Protein Science (2000), *9*:344–352. Cambridge University Press. Printed in the USA.

Proline inhibits aggregation during protein refolding

DHARMARAJ SAMUEL,[1] THALLAMPURANAM KRISHNASWAMY S. KUMAR,[1,2]
GOPAL GANESH,[1] GURUNATHAN JAYARAMAN,[1,3] PEY-WEN YANG,[1]
MEI-MING CHANG,[1] VISHWA DEO TRIVEDI,[1,2] SUE-LEIN WANG,[1]
KUO-CHU HWANG,[1] DING-KWO CHANG,[2] AND CHIN YU[1]

[1]Department of Chemistry, National Tsing Hua University, Hsinchu, Taiwan
[2]Institute of Chemistry, Academia Sinica, Taipei, Taiwan

(RECEIVED June 21, 1999; FINAL REVISION October 13, 1999; ACCEPTED November 3, 1999)

Abstract

The in vitro refolding of hen egg-white lysozyme is studied in the presence of various osmolytes. Proline is found to prevent aggregation during protein refolding. However, other osmolytes used in this study fail to exhibit a similar property. Experimental evidence suggests that proline inhibits protein aggregation by binding to folding intermediate(s) and trapping the folding intermediate(s) into enzymatically inactive, "aggregation-insensitive" state(s). However, elimination of proline from the refolded protein mixture results in significant recovery of the bacteriolytic activity. At higher concentrations (>1.5 M), proline is shown to form loose, higher-order molecular aggregate(s). The supramolecular assembly of proline is found to possess an amphipathic character. Formation of higher-order aggregates is believed to be crucial for proline to function as a protein folding aid. In addition to its role in osmoregulation under water stress conditions, the results of this study hint at the possibility of proline behaving as a protein folding chaperone.

Keywords: aggregation; proline; protein folding; supramolecular assembly

One of the important issues for both basic research and applied biotechnology is the problem of refolding denatured polypeptide chains in vitro (Goldberg, 1985; Goldberg et al., 1991). Fundamental interest in the protein folding problem arises because we do not yet clearly understand how a complex network of noncovalent interactions can specify an unique, compact conformation (Dill, 1990; Rozema & Gellman, 1996b). It is often observed that proteins tend to aggregate during the refolding reaction in vitro (Cleland & Wang, 1990; Georgia & DeBernardez, 1991; Cleland, 1993). A kinetic competition between interchain and intrachain interactions is proposed to occur during the refolding of proteins (Orsini et al., 1975; Orsini & Goldberg, 1978; Zetllmeissl et al., 1979; Goldberg et al., 1991). The interchain interactions are envisaged to predominate if the concentration of the refolding protein is higher and as a consequence the protein tends to aggregate (Orsini et al., 1975; Brems, 1988). Protein aggregation during in vitro refolding is contemplated to be mediated by the occurrence of "molten globule" like intermediates in the refolding pathway of proteins (Ptitsyn et al., 1990; Semisotnov et al., 1990; Kumar et al., 1995; Sivaraman et al., 1997a). In the "molten globule" state, the protein(s) show greater exposure of the nonpolar residues to the solvent, and this aspect promotes interchain hydrophobic interactions among the folding intermediates leading to aggregation (Ptitsyn, 1992; De Young et al., 1993; Sivaraman et al., 1997b). The problem of protein aggregation assumes practical importance because proteins overexpressed by genetically engineered cells are often obtained as insoluble aggregates termed as "inclusion bodies" (Marston, 1986; Mitraki & King, 1989; Wetzel, 1994; Thomas et al., 1995; Kumar et al., 1996). Recently, several experimental strategies have been devised to overcome the problem of aggregation (Tandon & Horowitz, 1986; Zardeneta & Horowitz, 1992; Karuppiah & Sharma, 1995; Rozema & Gellman, 1995, 1996a; Wetlaufer & Xie, 1995). However, to date, no general panacea exists to alleviate the problem of protein aggregation.

The production and accumulation of organic osmolytes are widespread among halophilic and halotolerant plants and microorganisms (Yancey et al., 1982; Hochachka & Somero, 1984; Wang & Bolen, 1996). Organic osmolytes are often accumulated in the cytoplasm up to concentrations well above 2 mol/kg in water (Galinski, 1993). The most common osmolytes, which tend to accumulate in large concentrations, are polyols, glycine, betaine, proline, and ectoines such as 1,4,5,6-tetrahydro-2-methyl-4-pyrimidine carboxylic acid (Schoffenials & Gilles, 1970; Wegmann, 1986; Wohlfarth et al., 1990). These solutes are often termed as compatible solutes since they do not interfere with the cell metabolism (Brown & Simpson, 1972; Borrowitzka, 1985). Under conditions of osmotic equilibrium, reduced water activity is believed to be a crucial factor for stability of enzymes and other cellular components (Lin & Timasheff, 1994). The osmolytes are proposed

Reprint requests to: Chin Yu, National Tsing Hua University, Department of Chemistry, Hsinchu, Taiwan.
[3]Present address: Institute of Life Sciences, National Tsing Hua University, Hsinchu, Taiwan.

to have a protectant role against the destructive effects of salt, freezing, heating, and possibly drying (Somero, 1986; Lin & Bolen, 1995; Wang & Bolen, 1997). Hence, the compatible solutes have promising biotechnological potential as protecting agents for industrial enzymes.

Among the compatible solutes, the imino acid proline has interesting physical properties (Schobert & Tschesche, 1978; Srinivas & Balasubramanian, 1995). First, its solubility in water is remarkably high—as much as 7 M at ambient temperatures. Second, proline at high concentrations (>1.0 M) is shown to behave as a hydrotrope, enhancing the solubility of hydrophobic compounds such as fluorescein diacetate (Srinivas & Balasubramanian, 1995). Recently, in vitro studies have indicated that proline at high concentrations promotes the solubility of sparingly soluble proteins (Samuel et al., 1997). These unusual properties of proline have prompted us to investigate the role of proline as a protein folding chaperone.

Results and discussion

Refolding of the protein in various concentrations of proline

Oxidative refolding of reduced and denatured lysozyme (r/d Lys) (1 mg/mL) was achieved by dilution with the refolding buffer (0.1 M Tris-HCl, pH 7.2, containing a redox mixture of 3 mM reduced glutathione and 0.3 mM oxidized glutathione) containing various concentrations of proline ranging upto 4.5 M. Refolding of r/d Lys in the refolding buffer containing proline (at concentrations greater than 1.5 M) yields clear solution(s), with no visible signs of aggregation. The percentage transmittance of these solutions at 600 nm ($\%T_{600}$) is close to 100%. At proline concentrations lower than 1.5 M, the solution(s) containing the "refolded protein" are faintly turbid. Interestingly, when r/d Lys is refolded in the refolding buffer lacking proline, the resultant protein solution turns turbid (with $\%T_{600} = 35\%$), indicating the formation of protein aggregates during oxidative refolding (Fig. 1). These results indicate that proline at higher concentrations (>1.5 M) effectively prevents aggregation during refolding of the protein.

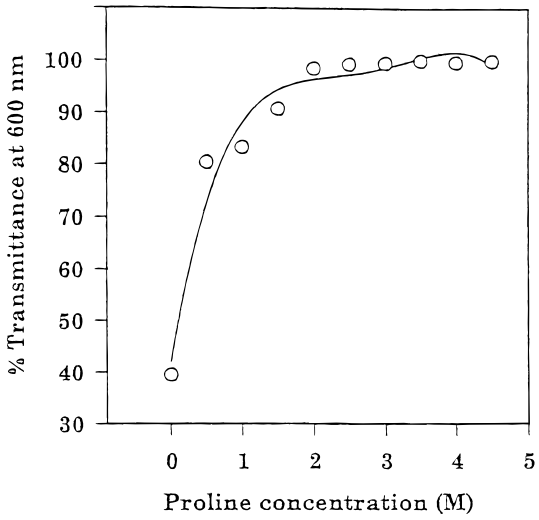

Fig. 1. Aggregation observed during the refolding of reduced and denatured lysozyme in various concentrations of proline dissolved in the refolding buffer (0.1 M Tris-HCl, pH 7.2). Concentration of the reduced and the denatured protein used for refolding is 1 mg/mL. The process of aggregation was monitored by the change(s) in the percentage transmittance ($\%T_{600}$) of the refolded protein solutions. Refolding was achieved by 10-fold dilution of the r/d Lys in the refolding buffer containing various concentration of proline at room temperature.

Protein concentration-dependent refolding

Concentration-dependent refolding of r/d Lys in the presence and absence of proline (in the refolding buffer) is shown in Figure 2. In the absence of proline in the refolding buffer, the $\%T_{600}$ values are observed to steadily decrease with the increase in the concentration of r/d Lys during refolding. Percent T_{600} values are found to drop below 10% when the protein concentration during refolding is increased beyond 3.5 mg/mL. However, in the presence of 2 M proline (in the refolding buffer), the $\%T_{600}$ value is close to 100% at all concentrations of the refolding protein below 5 mg/mL (Fig. S1). These results reveal that proline could inhibit visible protein aggregation even at high concentrations of the refolding protein (r/d Lys).

Comparison with other osmolytes

Compounds such as glycerol, sucrose, and ethylene glycol are well known for their ability to control aggregation during protein refolding (Wang & Bolen, 1997). To examine the uniqueness of proline in preventing protein aggregation, we studied the refolding of r/d Lys individually in the presence of ethylene glycol, glycine, glycerol, and sucrose. It could be noticed from Figure 3A that none of these compounds inhibit the process of aggregation of the protein during refolding. With the exception of sucrose, the $\%T_{600}$ of the protein solutions refolded in the presence of other osmolytes (glycerol, glycine, and ethylene glycol) are less than 10%, suggesting the formation of aggregates during refolding. At the high-

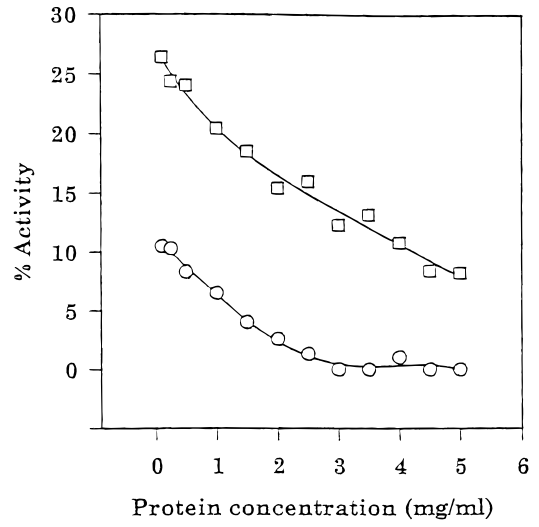

Fig. 2. Concentration dependent refolding of the reduced and denatured lysozyme in 2 M proline dissolved in the refolding buffer containing 2M proline (□) and in the refolding buffer alone (○). Aggregation of the protein was monitored from the changes in $\%T_{600}$ values of the refolded samples.

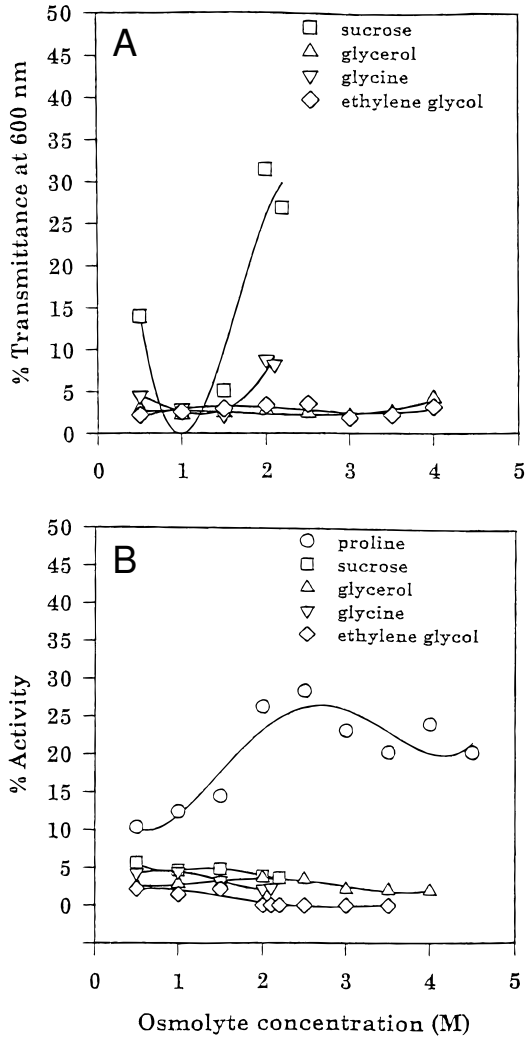

Fig. 3. Refolding of reduced and denatured lysozyme (r/d Lys) in various osmolytes. **A:** The percentage transmittance at 600 nm. **B:** The enzyme activity recovered upon refolding in various osmolytes. The percentage of bacteriolytic activity regained was estimated by considering the activity shown by the native lysozyme (in 0.1 M Tris-HCl, pH 7.2 containing 0.6 M GdnHCl, 0.2% β-mercaptoethanol, 3 mM reduced glutathione, and 0.3 M oxidized glutathione) as 100%. Bacteriolytic activity of the refolded protein mixture was measured after a incubation period of 60 min at room temperature.

est possible concentration of sucrose (>1.5 M), the $\%T_{600}$ of the refolded protein solution(s) is about 30%. Thus, among the various osmolytes used, proline appears to possess an unique property of preventing aggregation during refolding of protein(s).

Recovery of enzymatic activity

An ideal protein folding chaperone is not only expected to thwart protein aggregation during refolding but also aid in restoring the biological activity of the protein. Upon refolding r/d Lys, the percentage of bacteriolytic activity recovered is found to increase with the increasing concentrations of proline in the refolding buffer (Fig. 3B). In contrast, the percentage of enzyme activity regained upon refolding the protein (r/d Lys) in the presence of the other osmolytes used is insignificant (Fig. 3B). We also compared the bacteriolytic activity of r/d Lys samples refolded in the absence and presence of 2 M proline in the refolding buffer. Figure 2 shows the percentage of enzyme activity regained upon refolding various concentrations of r/d Lys in the presence of 2 M proline. In the absence of proline in the refolding buffer, the enzyme activity is found to decrease with the increase in the concentration of r/d Lys used for refolding (Fig. 2). A maximum of 10% activity is found to be recovered upon refolding r/d Lys in the absence of proline. Beyond a concentration (of r/d Lys) of 3 mg/mL, the percentage of enzyme activity recovered upon refolding r/d Lys in the absence of proline (in the refolding buffer) is close to zero. Interestingly, the percentage of bacteriolytic activity recovered upon refolding r/d Lys in the presence of 2 M proline is significantly higher than when the refolding was performed by dilution of r/d Lys in refolding buffer lacking proline (Fig. 2). Maximum enzyme activity (~30%) is recovered when the concentration of the protein (r/d Lys) used for refolding was 0.1 mg/mL. These results clearly suggest that presence of proline in the refolding buffer appreciably helps the protein to overcome aggregation and subsequently regain its bacteriolytic activity. It is important to note that although the protein refolded in 2 M proline shows no or very little aggregation, only a fraction of the protein molecules appear to have regained their biological active state upon refolding. In this context, it is presumed that a significant portion of the refolding protein molecules are locked into soluble, nonnative, and biologically inactive conformation(s) in the presence of proline (to be discussed later).

Folded/misfolded/aggregated species formed upon refolding

It is informative to evaluate the proportion of folded/misfolded or aggregated species formed upon refolding the protein from its reduced and denatured state. Analysis of the r/d Lys sample in the refolding buffer (in the absence of proline) using reversed-phase high performance liquid chromatography (HPLC), revealed at least three major components with retention times of 49.1, 60.3, and 60.9 min (Fig. 4A). Comparison with the HPLC profiles of the protein in its native and the denatured and reduced states (Fig. 4, inset) reveals that the peak with the retention time of about 49 min in the HPLC profiles of the sample (in the presence and absence of proline) represents the fraction of molecules that have regained the native state upon refolding. Interestingly, the HPLC profile of r/d Lys refolded in the absence of proline shows two other prominent peaks with retention times (60.3 and 60.9 min) longer than the protein in the completely unfolded state (in 6 M guanidinium hydrochloride (GdnHCl) containing 2% β-mercaptoethanol, Fig. 4). The presence of protein species (formed upon refolding) with retention times longer than that of the completely unfolded protein is perplexing but not unprecedented. 2,2,2, Trifluoro acetic acid (TFA), which is routinely used as an additive in the elution buffers (used in reversed phase HPLC) to improve the resolution of the HPLC profiles, is also known to affect the structure(s) of some proteins. This feature could consequently lead to in the distortion of the HPLC profile of the protein(s). However, the HPLC profiles of the refolded protein (in the presence and absence of proline) obtained using elution buffers with and without TFA are identical, implying that TFA per se has little or no effect on the structure of lysozyme (data not shown). In addition, fluorescence and circular dichroism (CD) spectra [near and far ultraviolet (UV)] of native lysozyme in water and in the presence of 30% acetonitrile (the gradient at

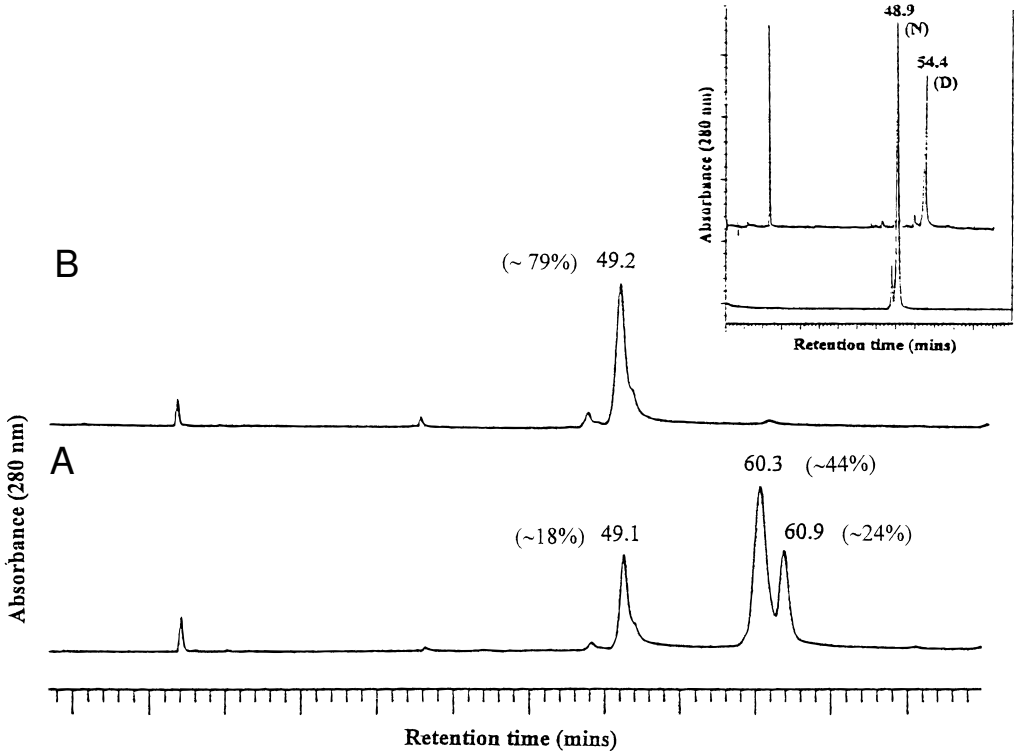

Fig. 4. HPLC profiles of r/d Lys refolded in (**A**) refolding buffer and (**B**) refolding buffer containing 2M proline. The inset figures represent the HPLC profiles of the native "N" (retention time 48.9 min) and the denatured and reduced "D" (retention time, 54.4 min) proteins. The elution of the protein was monitored by absorbance at 280 nm. Peaks eluting at 49.1/49.2 min corresponds to the refolded native state of the protein. The authenticity of this peak was verified by bacteriolytic activity and co-injection with the native protein. The refolded protein (in the presence and absence of proline) was desalted by ultrafiltration (at room temperature). Two hundred fifty microliters of the refolding buffer containing fixed amounts of the protein refolded in the presence (0.8 mg/mL) and absence of proline (0.92 mg/mL) were loaded. The values indicated in parenthesis beside the retention times indicate the percentage of protein recovered after refolding and subsequent desalting.

which native lysozyme elutes from the reversed-phase C_{18} column) containing 0.1% TFA were similar. These results suggest that the nonnative peaks observed in the HPLC profiles of r/d Lys refolded in the absence of proline are indeed protein species representing misfolded/aggregated molecules formed during refolding of the protein. Additionally, Kumar et al. (1994), examining the effect(s) of protein concentration on the disulfide refolding pathway of RNase A, reported the formation of protein species whose HPLC retention times are longer than the protein in the completely unfolded state. These protein fractions were attributed to soluble aggregates formed through intermolecular disulfide bonds. In this context, it appears that the peaks eluting later than the completely unfolded protein (in the HPLC profile, Fig. 4A) possibly represent the protein in the misfolded or soluble aggregated state(s) formed during refolding of the protein. In contrast, HPLC analysis of r/d Lys, refolded in the refolding buffer containing 2 M proline, shows only one major peak with a retention time (49.2 min) close to that of the native protein (Fig. 4B). The authenticity of this peak was verified by co-injecting the refolded sample (in 2 M proline) with the native lysozyme sample. If the major peak corresponds to the fraction of molecules that have regained the native state, then co-injection with native lysozyme is expected to produce only one major peak in the HPLC profile. On the contrary, if the major peak in the HPLC profile (of the protein sample refolded in 2 M proline) represents misfolded/aggregated species, then at least two major peaks are expected to be visualized in the HPLC profile obtained upon co-injection. It is observed that co-injection of the refolded protein sample along with the native protein shows a single peak with enhanced absorbance at 280 nm (data not shown). These results suggest that the major peak in the HPLC profile of r/d Lys, refolded in the presence of 2 M proline, corresponds to the protein in its native state. In addition, the major peak (49.2 min in the HPLC profile) upon isolation and subsequent purification yields a specific activity value 10,318 units/mg, which is about 84% of the value expected for the native protein (12,284 U/mg). These results indicate that proline effectively thwarts protein misfolding/aggregation and aids in the formation of the native biologically active state.

Mode of action of proline

It is important to note that although proline almost completely prevents protein aggregation during refolding of r/d Lys, the refolded protein sample shows a maximum of about 30% of enzyme activity expected for the native protein (lysozyme). The relatively low percentage of recovery of the bacteriolytic activity of the protein (refolded in the presence of proline) could be attributed to various reasons, such as (1) inhibition of the bacteriolytic activity of the refolded protein in high concentrations of proline; (2) formation of misfolded species or soluble aggregates during the re-

folding process; (3) due to binding of proline to the folding "intermediates" and temporarily trapping them into partially-structured biologically inactive states. The possibility of proline behaving as an inhibitor of the bacteriolytic activity of the protein could be discounted because it is found that native protein (lysozyme) incubated in 2 M proline shows no significant loss in its bacteriolytic activity (data not shown). The second possibility of the formation of irreversible misfolded species or soluble aggregates during refolding of the protein could also be ruled out based on the results of the HPLC experiments (discussed in the previous section). If a majority of the refolded protein molecules were to exist as misfolded species or soluble aggregates, then the HPLC profile of the refolded sample (in 2 M proline) is expected to yield multiple peaks. On the contrary, the HPLC profile of the protein refolded in the presence of 2 M proline shows a single peak (Fig. 4A), which co-elutes with the native protein. This aspect dispels the possibility that lower recovery of the enzymatic activity upon refolding (in 4 M proline) is due to formation of misfolded/aggregated species. If the third possibility of proline binding to the folding "intermediate(s)" and physically restricting the protein to refold back to its native conformation were to be valid, then removal or dilution of proline from the refolded sample mixture is expected to help the protein to revert back to its native, biologically active conformation. In this context, proline from the refolded samples was removed by ultrafiltration. It should be mentioned that 10-fold dilution of the protein refolded in 2 M proline is found to result in the increase in the percentage of bacteriolytic recovered (Fig. S2, see Supplementary material in Electronic Appendix). In addition, comparison of the bacteriolytic activity of the desalted, refolded protein with that of native lysozyme, under similar experimental conditions, revealed that the refolded protein (at 1 mg/mL concentration) regains about 75% of the enzymatic activity exhibited by the native protein (Fig. S3, see Supplementary material in Electronic Appendix). Although the results of these experiments suggest that proline prevents formation of aggregation by binding to the refolding intermediate(s), we are aware that these results only constitute an indirect evidence for the proposed mode of action of proline. More direct experimental proof is needed to support our conjecture. In light of these results obtained with proline, it would be of interest to learn that solvents/solutes such as polyethylene glycol, and certain surfactants have also been shown to prevent aggregation of protein(s) by binding to intermediate states formed during the refolding of proteins (Cleland & Wang, 1990; Cleland, 1993; Lin & Bolen, 1995). Cleland and Wang (1990), studying the effect of polyethylene glycol (PEG) on the refolding of bovine carbonic anhydrase found that PEG prevents aggregation during refolding of the protein by binding to the aggregation sensitive "molten globule" state. Similarly, Wetlaufer and Xie (1995), studying the effect(s) of various surfactants on the refolding of carbonic anhydrase II reported that surfactants and alkanols such as octyl glucoside and hexanol, prevent protein aggregation by binding to the hydrophobic sites on the "molten globule" state(s) formed during the refolding of the protein.

Probing the structure of the refolded protein

The conformation of the r/d Lys refolded in 2 M proline (after desalting) was probed using far-UV CD spectroscopy. The far-UV CD of lysozyme in 6 M GdnHCl containing 2% β-mercaptoethanol, as expected, shows a spectra resembling the protein in a random coil conformation (data not shown). On the other hand, native lysozyme in 0.1 M Tris-HCl (pH 7.2) shows two prominent negative ellipticity bands centered at 222 and 208 nm, indicative of the backbone of the protein being predominantly in a helical conformation (Fig. S4, see Supplementary material in Electronic Appendix). The far-UV CD spectra of the protein refolded in 2 M proline and subsequent removal of proline by ultrafiltration shows a CD curve that to a large extent resembles the spectra obtained for lysozyme in the native conformation (Fig. S4, see Supplementary material in Electronic Appendix). Thus, it appears that the presence of proline during refolding helps the protein to overcome aggregation and subsequently refold to biologically active, native conformation.

The conformation of the refolded protein was also monitored by intrinsic fluorescence measurements. Emission spectrum of native lysozyme in 0.1 M Tris-HCl (pH 7.2) shows an emission maxima at around 340 nm (Fig. S5, see Supplementary material in Electronic Appendix). The protein in its unfolded state (in 6 M GdnHCl containing 2% β-mercaptoethanol) showed a clear red shift (10 nm) in the emission maximum (to 350 nm), implying that the tryptophans in the protein are solvent exposed in the unfolded state of the protein (Fig. S5, see Supplementary material in Electronic Appendix). As proline was found to significantly interfere with the fluorescence properties of lysozyme, the emission spectra of the refolded protein sample (in 2 M proline) were measured after removal of proline by ultrafiltration. The desalted, refolded protein sample shows a fluorescence spectrum that closely resembles the native protein with an emission wavelength maximum of 341 nm (Fig. S5, see Supplementary material in Electronic Appendix). These results strongly suggest that a large population of the protein molecules refolded in 2 M proline regain their biologically active conformation.

Formation of supramolecular assembly

A detailed picture on the role of proline in the prevention of aggregation during protein refolding would only emerge when the structural properties of proline are clearly understood. In this context, we decided to examine the physical properties of proline in greater detail.

Laser light scattering experiments were performed at varying concentrations of proline (0.5–4.5 M) to investigate its molecular state in solution. The dependence of scattering intensity (counts per second) as a function of proline concentration at right angle is shown in Figure 5. The average scattering intensity does not show any significant change at lower concentration of proline ($<$1.0 M). However, increasing the proline concentration beyond 1.5 M results in a prominent increase in the scattering intensity, suggesting aggregation of the proline molecules in solution (Fig. 5. However, the scattering intensities tend to saturate at higher concentrations of proline beyond 2.5 M. As the change(s) in the average scattering intensity as a function of the proline concentration are not strong, it could be presumed that the supramolecular assembly is relatively loose. In conclusion, the results of the light scattering experiments appear to suggest that proline at higher concentration ($>$1.5 M) forms higher order assemblies.

Schobert and Tschesche (1978), studying the solution properties of proline, reported that proline increases the solubility of proteins in water. These authors attribute this property to the tendency of proline to form higher order aggregates. In this background, we attempted to obtain independent evidence of the formation of "molecular aggregates" at high concentration(s) in water.

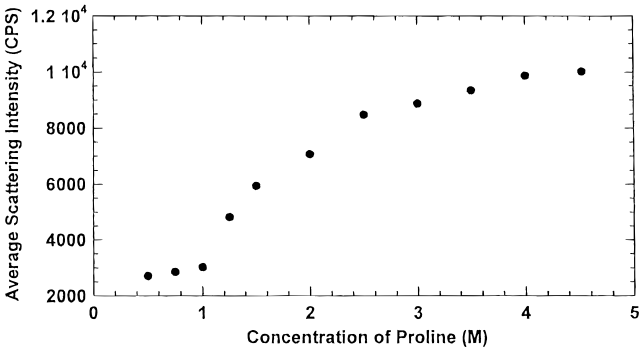

Fig. 5. A plot of the variation in the scattering intensity as a function of proline concentration. All measurements were made at a scattering angle of 90°.

1-Anilino-8-naphthalene sulfonic acid (ANS), a hydrophobic dye, is an useful probe to monitor the existence of any exposed hydrophobic surface(s) in the higher order aggregates/supramolecular assembly of proline. The emission maxima of ANS is blue-shifted by 25 nm from 524 to 499 nm in the proline concentration range of 0–4.5 M (Fig. S6). Such significant blue shifts in the emission maxima could be attributed to the availability of solvent exposed hydrophobic pockets in the supramolecular assembly of proline. The emission intensity of ANS shows a sevenfold increase (Fig. S6) in the range of proline concentration used (0–4.5 M). The increase in the ANS emission intensity is gradual at lower concentrations of proline (up to 1.5 M). Beyond this proline concentration, the emission intensity increases exponentially. Such an exponential increase in the emission intensities is generally believed to be due to the decrease in the internal relaxation of the fluorophore in a spatially restricted environment (Lakowicz, 1983). Control ANS binding experiments in glycerol show that ANS emission intensity increases steadily with increase in the concentration of glycerol (Fig. S6). The steady increase in ANS emission intensity in glycerol could be attributed to the progressive increase in the solvent dielectric constant. Thus, the ANS binding experiments suggest that proline forms an ordered supramolecular assembly with distinct hydrophobic surface(s) conducive for ANS binding.

It is generally believed that binding of ANS to proteins mainly occurs through its hydrophobic naphthalene moiety and the negatively charged sulfonate group merely acts as a water solubilizing group. However, recently Kirk et al. (1996), investigating the binding modes of ANS to proteins, reported that the electrostatic interactions contribute significantly to the binding of the hydrophobic dye (ANS) to the protein. Matulis and Lovrfen (1998), estimating the binding enthalpies of ANS to four proteins, suggested that the ANS anion is predominantly bound to cationic groups of water soluble proteins and polyamino acids through ion pair formation. However, proline is a neutral amino acid and lacks cationic charge(s) conducive to promote ion pair interaction between the dye and proline. Hence, it appears that nonpolar rather than charge–charge interactions play a dominant role in the binding of ANS to the supramolecular assembly.

We further investigated the microenvironment in the supramolecular assembly of proline using a hydrophobic probe such as pyrene. Pyrene is a polycyclic aromatic compound with very low solubility in water. This dye has been extensively used to study the hydrotropic character of various solutes (Dong & Winnik, 1982). Hence, if the supramolecular assembly of proline were to possess a hydrophobic character, it is expected to enhance the solubility of pyrene in water. Figure S7 (see Supplementary material in Electronic Appendix) shows the enhancement of the solubility of the dye (pyrene) in water, brought about by increasing concentrations of proline. Addition of proline up to almost 1.2 M does not solubilize pyrene to any significant degree. Further addition of proline results in a gradual increase in the solubility of pyrene. However, beyond 4.2 M concentration(s) of proline, the absorbance at 350 nm due to the dissolved pyrene reaches a plateau. Interestingly, proline in the concentration range of 0–5.4 M brings about a 12-fold increase in the solubility of pyrene in water. These results suggest that the supramolecular assembly of proline has a partial hydrophobic character through which proline could transiently interact with the nonpolar portions of the protein and consequently inhibit aggregation during refolding.

Although the crystal structure does not appropriately represent the properties exhibited by proline in solution, it could provide useful insights into the phenomenon of formation of high-order aggregates. In this context, we studied the crystal structure of proline. The crystal structure of proline reveals several interesting features. As predicted from ANS binding, pyrene solubility, and light scattering experiments, the crystal structure of proline reveals that proline exists as loose higher order aggregates (Fig. S8, see Supplementary material in Electronic Appendix). The pyrolidine rings of each proline molecule are observed to stack one over the other just as a stacked pile of coins. The crystal structure of proline determined at low resolution by Kayushina and Vainshtein (1983) also suggests an orderly packing and layering of the pyrolidine rings. The crystal structure determined at high resolution in the present study shows the supramolecular assembly of proline to be stabilized by the intermolecular hydrogen bonding between the imino group of the proline molecule with the negatively charged carboxylate group of the adjacent proline molecule. In addition, the carboxyl groups of proline are also found to form hydrogen-bonding with the solvent water molecules (Fig. S8, see Supplementary material in Electronic Appendix). The ordered stacking of the proline molecules appears to bestow an amphipathic character to the supramolecular assembly with the imino and the carboxyl groups facing on one side of the assembly providing the polar surface and the methylene groups of the pyrrolidine rings constituting the hydrophobic surface. Protein aggregation is believed to stem primarily due to the intermolecular hydrophobic interactions among the folding molecules. One of the ways to effectively inhibit protein aggregation during refolding would be to evolve a mechanism to effectively bind to the hydrophobic groups that tend to get solvent-exposed during protein refolding. In this respect, proline seems to be an ideal molecule. Due to the amphipathic nature of the supramolecular assembly, proline appears to provide the required nonpolar surface that can interact with the aggregation prone, solvent exposed hydrophobic residues in the protein during the refolding process. In such an event, protein aggregation is effectively blocked.

The organization of any supramolecular assembly is expected to be sensitive to temperature and is likely to undergo partial or complete collapse beyond a particular temperature. Thus, if proline were to organize into a supramolecular assembly at higher concentrations, then this higher order organization (of proline) is also liable to be disrupted at some higher temperature. In a disorganized state, proline is expected to be less effective in inhibiting protein

aggregation during refolding. To examine this aspect, we performed temperature-dependent refolding experiments in the presence and absence of proline. It can be visualized from Figure S9 (see Supplementary material in Electronic Appendix) that the $\%T_{600}$ value steadily drops when the refolding of the protein in 2 M proline is carried out beyond 55 °C indicating the onset of protein aggregation. Thus, it appears that beyond 55 °C the supramolecular assembly of proline tends to collapse and as a consequence it (proline) is unable to efficiently thwart protein aggregation. As expected, there is no significant change in $\%T_{600}$ values when the protein is refolded in the refolding buffer alone in the temperature range of 30–75 °C. Figure S10 (see Supplementary material in Electronic Appendix) depicts the bacteriolytic activity of lysozyme refolded in 2 M proline at various temperatures. It could be seen that the bacteriolytic activity of the refolded lysozyme shows a small increase in the temperature range between 30–50 °C (Fig. S10, see Supplementary material in Electronic Appendix). Maximal enzyme activity of about 32% is observed when the protein is refolded at 50 °C. Similarly, an increase in enzyme activity has been previously reported for lysozyme (Saxena & Wetlaufer, 1970) upon refolding of r/d Lys in aqueous solutions. Interestingly, when the refolding is carried out beyond 65 °C, a drastic decrease in the enzyme activity is noticed. The bacteriolytic activity is only about 8% (of the activity shown by native lysozyme) when the refolding was performed at 75 °C. It appears that the steep decline in the activity of lysozyme could be due to disorganization of the supramolecular assembly of proline at temperatures greater than 55 °C. The control experiments, wherein the $\%T_{600}$ of native lysozyme were monitored over the temperature range of 30–70 °C, showed no significant change (data not shown). Thus, these experiments clearly indicate that the unique property of proline to block protein aggregation stems from its ability to form an amphipathic supramolecular assembly.

The results presented in this paper strongly suggest that proline, in vitro, behaves as a protein folding chaperone. It is possible that, in vivo, under water stress conditions the accumulated proline in the cell not only has an osmoregulatory role in the cytoplasm but also possibly prevents aggregation of proteins during folding on the ribosome. Work is currently in progress to characterize protein–proline complex to elucidate the molecular mechanism underlying the chaperone activity of proline.

Materials and methods

Hen egg white lysozyme, oxidized and reduced glutathione, and *Micrococcus lysodeikticus* cells were purchased from Sigma Chemical Co. (St. Louis, Missouri). Proline and glycine were purchased from Lancester, England. Guanidinium hydrochloride was purchased from E. Merk (Germany). All other chemicals used were of high quality analytical grade.

Preparation of reduced and denatured lysozyme (r/d Lys)

Human egg white lysozyme was denatured and disulfide reduced in 100 mM Tris-HCl, pH 8.7 containing 6 M guanidinium hydrochloride and 2% β-mercaptoethanol. The solution was incubated overnight at room temperature before refolding experiments were carried out. The concentration of the reduced and denatured protein was calculated based on the extinction coefficient at 280 nm. All UV spectroscopic measurements were made on a Hitachi (Model U-3300) spectrometer.

Protein refolding monitored by UV-absorption spectroscopy

Protein refolding was carried out at room temperature by a 10-fold dilution of the reduced and denatured lysozyme into the refolding buffer (100 mM Tris-HCl pH 7.2, containing 3 mM reduced glutathione and 0.3 mM oxidized glutathione) or with the relevant osmolytes dissolved in the refolding buffer. The tubes were constantly flushed with nitrogen during refolding of the protein. The extent of aggregation in each case was monitored by the percentage transmittance (at 600 nm) of the solutions containing the refolded protein. The protein aggregates formed upon refolding were removed by centrifugation at 3,000 rpm, and the supernatant solutions were used further for the enzyme assay.

Ultrafiltration

r/d Lys samples refolded in the refolding buffer (with and without 2 M proline) were desalted using an Amicon ultrafiltration set-up. The refolded protein sample was buffer-exchanged with five volumes of 0.1 M Tris-HCl (pH 7.2). Ultrafiltration of the refolded protein samples was carried out under a pressure of N_2 gas. The protein samples were later centrifuged at 3,000 rpm, and the clear supernatant were chosen for further analysis.

Enzyme activity assay

The bacteriolytic activity of the refolded lysozyme samples was estimated as per the method reported by Jolles (1962). *Micrococcus lysodeikticus* suspension was prepared by mixing 0.2 mg of dried bacterial cells per mL of 100 mM phosphate buffer, pH 6.2. One milliliter of *M. lysodeikticus* cell suspension was added to 5 μL of refolded lysozyme solution in a 1 mL methacrylate cuvette. The mixture was mixed well and the activity measured as a function of time. The decrease in light scattering intensity of the solution was measured by following the increase in the percentage transmittance ($\%T_{600}$) of the solution at 600 nm. The percentage of enzyme activity regained upon refolding was estimated by comparing with the activity of native lysozyme under identical conditions containing appropriate amounts of guanidinium hydrochloride and the redox mixture (oxidized and reduced glutathione). One unit of enzyme activity corresponds to an absorbance (at 450 nm) decrease of 0.0026 per minute.

High performance liquid chromatography (HPLC)

The native and r/d Lys sample refolded separately in the refolding buffer in the absence and presence of 2 M proline was analyzed using C_{18} reversed-phase columns. The proteins were eluted using a linear water-acetonitrile gradient (0–60% v/v) containing 0.1% TFA. The proteins eluted out were detected by their 280 nm absorption. All eluents were degassed prior to use. All experiments were performed on a HPLC (Model L4000H) procured from Hitachi Co. (Japan).

Circular dichroism (CD)

Far-UV CD experiments were performed on a Jasco J720 spectropolarimeter equipped with a NESLAB circulating water bath. All experiments were performed at 25 °C. Measurements were made using a 0.02 cm pathlength water-jacketed quartz cell. The concentration of the protein samples were set to 100 μM. The

mean residue ellipticity values were estimated using standard procedures. The CD instrument was calibrated using camphor-10-sulfonic acid (ammonium salt) periodically prior to use.

Temperature-dependent refolding

For the temperature-dependent refolding experiments, reduced and denatured lysozyme was diluted with the refolding buffer alone or with the refolding buffer containing 2 M proline at appropriate temperatures. The temperature was maintained using a variable temperature water-bath (Neslab RTE 111). The percentage of enzyme activity regained upon refolding was calculated by comparison with the activity of native lysozyme treated under similar temperature conditions and at the same denaturant concentrations as that of the refolded protein.

ANS-binding experiments

To appropriate concentrations of proline (0–5.5 M), 250 μM of 1-anilino-8-naphthalene sulfonic acid was added, and the fluorescence spectra were recorded at room temperature between 450 and 600 nm using an excitation wavelength of 400 nm. All necessary background corrections were made. The fluorescence measurements were made on a Hitachi spectrofluorimeter (Model F-2500).

Solubilization of pyrene

Solubility of pyrene in water containing various concentrations of proline (0–5.4 M) was monitored by the increase in the absorbance of the solutions at 350 nm. Five milligrams of pyrene was dissolved in 5 mL of water and equilibrated overnight (at room temperature) in a constant temperature water bath. The samples were centrifuged, and the solubility was assessed from the absorbance at 350 nm after necessary background corrections.

Measurement of intrinsic fluorescence

The intrinsic tryptophan fluorescence spectra of the native, refolded, and desalted lysozyme samples in the relevant solvents were recorded at room temperature between 300 and 500 nm. The excitation wavelength used was 280 nm. The excitation and emission slit widths were set to 10 and 2.5 nm, respectively.

Light scattering

Measurements were performed on a super dynamic light scattering photometer model DLS-7000 with control unit, Model LS-71 and pump controller, Model LS-72 (Photal Otsuka Electronics, Japan) using a He-Ne laser light source of 10 mW at 643.0 nm. The instrument was operated in conjunction with a software controlled, stepping-motor driven goniometer (with an angle range of 5–160° and an accuracy of $\pm 0.1°$), a correlator, photo counting detector, and a temperature control device. The measurements were made at 10° interval in the scattering angle range of 50–140° at room temperature. The concentration of proline used in this study was 0.5–4.5 M. Samples were filtered twice before measurements.

Proline crystallization and structure determination

Crystals of proline were grown from a saturated solution of proline (~7 M) by air evaporation of the solvent. The crystals obtained were needle-shaped, transparent, bifringent, and approximately rectangular in cross section. The needles were elongated. All crystal data were acquired using Seiman's R3M/V single crystal diffractometer and were processed using SHELXTL Plus software.

Supplementary material in Electronic Appendix

Description of the supplementary figures included in Electronic Appendix:

Fig. S1. Concentration dependent aggregation of r/d lysozyme in 2 M proline.

Fig. S2. Recovery of enzymatic activity upon 10-fold dilution with the refolding buffer.

Fig. S3. Time course of bacteriolytic activity of native and refolded lysozyme.

Fig. S4. Far-UV CD spectra of native and refolded lysozyme.

Fig. S5. Emission spectra of native and refolded lysozyme.

Fig. S6. ANS binding profiles at various concentrations of proline.

Fig. S7. Solubility curve of pyrene in proline solutions.

Fig. S8. Crystal structure of the molecular assembly of proline.

Fig. S9. Temperature dependent aggregation of r/d lysozyme.

Fig. S10. Recovery of enzymatic activity as a function of temperature.

Acknowledgments

This work was supported by the grants from the National Science Council of Taiwan, Academia Sinica (Taiwan), and the Dr. C.S. Tsong Memorial Medical Research Foundation.

References

Borrowitzka LJ. 1985. In: Gilles RGB, ed. *Transport processes, iono- and osmoregulation.* Berlin: Springer-Verlag. pp 437–453.

Brems DN. 1988. Solubility of different folding conformers of bovine growth hormone. *Biochemistry 27*:4541–4545.

Brown AD, Simpson JR. 1972. Water relations of sugar-tolerant yeasts: The role of intracellular polyols. *J Gen Microbiol 72*:589–591.

Cleland JL. 1993. *Protein folding in vitro and in vivo*. ACS Symposium Series 470. Washington, D.C.: American Chemical Society.

Cleland JL, Wang DIC. 1990. Refolding and aggregation of bovine carbonic anhydrase B: Quasi-elastic light scattering analysis. *Biochemistry 29*:11072—11078.

De Young LR, Dill KA, Fink AL. 1993. Aggregation and denaturation of apomyoglobin in aqueous urea solutions. *Biochemistry 31*:12345–12352.

Dill KA. 1990. Dominant forces in protein folding. *Biochemistry 29*:7133–7155.

Dong DC, Winnik MA. 1982. The PY scale of solvent polarities—Solvent effects on the vibronic fine structure of pyrene fluorescence and empirical correlations with Et-value and Y-value. *Photochem Photobiol 35*:17–21.

Galinski EA. 1993. Compatible solutes of halophilic eubacteria: Molecular principles, water solute interaction, stress protection. *Experimentia 49*:487–494.

Georgia G, DeBernardez CE. 1991. *Protein refolding*. ACS Symposium Series 470. Washington, D.C.: American Chemical Society.

Goldberg ME. 1985. The second translation of the genetic message: Protein folding and assembly. *Trends Biochem Sci 10*:388–392.

Goldberg ME, Rudolph R, Jaenicke R. 1991. A kinetic study of the competition between renaturation and aggregation during the refolding of denatured-reduced hen egg white lysozyme. *Biochemistry 30*:2790–2797.

Hochachka PW, Somero GN. 1984. *Water-solute adaptations: The evolution and regulation of biological solutions*. Princeton, NJ: Princeton University Press. pp 304–354.

Jolles P. 1962. Lysozymes from rabbit spleen and dog spleen. *Methods Enzymol 5*:137–140.

Karuppiah N, Sharma A. 1995. Cyclodextrins as protein folding aids. *Biochem Biophys Res Commun 271*:60–66.

Kayushina RL, Vainshtein BK. 1983. Structure determination of L-proline by X-ray diffraction. *Kristallographia 10*:834–844.

Kirk G, Kurian E, Pendergast F. 1996. Characterization of the source of protein-ligand affinity, 1-sulfonate 8-(1′) anilino naphthalene binding to intestinal fatty acid binding protein. *Biophys J 70*:69–83.

Kumar TKS, Gopalakrishna K, Ramakrishna T, Pandit MW. 1994. Refolding of Rnase-A at high concentrations—Identification of non-native species. *Int J Biol Macro Mol 16*:171–176.

Kumar TKS, Jayaraman G, Lee CS, Sivaraman T, Lin WY, Yu C. 1995. Identification of "molten globule" like state in an all β-sheet protein. *Biochem Biophys Res Commun 207*:536–543.

Kumar TKS, Yang PW, Lin SH, Wu CY, Lei B, Lo SJ, Tu SC, Yu C. 1996. Cloning, direct expression and purification of snake venom cardiotoxin in *Escherichia coli*. *Biochem Biophys Res Commun 219*:450–456.

Lakowicz JR. 1983. *Principles of fluorescence spectroscopy*. New York: Plenum Press. pp 341–381.

Lin TY, Bolen DW. 1995. The peptide backbone plays a dominant role in protein stabilization by naturally occurring osmolytes. *Biochemistry 34*:12884–12891.

Lin TY, Timasheff SN. 1994. Why do some organisms use a urea methylamine mixture as osmolyte? Thermodynamic compensation of urea and trimethylamine N-oxide interactions with protein. *Biochemistry 33*:12695–12701.

Marston FA. 1986. The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. *Biochem J 240*:1–12.

Matulis D, Lovrfen R. 1998. 1-Anilino-8-naphthalene sulfonate anion-protein binding depends primarily on ion-pair formation. *Biophys J 74*:422–429.

Mitraki A, King J. 1989. Protein folding intermediates and inclusion body formation. *Bio/Technology 7*:690–697.

Orsini G, Goldberg ME. 1978. The renaturation of reduced chymotrypsinogen A in guanidine HCl: Refolding versus aggregation. *J Biol Chem 253*:3453–3458.

Orsini G, Skrzynia C, Goldberg ME. 1975. The renaturation of reduced polyalanyl-chymotrypsinogen and chymotrypsinogen. *Eur J Biochem 59*:433–440.

Ptitsyn OB. 1992. In: Creighton TE, ed. *Protein folding*. New York: Freeman. pp 243–300.

Ptitsyn OB, Pain RH, Semisotnov GV, Zerovnik F, Razgulyaev OI. 1990. Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett 262*:20–24.

Rozema D, Gellman SH. 1995. Artificial chaperones: Protein refolding via sequential use of detergent and cyclodextrin. *J Am Chem Soc 117*:2373–2374.

Rozema D, Gellman SH. 1996a. Artificial chaperone assisted refolding of denatured-reduced lysozyme: Modulation of the competition between renaturation and aggregation. *Biochemistry 35*:15760–15771.

Rozema D, Gellman SH. 1996b. Artificial chaperone assisted refolding of carbonic anhydrase B. *J Biol Chem 271*:3478–3487.

Samuel D, Kumar TKS, Jayaraman G, Yang PW, Yu C. 1997. Proline is a protein solubilizing solute. *Biochem Mol Biol Int 41*:235–242.

Saxena VP, Wetlaufer DB. 1970. Formation of three-dimensional structure in proteins. Rapid nonenzymatic reactivation of reduced lysozyme. *Biochemistry 9*:5015–5023.

Schobert B, Tschesche H. 1978. Unusual solution properties of proline and its interaction with proteins. *Biochim Biophys Acta 541*:270–277.

Schoffenials E, Gilles P. 1970. In: Floricin M, Scheer B, eds. *Chemical zoology*, London: Academic Press. pp 343–420.

Semisotnov GV, Uversky VN, Sokolovsky IV, Gutin AM, Razgulyaev OI, Rodionova NA. 1990. Two slow stages in refolding of bovine carbonic anhydrase B are due to proline isomerization. *J Mol Biol 213*:561–568.

Sivaraman T, Kumar TKS, Jayaraman G, Han CC, Yu C. 1997a. Characterization of partially structured state in an all β-sheet protein. *Biochem J 321*:457–464.

Sivaraman T, Kumar TKS, Jayaraman G, Yu C. 1997b. The mechanism of 2,2,2-trichloro acetic acid–induced protein precipitation. *J Prot Chem 16*:291–297.

Somero GN. 1986. Protons, osmolytes and fitness of internal milieu for protein function. *Am J Physiol 251*:197–213.

Srinivas V, Balasubramanian D. 1995. Proline is a protein compatible hydrotrope. *Langmuir 11*:2830–2833.

Tandon S, Horowitz PM. 1986. Detergent-assisted refolding of guanidium chloride denatured rhonase. The effect of lauryl maltoside. *J Biol Chem 261*:15615–15618.

Thomas PJ, Qu BH, Pedersen PL. 1995. Defective protein folding as a basis of human disease. *Trends Biochem Sci 20*:456–459.

Wang A, Bolen DW. 1996. Effect of proline on lactate dehydrogenase activity: Testing the generality and scope of the compatibility paradigm. *Biophys J 71*:2117–2122.

Wang A, Bolen DW. 1997. A naturally occurring protective system in urea rich cells: Mechanism of osmolyte protection of proteins against urea denaturation. *Biochemistry 36*:9101–9108.

Wegmann K. 1986. Osmoregulation in eukaryotic algae. *FEMS Microbiol Rev 39*:37–43.

Wetlaufer DB, Xie Y. 1995. Control of aggregation in protein refolding: A variety of surfactants promote renaturation of carbonic anhydrase II. *Protein Sci 4*:1535–1543.

Wetzel R. 1994. Mutations and off-pathway aggregation of proteins. *Trends Biotechnol 12*:193–198.

Wohlfarth K, Severin J, Galinski EV. 1990. The spectrum of compatible solutes in heterotrophic halophilic eubacteria of the family Halomonadaceae. *J Gen Microbiol 136*:705–712.

Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN. 1982. Living with water stress: Evolution of osmolyte systems. *Science 217*:1214–1222.

Zardeneta G, Horowitz PM. 1992. Micelle assisted protein folding. Denatured rhodanese binding to cardiolipin containing lauryl maltoside micellus results in slower refolding kinetics but greater enzyme reactivation. *J Biol Chem 267*:5811–5816.

Zetllmeissl G, Rudolph R, Jaenicke R. 1979. Reconstitution of lactic dehydrogenase: Noncovalent aggregation vs. reactivation. Physical properties and kinetics of aggregation. *Biochemistry 18*:5567–5571.