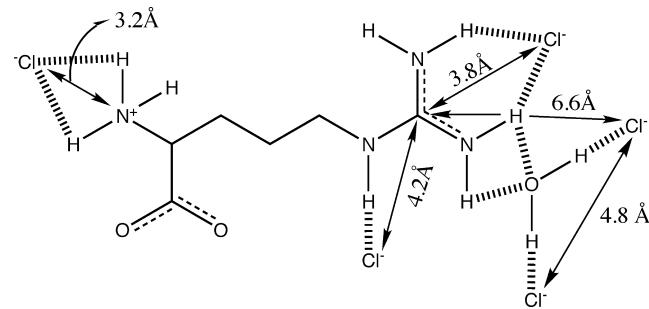


Figure 5. Arginine–chloride ion pairing found in the simulations. The distances reported in the figure correspond to the peaks in the radial distribution functions shown in Figure 4. Hatched lines denote a hydrogen bond, and dotted lines denote a partial double bond.

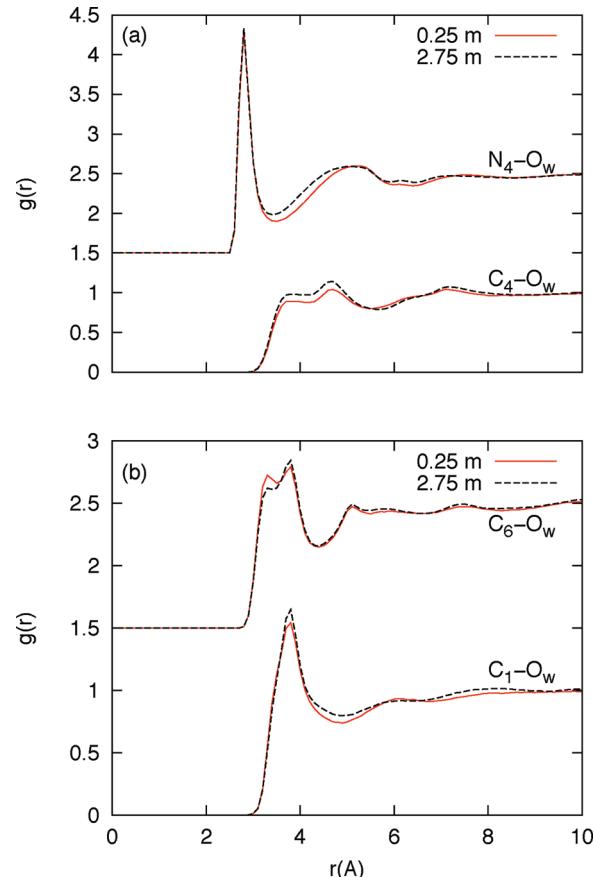


Figure 6. Arginine–water, site–site radial distribution functions. RDFs for the N-terminal nitrogen and carboxylate carbon (C_6) are offset 1.5 units along the ordinate.

surface area is $\sim 45\%$. With increasing temperature, the extent of clustering (for 2.50 m) is observed to decrease, due to a higher solubility and an increase in the overall volume of the box. However, this approach overestimates the extent of clustering due to the fact that, in the absence of any interaction, random contacts between arginine molecules would reduce this solvent-accessible surface area. Stumpe and Grubmüller¹⁸ have proposed a procedure to distinguish this from real self-aggregation for the clustering in aqueous urea solutions. Such area loss due to random contacts is expected to be small in the case of arginine due to the low diffusivity and strong interaction between arginine molecules.

The populations of clusters of various sizes were calculated where the criteria for a molecule to be in a cluster are (a) that it is connected by one or more hydrogen bonds and (b) the

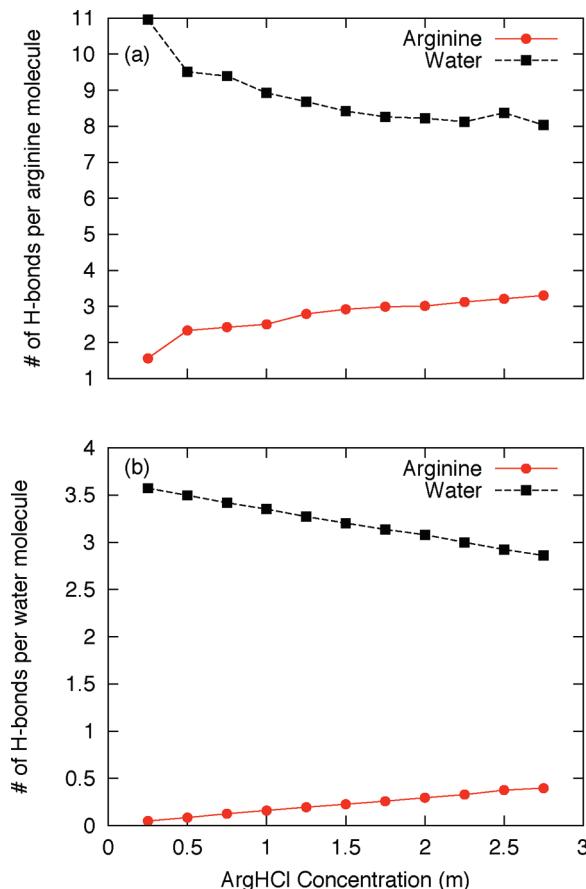


Figure 7. Average number of hydrogen bonds (a) per arginine molecule and (b) per water molecule.

specified atoms are within a given distance range, which is defined by the position of the first minimum between the atoms in the appropriate $g(r)$. The second constraint applies only to the clusters with Gdm–Gdm stacking. The cluster size is calculated by counting all the molecules that are connected to one other molecule in a cluster. The number of contacts between the groups in the interacting arginine molecules is calculated using the same criteria. The number of Gdm–carboxylate, Gdm–Gdm, and N-terminal–C-terminal contacts per arginine molecule as a function of concentration is shown in Figure 10. Gdm–Gdm and Gdm–carboxylate are the dominant modes of contact between arginine molecules in solution. A single arginine molecule can form up to four contacts. The number of contacts is observed to increase with increasing concentration. The total number of contacts at 2.75 m concentration is 2.3. The population of clusters as a function of the size and the probability to find an arginine molecule in a cluster of a given size is shown in Figure 11. It is found that, for the 0.25 m system (Figure 11b), 45% of the arginine molecules are in dimers or higher order n -mers. On average, only 13% of the arginine was found to be in n -mers larger than a dimer. If a number of clusters of various sizes are considered, only 25% of the clusters are dimers or higher order n -mers. For the 2.75 m system, the number of arginine molecules present as monomers is as low as 18%. n -mers from monomers to decamers account for ~60% of the total arginine molecules. From the plot of the number of clusters of various sizes, it can be seen that there are few very large clusters present in the solution at 2.75 m. For low concentrations, the solution is dominated by the arginine monomers and dimers, but as the concentration increases n -mers of larger sizes are formed. These n -mers are nanoscale clusters held together by

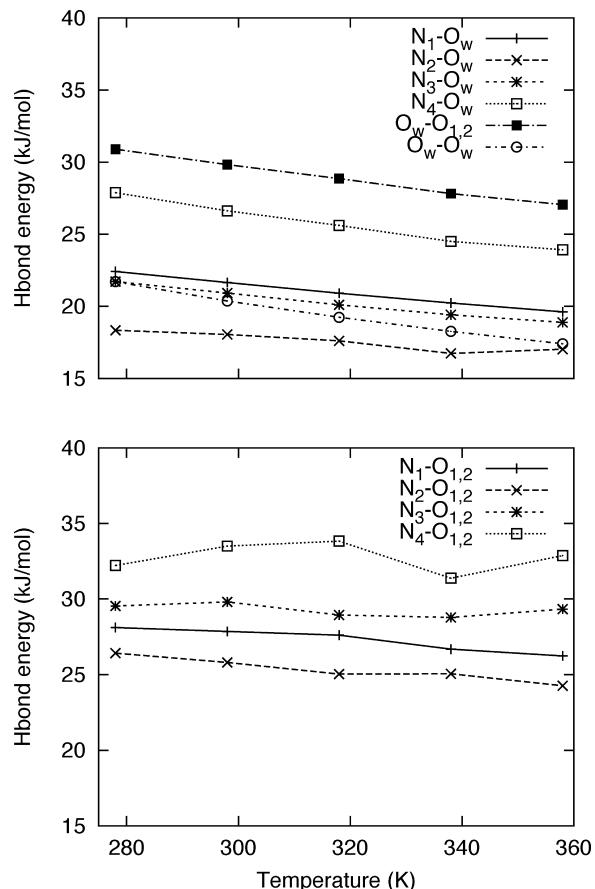


Figure 8. Mean energy per hydrogen bond for all donor–acceptor combinations at 2.50 m concentration.

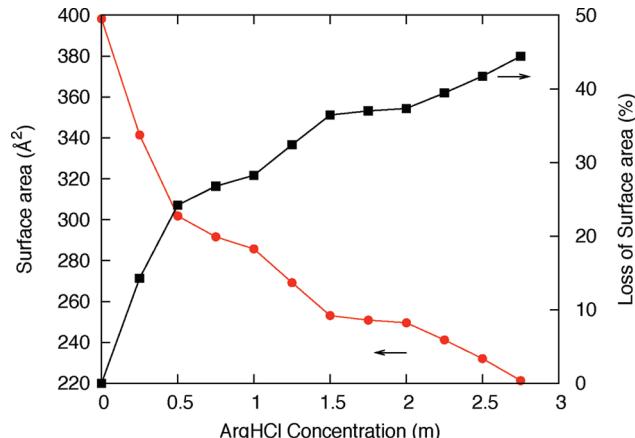


Figure 9. Solvent-accessible surface area as a function of the ArgHCl molality normalized by the total number of arginine molecules in the system (left). Percentage loss in surface area as a function of the ArgHCl molality (right).

hydrogen bonds, hetero-ion pairing, and Gdm–Gdm stacking. These clusters are significantly larger in size as compared to a single arginine molecule, and as such, they are expected to play a significant role in stabilizing the protein molecules in mixed solvents.

Protein–Arginine Interaction. To understand the role that interactions between the protein surface and the arginine play in the mechanism by which arginine inhibits aggregation, a molecular dynamics simulation of the protein α -chymotrypsinogen A (α -CgnA) in a 1 m aqueous arginine solution was performed. Arginine is found to interact with the protein surface mainly via the guanidinium group. Radial distribution functions

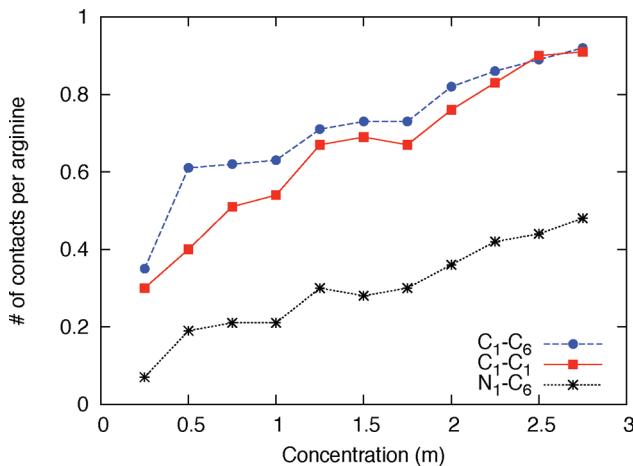


Figure 10. Number of contacts per arginine molecule between the Gdm and carboxylate groups (C_1-C_6), Gdm groups (C_1-C_1), and the N- and C-termini of arginine molecules in solution.

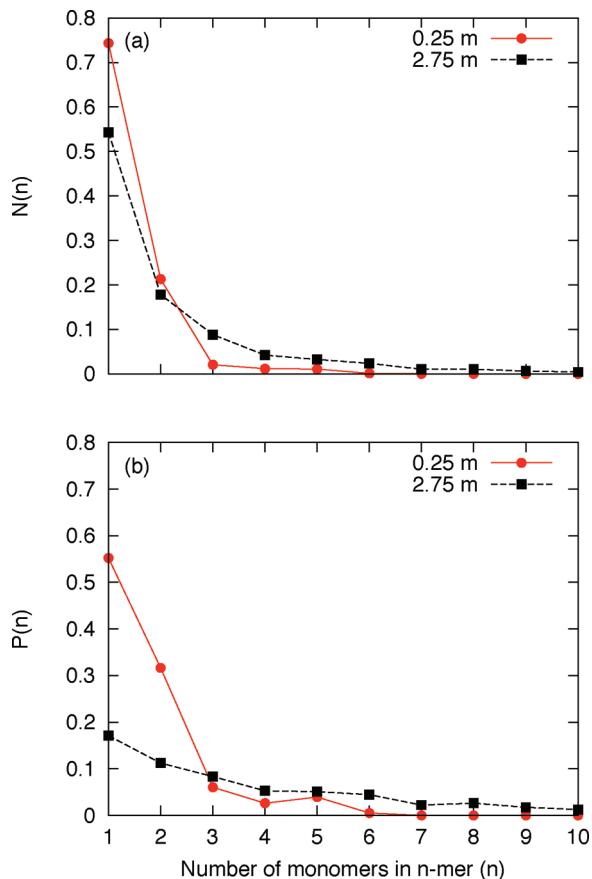


Figure 11. Distribution of the population of arginine clusters: (a) normalized population of arginine clusters of various sizes; (b) probability of finding an arginine molecule in a cluster of a particular size.

of the guanidinium and the carboxylate carbon in the arginine molecule around the protein α -CgnA are shown in Figure 12. It can be seen that arginine is preferentially oriented with respect to the protein surface. The preferential binding behavior of arginine with each type of amino acid and the backbone in α -CgnA is shown in Figure 13. The contact coefficient is defined as the ratio of the arginine concentration around a particular amino acid and the bulk concentration. Stumpe and Grubmüller⁴² have reported contact coefficient values for interaction of urea with glycine-capped tripeptides. The normalization with the bulk

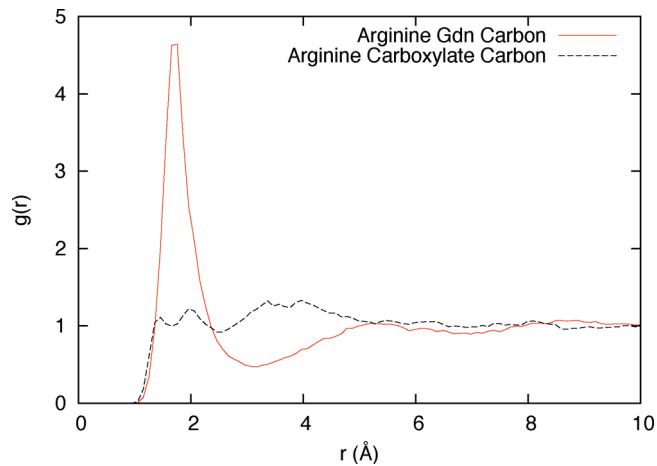


Figure 12. Radial distribution functions of the guanidinium carbon (C_1) and the carboxylate carbon (C_6) in the arginine molecule around the protein α -CgnA.

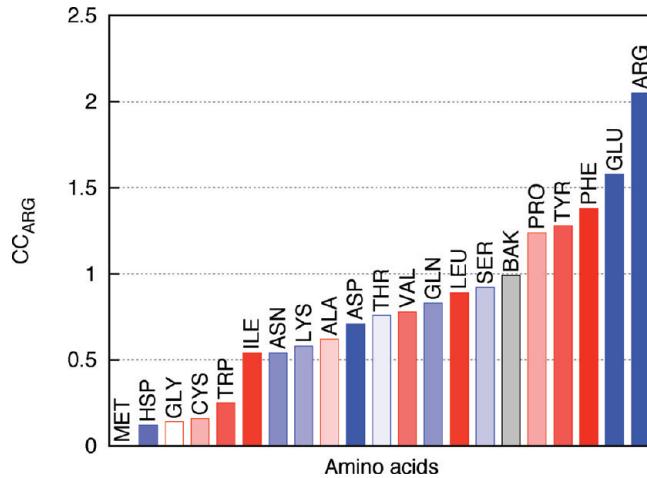


Figure 13. Contact coefficient, CC_{ARG} , for each amino acid in α -CgnA. The colors represent the hydrophobicity of amino acids: red, hydrophobic; blue, hydrophilic. The intensity of the bars depends on the normalized hydrophobicity values⁴¹ of each amino acid. The label BAK denotes the protein backbone. It can be seen that there is no trend between the CC values and the hydrophobicity of the residues.

concentration gives a better idea of the tendency of the cosolvent to interact preferentially with the surface of a particular amino acid. The number of arginine molecules and chloride ions bound to each amino acid on the protein surface is included in the Supporting Information. Arginine is observed to interact strongly with the aromatic and charged residues on the protein surface. The charged groups in arginine can interact with both the positively and negatively charged amino acid side chains. Arginine is observed to interact with charged residues via hydrogen bonding, similar to the behavior of arginine in aqueous solution. The average concentration around the arginine and glutamic acid residues on the protein surface is found to be 2.1 and 1.6 times the bulk concentration, respectively. The Gdm group interacts with aromatic side chain residues via cation–π interactions. Figure 14a shows an arginine molecule interacting with a tryptophan residue on the protein surface. The average surface concentration of arginine around tyrosine and phenylalanine side chains in α -CgnA is found to be 1.4 and 1.3 times the bulk concentration, respectively. The backbone atoms can hydrogen bond with the arginine molecules. The average number of arginine molecules bound to the backbone atoms is 5.8, which corresponds to a contact coefficient value of 1.0. In our

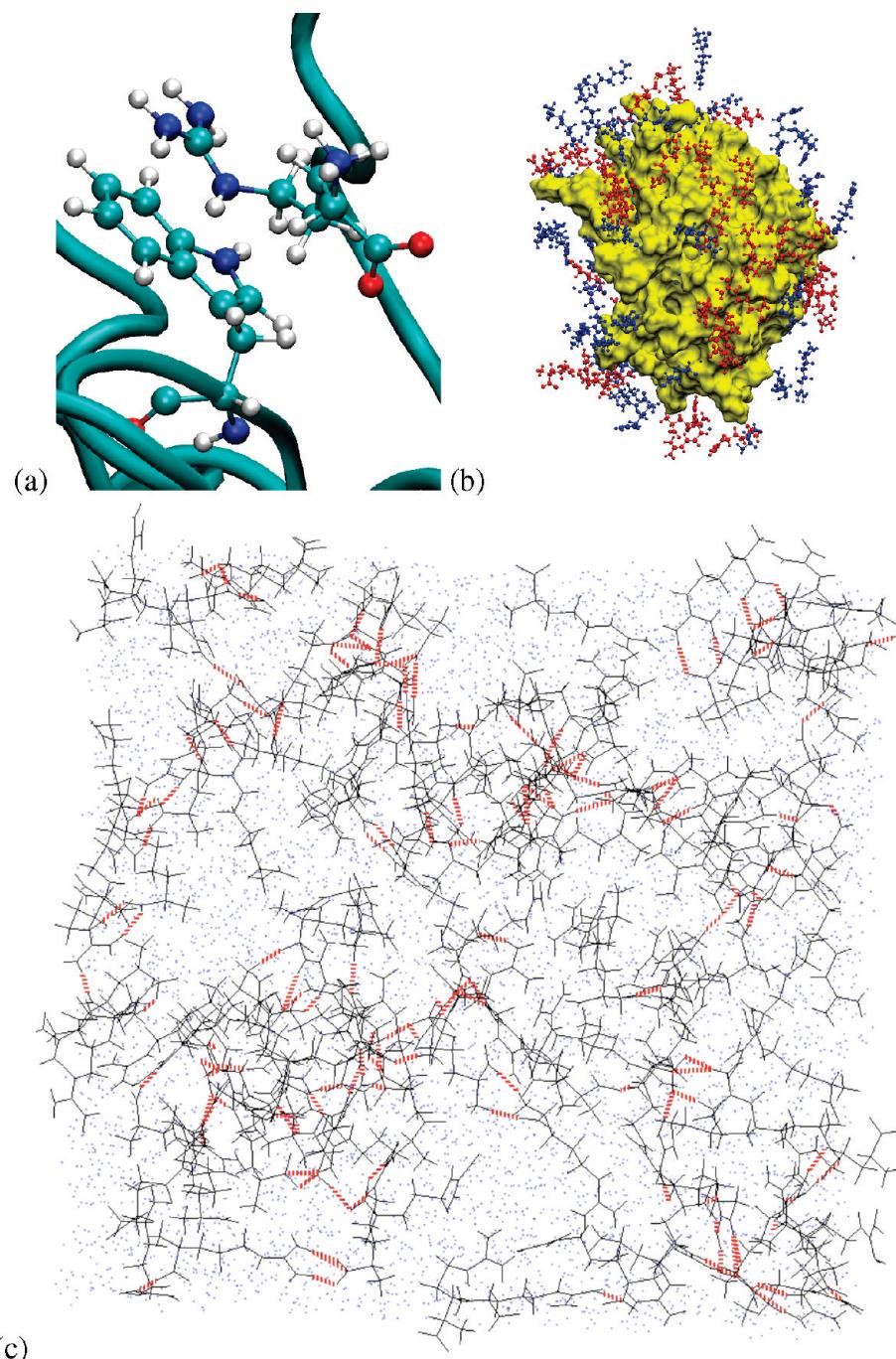


Figure 14. (a) Cation- π interaction between the tryptophan residue on the surface of α -CgnA and the Gdm group of an arginine molecule in solution. (b) Snapshot of arginine molecules present in the local domain (6.0 \AA) of α -chymotrypsinogen A. Arginine molecules present as dimers are shown in red. (c) Snapshot of the MD simulation box containing aqueous arginine solutions at 1 M concentration (right). Water molecules are shown as blue dots for clarity. Hydrogen bonds between arginine molecules are shown as hatched red lines.

simulations, interactions between hydrophobic residues and arginine are observed. The methylene groups in the arginine molecule can interact with the hydrophobic residues on the protein surface. If we consider all the nonaromatic hydrophobic residues (Val, Ile, Leu, Met, Cys, Ala, Pro, and Gly), the average local concentration around these residues is 0.5 times the bulk concentration, which indicates that arginine does not interact strongly with the hydrophobic residues. In the native structure of the protein, the hydrophobic residues are not accessible to the solvent. Methionine, histidine, and tryptophan have low contact coefficient values due to the limited exposure of these residues to the solvent. Water molecules can interact with these partially exposed residues, but large arginine molecules cannot. There-

fore, a simulation of the polypeptide with exposed hydrophobic groups (melittin) was performed to assess the interactions of arginine with the hydrophobic regions. The total number of arginine molecules coordinated with the protein is 9.0. The number of arginine molecules coordinated with the hydrophobic residues (Val, Ile, Leu, Pro, and Gly) is 2.9, corresponding to a local concentration of 0.54 relative to the bulk concentration. Although the number of aromatic and charged amino acids in melittin is less than the number of hydrophobic amino acids, these amino acids account for 6.1 molecules out of 9 molecules coordinated with the protein. Furthermore, arginine clusters were not found to interact with the hydrophobic residues by stacking the methylene groups to form a surface as suggested by Das

et al.¹⁴ Therefore, simulation of a polypeptide with exposed hydrophobic groups also provides similar results for the interaction of arginine with protein surface residues.

The preferential interaction coefficient, Γ_{23} , measures the excess number of cosolvent molecules within the local domain of the protein as compared to that in bulk solution. Recently, Schneider and Trout¹ have reported preferential binding coefficient data for α -chymotrypsinogen A. The experimental preferential binding coefficient value at 1 m concentration is -8.7 ± 2.9 . The computed value of -8.3 ± 0.7 is in good agreement with the experimental data. Arginine solutions are comprised of monomers and higher order n -mers. It would be interesting to look at the preferential binding data for the monomers ($\Gamma_{23,m}$) and higher order n -mers. $\Gamma_{23,m}$ can be calculated by counting the number of monomers in the local and bulk domains. Similar calculations can also be performed for the higher order n -mers. However, the standard deviation of these coefficients would be large due to the small population of these n -mers. Therefore, longer simulations would be required to calculate these coefficients. However, the difference between the overall Γ_{23} and $\Gamma_{23,m}$ would give an average value for all n -mers ($n \geq 2$). For α -CgnA, $\Gamma_{23,m}$ was found to be -5.08 . The value for $\Gamma_{23,m}$ is higher than the overall Γ_{23} . This implies that the concentration of monomers within the local domain is higher than the n -mer ($n \geq 2$) concentration. This is expected because the local concentration around the protein is less than the bulk concentration and the self-interaction between the arginine molecules increases with increasing concentration.

Mechanism. The three proposed mechanisms by which arginine inhibits protein aggregation should be reassessed on the basis of the molecular level insights obtained from the simulations performed in this study.

Neutral crowders are cosolvents which have a preferential binding coefficient value of zero (or slightly positive) and are larger than water in size. The preferential binding data of arginine from Schneider and Trout¹ clearly show that arginine is highly excluded from the protein surface and therefore cannot be classified as neutral. However, arginine can still be classified as a crowder. Arginine molecules which accumulate on the protein surface are responsible for crowding out the protein–protein interactions. According to the gap effect theory, the rate of association decreases with increasing additive size and preferential interaction coefficient. Arginine has a negative preferential binding coefficient, which would increase the association rate of proteins. On the basis of the simulations of aqueous arginine solutions, it is concluded that arginine forms clusters in solution and these clusters have sizes larger than that of the arginine molecule. This effective increase in the size of arginine due to self-interaction counters the change in the association rate due to exclusion of arginine from the protein surface. Therefore, the observed aggregation suppression would be equivalent to a neutral crowder of the size of an arginine molecule. The aggregation suppression effect of arginine on the proteins can be illustrated clearly using the association model described in the Methods. The relative change in the association rate as a function of the additive size for two different values of Γ_{23} is shown in Figure 15. Arginine has a radius of gyration of 3.6 Å. If arginine is assumed to be neutral ($\Gamma_{23} = 0$), then the relative change in the association rate is 0.57. If the arginine is excluded from the protein surface ($\Gamma_{23} < 0$), then the ratio k_a/k_0 is greater than 0.57. However, the size of arginine is most likely enhanced due to the self-interaction, which compensates for the negative contribution due to the preferential interaction. Experimental or theoretical preferential interaction coefficients

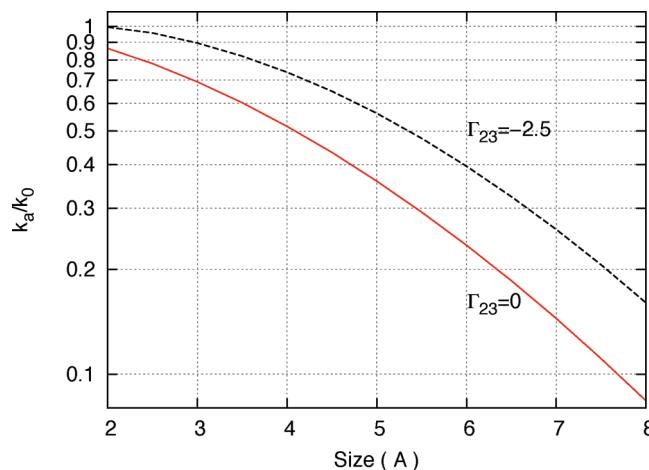


Figure 15. Relative change in the association rate for 20 Å spherical proteins caused by a 0.5 m cosolvent solution as a function of the cosolvent size.

for arginine hydrochloride are reported for four proteins.^{1,31} The values at 0.5 m ArgHCl concentration lie in the range -4 to -1 . Therefore, a value of -2.5 is chosen as the preferential interaction coefficient for the model protein. For $\Gamma_{23} = -2.5$, the value of k_a/k_0 is 0.8. From the $\Gamma_{23} = -2.5$ plot in Figure 15, it can be seen that the effective size of the arginine should be about 4.8 Å for k_a/k_0 to be equal to that of the neutral crowder ($\Gamma_{23} = 0$) with the size of an arginine molecule. On the basis of the loss of the surface area per arginine molecule, the population of monomers and higher n -mers can be calculated. At 0.5 m , the percentage of monomers is 68%. Therefore, the effective size of the arginine in the solution is calculated to be 4.77 Å, which is the same as the effective size of the arginine required to compensate for the exclusion of arginine from the protein surface. The effective size is calculated on the basis of the population of monomers and dimers in solution. At low concentrations, the solution predominately contains monomers and dimers. The size of the dimer is assumed to be equal to the size of the two arginine molecules separated by the distance equal to the location of the first minimum in the RDF between arginine molecules in solution. Arginine is preferentially excluded from the protein surface. Therefore, the local concentration around the protein is less than the bulk concentration. The extent of clustering of arginine increases with increasing concentration, which implies that $\Gamma_{23,m}$ will always be higher than Γ_{23} . The $\Gamma_{23,m}$ value calculated for the simulation of α -CgnA in 1 m arginine solution supports this point. From the viewpoint of aggregation suppression, monomers have a small size but a higher Γ and n -mers have a low Γ but are large in size. This observation shows the trade-off between Γ and the size of the cosolvent in terms of their effect on the association rate.

The relative change in the association rate as a function of the arginine concentration for the two extreme cases of spherical and planar proteins is shown in Figure 16. To calculate the rate of association in the presence of arginine, experimental preferential interaction data¹ for lysozyme in arginine solution and the effective size of the arginine molecule in solution are used. The relative change in the association rate is calculated using eq 7. Hirano et al.⁴³ have reported the aggregation rate constant for lysozyme with and without arginine. The experimental value at 600 mM is found to be 0.14, which is in agreement with our predictions of 0.11–0.30.

On the basis of the simulation results, it can be concluded that arginine interacts strongly with aromatic residues. The

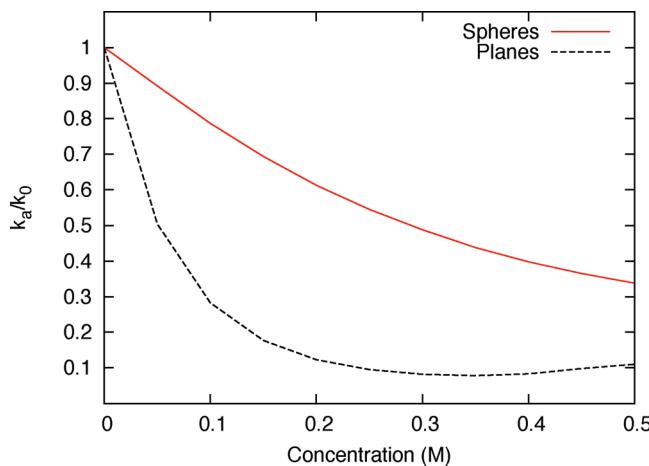


Figure 16. Relative change in the association rate for the spherical and planar proteins due to arginine solution as a function of the concentration.

number of aromatic amino acids in any protein is typically less than 10% of the total number of amino acids.⁴⁴ Therefore, the number of arginine molecules bound to the aromatic residues would be a small fraction of the total number of arginine molecules associated with the protein. If we consider a situation in which all the aromatic residues in α -CgnA are exposed, then the number of arginine molecules bound to them can be calculated on the basis of the number of arginine molecules coordinated per residue for a particular amino acid. The number of arginine molecules associated with the aromatic residues would be 12. However, if the same analysis is done for nonaromatic residues, the number of arginine molecules associated with the nonaromatic residues is found to be 77. The number of arginine molecules associated with the aromatic residues is a small fraction of the total arginine molecules associated with the protein. Therefore, the interactions between arginine and aromatic residues alone cannot account for the aggregative suppression behavior of arginine.

Simulation results and experimental light scattering data suggest that arginine forms molecular clusters in solutions.¹⁴ However, the hypothesis that these clusters present a large hydrophobic surface by the alignment of their methylene groups is not supported by the simulations. The simulation involving melittin which has an exposed hydrophobic surface does not show any such interactions. No large hydrophobic surfaces similar to the hydrophobic columns present in L-arginine crystals is observed in simulations of aqueous arginine solutions. A typical snapshot of the MD simulation box containing aqueous arginine solution with and without protein is shown in Figure 14b,c. It can be seen that n -mers ($n > 2$) of arginine are formed by hydrogen bonding between charged groups without any alignment of methylene groups to form a large hydrophobic surface. To support the hypothesis that the interactions between methylene groups in arginine clusters and the protein surface are responsible for the aggregation suppression, Das et al.¹⁴ measured the solubility of pyrene and performed ANS fluorescence emission intensity studies of aqueous arginine solutions. The solubility of pyrene was observed to increase in the presence of arginine. The solubility of pyrene increases with a decrease in the polarity of the solvent. The ANS fluorescence intensity also increased, and a blue shift in the maximum wavelength was observed. These changes in ANS fluorescence are observed when the fluorophore is in a hydrophobic environment. On the basis of these observations, the authors suggested that methylene groups in arginine are responsible for these effects on nonpolar

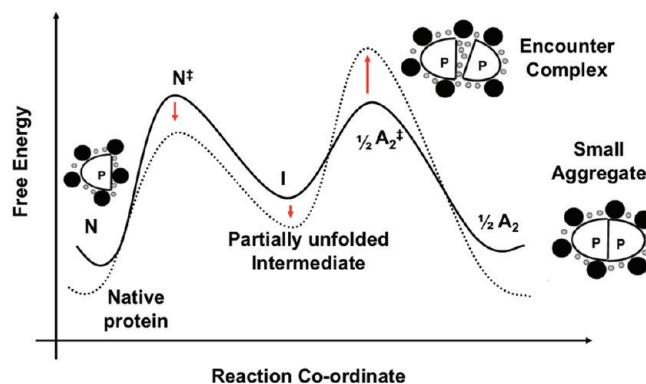


Figure 17. Effect of arginine on the free energy protein states along the refolding/aggregation reaction coordinate. The solid line represents the free energy in the absence of cosolvent and the dotted line that in the presence of cosolvent. Additive molecules are shown as large black circles and water molecules as small gray circles.

compounds. However, pyrene is a polycyclic aromatic hydrocarbon which forms a large flat aromatic system, and ANS is a charged molecule with phenyl and naphthalene rings. It has been shown that the Gdm group interacts strongly with aromatic residues due to cation– π interactions.^{11,45,46} Recently, Mason et al.⁴⁶ have reported that GdmCl significantly suppresses aromatic interactions between pyridine molecules but has no effect on the 2-propanol aggregation. Therefore, the Gdm group is more likely to be responsible for both the phenomena observed by the authors. Das et al. also showed that arginine increases the solubility of the $\text{A}\beta_{1-42}$ peptide and decreases fibrillar formation. They suggested that arginine clusters with aligned methylene groups mask the hydrophobic surface of $\text{A}\beta_{1-42}$. However, $\text{A}\beta_{1-42}$ has 4 aromatic and 12 charged residues which can interact with the arginine molecules via cation– π interaction and salt-bridge formation, respectively. The authors have shown the effect of arginine on various compounds, but there is no conclusive evidence for the presence of a hydrophobic surface formed due to stacked methylene groups or relating the observed effects to such a hydrophobic surface. Furthermore, the simulation results reported in this paper show that there is no such surface formed due to the stacking of methylene groups. Therefore, the antiaggregative property of arginine is most likely not due to hydrophobic interactions between the clusters and the protein surface.

The proposed effect of the arginine on the free energy of the protein along the refolding/aggregation reaction coordinate is shown in Figure 17. In the absence of arginine (solid line), the small aggregate (A_2) is formed from an unfolded or partially unfolded intermediate (I). These unfolded intermediates are formed due to the exposure of hydrophobic residues which are not exposed in the native state (N). In the presence of arginine (dotted line), the interaction of arginine with aromatic groups of the partially unfolded intermediates stabilizes these intermediates. Crowding around the macromolecule leads to an increase in the height of the barrier corresponding to the association of the partially unfolded proteins to form an aggregate. The free energy of the encounter complex increases due to the exclusion of the arginine from the gap between the associating macromolecules. The free energy of the native and the aggregated states also changes due to the presence of arginine. The interaction between the hydrophilic groups on the protein surface and arginine (enthalpic) lowers the free energy, whereas the exclusion of arginine from the protein surface (entropic) increases the free energy. However, arginine is not a highly excluded cosolvent (as compared to sucrose and other sugars).

Therefore, it is expected that the free energy of the native and the aggregated states will decrease in the presence of arginine.

It has been shown that the charge-based interactions between peptides are the dominant force for association in aqueous solution.⁴⁷ The self-interaction of arginine is also due to the hydrogen bonding between the oppositely charged groups. This raises an interesting question: Why do other charged amino acids not show such strong self-interaction? The possible reason for such a behavior can be the absence of a Gdm-like side chain in other amino acids. The Gdm side chain can have strong, multiple interactions and forms a planar structure, which helps in Gdm–Gdm stacking and interaction with neighboring molecules in aqueous solution.¹⁴ Lysine has a positively charged side chain, but it does not form clusters in solution.⁴⁸ Another possible reason for such a behavior could be the presence of a highly flexible methylene chain which would significantly reduce the stability of the dimers. Recently, Schulund et al.^{35,36} have shown that molecular rigidity is critical for the stability of the dimers formed by arginine in the gas phase. They showed that 2-(guanidinocarbonyl)-1*H*-pyrrole-5-carboxylate forms more stable clusters than arginine due to the rigidity of the groups holding the terminal carboxylate and Gdm groups. Therefore, making the chain less flexible might lead to better inhibition of protein aggregation. On the basis of the mechanism proposed in this paper, any change in the cosolvent structure which promotes self-interaction, keeping the preferential interaction the same, would improve the aggregation inhibition ability of the cosolvent.

Conclusions

This study was aimed at understanding the structure and interactions in aqueous arginine solutions and the interaction of arginine with proteins and their implications for the mechanism by which arginine inhibits aggregation. The radial distribution functions obtained from the simulations suggested a tendency for the arginine molecules to self-associate. There are several ways in which arginine molecules can self-associate. The hydrogen bonding between the Gdm and carboxylate group and the stacking of Gdm groups in adjacent arginine molecules are the two important contacts between arginine molecules. The weak hydrogen bonds between arginine and water are replaced by stronger arginine–arginine hydrogen bonds on self-association. The total number of hydrogen bonds per water molecule remains almost constant, which shows that arginine substitutes well for water in the hydrogen bond network. The population of the first solvation shell (within 0.5 Å) around water molecules decreases by a small amount with increases in the arginine concentration. This shows that the water structure is not perturbed significantly by the presence of arginine molecules. From our results, it can be seen that the self-association of arginine plays an important role in its binding and inhibition of protein aggregation. These results also highlight the role of the carboxylate group in the arginine molecule. Guanidinium salts (e.g., GdmHCl) interact too strongly with the protein and thereby unfold the protein, which promotes aggregation. The presence of the carboxylate group in arginine limits the interaction of the Gdm group with the protein surface. Therefore, the interaction is strong enough to stabilize partially unfolded intermediates but not strong enough to unfold the protein. From the simulation of proteins in an aqueous arginine solution, the aromatic and charged residues were found to interact with arginine via cation–π interactions and hydrogen bonding, respectively. These results demonstrate that, to explain how cosolvents work, all the possible interactions between the components of a mixed

solvent have to be analyzed. The existing mechanisms for the effect of arginine on protein aggregation are analyzed on the basis of the information available from the simulation results and recent experimental preferential binding data of arginine. A detailed analysis of all the existing mechanisms shows that none of them by themselves are completely consistent with the simulation results. The mechanism proposed in this paper takes into account the intrasolvent and solvent–protein interactions. The crowding due to the arginine molecules within the local domain of the protein and the interactions due to the Gdm group are mainly responsible for the aggregation suppression by arginine. The crowding effect of arginine is enhanced due to its self-association.

Acknowledgment. We acknowledge funding support from the Singapore-MIT Alliance. We also thank Curtiss P. Schneider of the Department of Chemical Engineering at MIT for his careful reading of the manuscript and helpful suggestions.

Supporting Information Available: Table containing the number of arginine molecules bound to each amino acid in α-CgnA and melittin. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- (1) Schneider, C. P.; Trout, B. L. *J. Phys. Chem. B* **2009**, *113*, 2050–2058.
- (2) Arakawa, T.; Tsumoto, K. *Biochem. Biophys. Res. Commun.* **2003**, *304*, 148–152.
- (3) Shiraki, K.; Kudou, M.; Fujiwara, S.; Imanaka, T.; Takagi, M. *J. Biochem.* **2002**, *132*, 591.
- (4) Reddy, K. R. C.; Lilie, H.; Rudolph, R.; Lange, C. *Protein Sci.* **2005**, *14*, 929.
- (5) Tsumoto, K.; Umetsu, M.; Kumagai, I.; Ejima, D.; Philo, J. S.; Arakawa, T. *Biotechnol. Prog.* **2004**, *20*, 1301–1308.
- (6) Baynes, B. M.; Wang, D. I. C.; Trout, B. L. *Biochemistry* **2005**, *44*, 4919–4925.
- (7) Okanojo, M.; Shiraki, K.; Kudou, M.; Nishikori, S.; Takagi, M. *J. Biosci. Bioeng.* **2005**, *100*, 556–561.
- (8) Shiraki, K.; Kudou, M.; Nishikori, S.; Kitagawa, H.; Imanaka, T.; Takagi, M. *Eur. J. Biochem.* **2004**, *271*, 3242–3247.
- (9) Hofmeister, F. *Arch. Exp. Pathol. Pharmakol.* **1888**, *24*, 247.
- (10) Kirkwood, J. G.; Buff, F. P. *J. Chem. Phys.* **1951**, *19*, 774.
- (11) Arakawa, T.; Ejima, D.; Tsumoto, K.; Obeyama, N.; Tanaka, Y.; Kita, Y.; Timasheff, S. N. *Biophys. Chem.* **2007**, *127*, 1–8.
- (12) Timasheff, S. N. *Adv. Protein Chem.* **1998**, *51*, 355–432.
- (13) Baynes, B. M.; Trout, B. L. *Biophys. J.* **2004**, *87*, 1631–1639.
- (14) Das, U.; Hariprasad, G.; Ethayathulla, A. S.; Manral, P.; Das, T. K.; Pasha, S.; Mann, A.; Ganguli, M.; Verma, A. K.; Bhat, R.; Chandrayan, S. K.; Ahmed, S.; Sharma, S.; Kaur, P.; Singh, T. P.; Srinivasan, A. *PLoS ONE* **2007**, *2*, e1176.
- (15) Asakura, S.; Oosawa, F. *J. Chem. Phys.* **1954**, *22*, 1255–1256.
- (16) Asakura, S.; Oosawa, F. *J. Polym. Sci.* **1958**, *33*, 183–192.
- (17) Karle, I. L.; Karle, J. *Acta Crystallogr.* **1964**, *17*, 835–841.
- (18) Stumpe, M. C.; Grubmüller, H. *J. Phys. Chem. B* **2007**, *111*, 6220.
- (19) Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Saboungi, M. L.; Dempsey, C. E.; MacKerell, A. D.; Brady, J. W. *J. Am. Chem. Soc.* **2004**, *126*, 11462–11470.
- (20) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (21) Brooks, B. R.; Bruckolieri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (22) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- (23) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry, Part I: The Conformation of Biological Macromolecules*; W. H. Freeman: San Francisco, 1980.
- (24) Pfeil, W.; Privalov, P. L. *Biophys. Chem.* **1976**, *4*, 41–50.
- (25) Schaefer, M.; Sommer, M.; Karplus, M. *J. Phys. Chem. B* **1997**, *101*, 1663–1683.
- (26) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
- (27) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613.
- (28) Mason, P. E.; Brady, J. W.; Neilson, G. W.; Dempsey, C. E. *Biophys. J.* **2007**, *93*, L04.

- (29) Terwilliger, T. C.; D., E. *J. Biol. Chem.* **1982**, *257*, 6010–6015.
(30) Baynes, B. M.; Trout, B. L. *J. Phys. Chem. B* **2003**, *107*, 14058–14067.
(31) Shukla, D.; Shinde, C.; Trout, B. L. *J. Phys. Chem. B* **2009**, *113*, 12546–12554.
(32) Tang, K. E. S.; Bloomfield, V. A. *Biophys. J.* **2002**, *82*, 2876–2891.
(33) Vagenende, V.; Yap, M. G. S.; Trout, B. L. *J. Phys. Chem. B* **2009**, *113*, 11743–11753.
(34) Vagenende, V.; Yap, M. G. S.; Trout, B. L. *Biochemistry* **2009**, *48*, 11084–11096.
(35) Schlund, S.; Schmuck, C.; Engels, B. *Chem.—Eur. J.* **2007**, *13*, 6644.
(36) Schlund, S.; Muller, R.; Grabmann, C.; Engels, B. *J. Comput. Chem.* **2008**, *29*, 407.
(37) Vondrášek, J.; Mason, P. E.; Heyda, J.; Collins, K. D.; Jungwirth, P. *J. Phys. Chem. B* **2009**, *113*, 9041–9045.
(38) Angelini, T. E.; Liang, H.; Wriggers, W.; Wong, G. C. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 8634–8637.
(39) Guardia, E.; Rey, R.; Padró, J. A. *Chem. Phys.* **1991**, *155*, 187–195.
(40) Rapaport, D. *Mol. Phys.* **1983**, *50*, 1151–1162.
(41) Black, S.; Mould, D. *Anal. Biochem.* **1991**, *193*, 72–82.
(42) Stumpe, M. C.; Grubmüller, H. *J. Am. Chem. Soc.* **2007**, *129*, 16126–16131.
(43) Hirano, A.; Hamada, H.; Okubo, T.; Noguchi, T.; Higashibata, H.; Shiraki, K. *Protein J.* **2007**, *26*, 423–433.
(44) Barnes, M. R. *Bioinformatics for Geneticists: A Bioinformatics Primer for the Analysis of Genetic Data*; John Wiley and Sons Ltd.: West Sussex, England, 2007.
(45) Flocco, M. M.; Mowbray, S. L. *J. Mol. Biol.* **1994**, *235*, 709–17.
(46) Mason, P. E.; Dempsey, C. E.; Neilson, G. W.; Kline, S. R.; Brady, J. W. *J. Am. Chem. Soc.* **2009**, *131*, 16689–16696.
(47) McLain, S. E.; Soper, A. K.; Daidone, I.; Smith, J. C.; Watts, A. *Angew. Chem., Int. Ed.* **2008**, *47*, 9059–9062.
(48) Julian, R. R.; Beauchamp, J. L.; Goddard, W. A., III. *J. Phys. Chem. A* **2002**, *106*, 32–34.

JP108399G