Role of Arginine in the Stabilization of Proteins against Aggregation†

Brian M. Baynes, Daniel I. C. Wang, and Bernhardt L. Trout*

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

Received November 23, 2004; Revised Manuscript Received January 28, 2005

ABSTRACT: The amino acid arginine is frequently used as a solution additive to stabilize proteins against aggregation, especially in the process of protein refolding. Despite arginine's prevalence, the mechanism by which it stabilizes proteins is not presently understood. We propose that arginine deters aggregation by slowing protein−protein association reactions, with only a small concomitant effect on protein folding. The associated rate effect was observed experimentally in association of globular proteins (insulin and a monoclonal anti-insulin) and in refolding of carbonic anhydrase. We suggest that this effect arises because arginine is preferentially excluded from protein−protein encounter complexes but not from dissociated protein molecules. Such an effect is predicted by our gap effect theory [Baynes and Trout (2004) *Biophys. J. 87*, 1631] for "neutral crowder" additives such as arginine which are significantly larger than water but have only a small effect on the free energies of isolated protein molecules. The effect of arginine on refolding of carbonic anhydrase was also shown to be consistent with this hypothesis.

Proteins are inherently unstable in aqueous solution and degrade by a variety of routes, the most common of which is aggregation. Aggregation is the assembly of non-native protein conformations into multimeric states, often leading to phase separation and precipitation. The problem of aggregation is especially grave in the pharmaceutical industry and in biotechnology, where it can be necessary to handle and store proteins at high concentrations and temperatures and for long periods of time.

In addition to aggregation being important in the storage of proteins, it is the dominant mode of protein degradation in protein refolding. Overproduction of recombinant proteins often results in a majority of the protein being produced in the form of phase-separated inclusion bodies (*1*). When this occurs, the inclusion bodies must be harvested, solubilized with a strong denaturant, and then refolded by removal of the denaturant to yield active protein. When the denaturant is removed, the hydrophobic effect drives the unfolded protein molecules to sequester their hydrophobic groups (*2*). This can occur either in an intramolecular fashion (proper protein folding) or in an intermolecular fashion (aggregation), as illustrated schematically by the reactions:

$$U \rightarrow N \quad (1)$$

$$U + U \rightarrow A_2 \quad (2)$$

where U, N, and A_2 represent the unfolded protein, native protein, and a dimer, respectively. Thus, there is direct competition between proper protein refolding and aggregation (*3*).

Empirically, it has been observed that, by adding low molecular weight components, such as salts, sugars, or polyols, to protein solutions, the propensity of a protein to aggregate can often be affected significantly (*4*, *5*). Unfortunately, because proteins are diverse in chemistry and structure, additives that work well for a particular protein may not work universally. In addition, current understanding of the mechanisms by which additives confer stability on proteins is limited. Thus, there is often no theoretical guidance to aid in selection of optimal additives, necessitating that protein stabilization be carried out on a case-by-case basis using heuristic experimental screens.

A prevalent antiaggregation additive whose mechanism of function is unknown is the amino acid arginine. Arginine has very little effect on the folding equilibrium (*6*−*8*), yet it facilitates refolding of several types of proteins from the unfolded state, such as tPA (*9*), interferon γ (*10*), lysozyme (*11*), carbonic anhydrase II (*11*), factor XIII (*12*), and antibodies (*13*). While a mechanism which can explain how arginine functions has not been proposed (*1*, *14*), these results suggest that arginine selectively slows protein−protein association (eq 2) while having little effect on protein folding (eq 1).

In recent theoretical studies of the effects of solution additives on protein aggregation and association, we developed a theory that may explain how arginine deters aggregation (*15*). This theory builds on previous molecular level understanding of additive effects on protein thermodynamics (*16*), preferential binding (*17*), osmotic stress (*18*), and Kirkwood−Buff theory (*19*−*22*).

"Gap effect theory" suggests that solution additives much larger than water which do not affect the free energy of isolated protein molecules will selectively increase the free energy of protein−protein encounter complexes. This effect will increase the activation free energy for association and, therefore, slow protein−protein association reactions. The accompanying effect on intramolecular reactions such as refolding is predicted to be small.

† The authors acknowledge funding support from the National Institutes of Health, the National University of Singapore, and the Singapore−MIT Alliance.

* Corresponding author. E-mail: trout@mit.edu. Phone: (617) 258-5021. Fax: (617) 258-5042.

10.1021/bi047528r CCC: $30.25
Published on Web 03/05/2005

We termed solution additives that have the above properties "neutral crowders" because of their size (crowder) and affinity for isolated protein molecules (neutral). The effect of such molecules on protein association reactions contrasts with that of excluded or hard-sphere crowders, which can accelerate association and generally shift the association equilibrium toward the associated state (*23*, *24*).

On the basis of the above theoretical developments and the existing experimental data on arginine systems, we hypothesized that arginine is a neutral crowder, and it exerts its beneficial effect on protein refolding by slowing protein association reactions with only a small concomitant effect on the rate of protein refolding. The goal of the present investigation was to test this hypothesis experimentally, compare the experimental results to gap effect model predictions, and determine the extent to which arginine can decrease the degree of aggregation of a globular protein.

Because gap effect theory predicts that arginine should decrease protein−protein association rates in general, this effect can be tested on any convenient system. We selected two types of protein association reactions for study: the association of insulin with a monoclonal antibody to insulin (globular protein association) and association of folding intermediates and aggregates of carbonic anhydrase II (aggregation during refolding). By performing these association tests in different buffers, the effect of arginine in the buffer can be deduced by comparison. In parallel, the effects of guanidinium chloride on the same association/aggregation systems was assessed. Finally, the experimental results were reconciled with gap effect theory.

MATERIALS AND METHODS

Proteins and Reagents. Human insulin (I8530), bovine carbonic anhydrase II (CA)[1] (C2522), hen egg white lysozyme (L7651), and bovine serum albumin (B4287) were obtained from Sigma-Aldrich (St. Louis, MO). Monoclonal anti-insulin (10-I30 clone M322214) was obtained from Fitzgerald Industries (Concord, MA). Consumable reagents for Biacore experiments (NHS, EDC, ethanolamine, glycine, and HBS-EP buffer) were obtained from Biacore AB (Switzerland). Guanidinium chloride, arginine hydrochloride, and sodium chloride were attained from Sigma-Aldrich in the highest available grade.

The concentration of carbonic anhydrase in solution was determined by absorbance at 280 nm using an extinction coefficient of 54000 M^{-1} cm^{-1} (*25*).

Globular Protein Association Kinetics. Protein association and dissociation rate constants, k_a and k_d, were measured for globular proteins via surface plasmon resonance on a Biacore 3000 instrument. Monoclonal anti-insulin was immobilized on a Biacore CM5 sensor chip via amine coupling. The amount of immobilized antibody was selected to give a detector response in the range of 50−100 RU when antigen was present. A reference surface was created by activating and deactivating the surface without coupling an antibody to it.

Different concentrations of insulin in the nanomolar range (1−200 nM) were prepared by dilution and injected serially into the antibody-containing and reference flow cells. Such low concentrations were used to ensure that multimerization of insulin did not affect the results (*26*). The dissociation rate was sufficiently fast in buffer that a regeneration buffer was not required. Kinetic constants were extracted by simultaneous fitting of k_a and k_d to each set of sensorgrams using a 1:1 kinetic model in the BIAevaluation 3.0 software package.

Refolding of Carbonic Anhydrase. Refolding of carbonic anhydrase was accomplished by dilution from high concentrations of the denaturant guanidinium chloride (GuHCl) as done previously (*27*, *28*). High concentrations of carbonic anhydrase (>300 μM) were denatured in 6 M GuHCl and equilibrated overnight. Refolding was initiated by dilution to 0.5 M GuHCl with 50 mM Tris-HCl buffer, pH 7.5. This final GuHCl concentration was selected because it yields a mixture of active, refolded protein and aggregates. The distribution of this mixture was analyzed via esterase activity, size exclusion HPLC, and dynamic light scattering as described below.

Carbonic Anhydrase Esterase Activity. Esterase activity of carbonic anhydrase was assessed using *p*-nitrophenyl acetate (pNPA) as the substrate as described previously (*25*). Briefly, 10 μL samples of carbonic anhydrase solution were added to 500 μL of Tris-HCl, pH 7.5, and 50 μL of 50 mM pNPA in acetonitrile. The kinetics of hydrolysis of pNPA was observed by the increase in absorbance at 400 nm due to the appearance of the *p*-nitrophenolate ion (pNP^-). In all cases, the observed hydrolysis rate in absorbance units per second (AU/s) under these conditions was constant (pseudo zero order). Hydrolysis rates were corrected for the hydrolysis of pNPA by the buffer for each type of buffer used. Hydrolysis rates were converted to concentration of active protein via a standard curve constructed from dilutions of known concentrations of native protein. The active protein concentration data were reproducible to within 5−8% in replicated experiments.

Size Exclusion HPLC. Size exclusion HPLC (SE-HPLC) experiments were performed on a Beckman System Gold HPLC instrument equipped with a Tosohaas G3000SWXL size exclusion column and a UV detector. Samples of 30 μL were introduced to the column by a constant flow of 1 mL/min mobile phase. Each sample ran for 15 min, with carbonic anhydrase eluting between 6 and 10 min, depending on its molecular weight and buffer. Protein was observed at the exit of the column via absorbance at 280 nm. For samples that did not contain large submicron or micron-sized aggregates (which do not pass through the column), the total chromatogram areas at 280 nm were consistent to within 2−3% during the entire refolding process, indicating that the extinction coefficients of different-sized aggregates did not vary significantly on a mass basis. A mixture of lysozyme, carbonic anhydrase, and bovine serum albumin (monomer and dimer) was used as a standard to calibrate molecular weight to retention time. Using this calibration curve and the breakthrough time of the column, the largest multimer that could pass through the column was a 15-mer. When significant mass was missing from a chromatogram, large multimers were quantitated by difference. The presence of large multimers was confirmed via turbidity or dynamic light scattering for each buffer. The instrument was cleaned with 30 μL injections of 4 M GuHCl, a denaturing

[1] Abbreviations: CA, carbonic anhydrase; NHS, *N*-hydroxysuccinimide; EDC, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide; pNP^-, *p*-nitrophenolate ion; pNPA, *p*-nitrophenyl acetate.

concentration found to dissociate and elute precipitates and large soluble carbonic anhydrase multimers.

Modeling of Association and Dissociation. Transfer free energies for pairs of proteins into 1 M arginine hydrochloride and 1 M guanidinium chloride solutions were computed by a method described previously (*15*). Associating proteins were modeled as spheres of radius 20 Å or as planes of surface area 400π Å^2. (While these shapes may seem like drastic approximations, interaction parameters used below to calculate additive effects were obtained from all-atom molecular simulation data.) The distance between the surfaces of the proteins in any configuration was defined as the reaction coordinate, x, for association and dissociation. The associated state was taken to be the point at which the proteins are in contact with each other ($x = 0$), the dissociated state at infinite separation, and the transition state at a separation distance of 6 Å, or about one shell of water around each protein.

The free energy and the activation free energy of association were defined to be −8 and 2 kcal/mol, respectively. An empirical reaction coordinate free energy surface between these points was constructed from Gaussian functions for the dimer and transition states and an inverse sixth power repulsive term ($x < 0$). The exact function used was

$$\mu = -9.05e^{-(x+0.22/3)2} + 1.98e^{-(x-6/2)2} + \left(\frac{15}{x+15}\right)^6 \quad (3)$$

where μ is the free energy.

Additive-induced perturbations to this free energy function were computed via (*15*):

$$\Delta\mu_P^{tr} = -RTc_X\int(e^{-\langle U_{XP}\rangle/RT} - e^{-\langle U_{WP}\rangle/RT})\,dV \quad (4)$$

where $\Delta\mu_P^{tr}$ is the transfer free energy, RT is the gas constant times absolute temperature, c_X is the additive concentration, $\langle U_{XP}\rangle$ is the additive−protein potential of mean force, $\langle U_{WP}\rangle$ is the water−protein potential of mean force, and the integral is over the solvent volume. The potentials of mean force were modeled as exponential-6 potentials and fit to radial distribution data obtained from all-atom molecular dynamics simulation (*16*). The model for water was taken directly from ref *15*. Guanidinium was modeled as urea from the same reference, but with double the free energy change, since protein free energy effects due to guanidinium chloride are on average double that of urea (*29*). Arginine was modeled as having a characteristic radius of 4 Å and no effect on the free energy of the dissociated state.

RESULTS

Effect on Globular Protein Association. Surface plasmon resonance experiments were conducted to measure the effect of added ArgHCl and GuHCl on the kinetics of globular protein association and dissociation versus an equimolar salt control (NaCl). A typical experimental data set for a binding interaction at one buffer condition is shown in Figure 1.

The data set shown in the figure is a composition of 8 different concentration runs plus replicates, for a total of 16 runs. At $t = 140$ s, the flow cell with immobilized anti-insulin was exposed to a constant concentration of insulin in the range of 2−188 nM for 3 min. During this 3 min, the antibody and antigen were free to associate and dissociate.

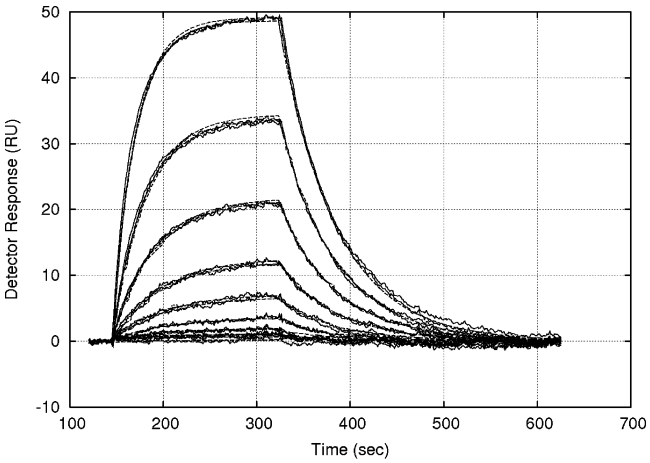


FIGURE 1: Biacore 3000 surface plasmon resonance data for insulin binding to immobilized anti-insulin. Raw binding data (solid curves) are shown with a three-parameter, least-squares fit to all of the data (dashed curves). The detector response is proportional to the mass of antigen bound to the antibody immobilized in the flow cell.

Table 1: Effect of Arginine on Association and Dissociation Rate Constants for Insulin with Monoclonal Antibodies[a]

buffer additive	k_a ($M^{-1}s^{-1}$)	k_d (s^{-1})	K_D (μM)	k_a/k_{a0}	k_d/k_{d0}
0.5 M NaCl	4.4×10^4	1.4×10^{-2}	0.32		
0.5 M ArgHCl	1.2×10^4	2.2×10^{-2}	1.8	0.27	1.6
0.5 M GuHCl	4.0×10^4	9.4×10^{-2}	2.4	0.91	6.7

[a] The base buffer was Biacore HBS-EP (10 mM HEPES, 0.15 M NaCl, 3 mM EDTA, 0.005% polysorbate 20, pH 7.4). k_{a0} and k_{d0} are the association and dissociation rate constants in HPS-EP + 0.5 M NaCl. $K_D \equiv k_d/k_a$. The estimated error in the absolute values of k_a and k_d is 15%.

The net reaction is the binding of free antigen in solution, resulting in an increase in detector response proportional to the mass of antigen bound. At $t = 320$ s, the insulin concentration in the flow cell inlet is returned to zero, and the bound antigen then dissociates from the surface. All 16 runs were simultaneously fit to a binding model by minimizing the squared residuals to yield the association and dissociation rate constants, k_a and k_d. This process was repeated to yield association, dissociation, and equilibrium constant data for the model systems in various buffers as shown in Table 1.

Relative to the 0.5 M NaCl control, 0.5 M GuHCl significantly increases the dissociation rate of insulin and anti-insulin and has an insignificant effect on the association rate. This effect of GuHCl on dissociation rate is consistent with its well-known behavior as a strong denaturant. Small denaturants such as guanidinium chloride and urea bind uniformly to protein surfaces and thermodynamically favor protein states which have the largest solvent-accessible area, such as denatured states (in folding equilibria) and dissociated states (in association equilibria). Since GuHCl does not significantly affect the rate of association of insulin and anti-insulin, it is likely that the association transition state does not have a significantly different solvent-accessible area than the dissociated state.

MECHANISTIC INTERPRETATION

In the preceding section, we observed that arginine slowed protein−protein association and accelerated dissociation,

while guanidinium accelerated dissociation and had little effect on association (Table 1). Here, we wish to relate these observations to a mechanistic model of additive effects on protein association reactions.

We begin by considering the change in a protein reaction rate due to an additive:

$$k = k_0 e^{(\Delta\mu_P^{tr} - \Delta\mu_P^{tr,\ddagger})/RT} \quad (5)$$

where k is the rate constant in the presence of an additive, k_0 is the same rate constant the absence of the additive, $\Delta\mu_P^{tr}$ is the transfer free energy of the reactant into the additive solution, $\Delta\mu_P^{tr,\ddagger}$ is the transfer free energy of the transition state into the additive solution, R is the gas constant, and T is the absolute temperature. The effect of a particular additive enters into the above equation entirely through the difference in the transfer free energies.

When a high concentration of an additive (>0.1 M) is required to have a significant effect on a protein reaction rate or equilibrium constant, such as has been observed in this study for arginine and guanidinium (data at low concentration not shown), the strength of the additive effect can be termed "weak". If, in addition to being weak, the additive interacts with the protein at a large number of sites distributed uniformly over the protein's surface, or does not act in a site-specific manner, the transfer free energy due to the additive is proportional to the solvent-accessible area of the protein (a_P) and an additive-dependent constant (γ_X) related to the preferential binding coefficient (*15*, *30*−*34*):

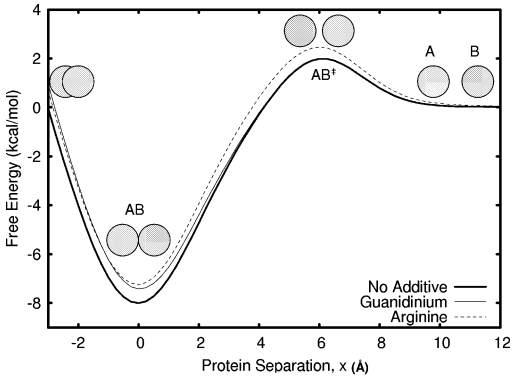
$$\Delta\mu_P^{tr} = -RT\gamma_X a_P c_X \quad (6)$$

where c_X is the concentration of additive. Analogous expressions are frequently used to model the effects of additives such as guanidinium, trehalose, and sorbitol.

The experimental observation that guanidinium does not significantly alter the rate of association of insulin and anti-insulin suggests that the surface area of the pair of molecules accessible to guanidinium does not change significantly from the dissociated state to the association transition state. If this is the case, and if arginine interacts with proteins in the same way that guanidinium does, it should not be possible for arginine, acting in a weak and nonspecific manner, to exert any effect either, yet we observe 0.5 M arginine induces approximately a factor of 3 depression in the association rate (Table 1). This suggests that arginine acts via a mechanism distinct from that of guanidinium.

Recently, we have developed statistical−mechanical methods to calculate transfer free energies for association and dissociation reactions in general (*15*, *16*). This led to the understanding that additive−protein steric effects can influence the energetics of association and dissociation reactions. Specifically, if an additive is much larger than water but does not significantly affect the free energy of dissociated protein molecules, the additive will increase the activation free energy for the molecules to associate. This steric effect, which we call "the gap effect", slows protein association and may either speed or slow dissociation.

The model developed in ref *15* can be used to calculate the effects of guanidinium and arginine as described in Materials and Methods. The results of such a calculation are shown in Figure 2.



Additive	k_a	k_d
Arginine	-	+
Guanidinium	0	+

FIGURE 2: Calculated free energies are shown for a pair of 20 Å spherical proteins into 1 M arginine and guanidinium solutions as a function of the separation between the proteins. Free energies are normalized to the free energy of the dissociated pair (x > 10 Å). The gray spheres indicate the geometry of the protein pair as a function of protein separation. The table at the bottom shows the magnitudes of the changes in the association and dissociation rate constants (k_a and k_d).

In the presence of arginine, the model predicts that the free energy of the transition state will increase relative to the dissociated state. This causes the association rate constant to decrease. Inversely, the free energy of the associated state increases relative to the free energy of the transition state, causing the dissociation rate constant to increase. In stark contrast to the arginine effect, the presence of guanidinium has little effect on the transition state free energy relative to the dissociated state; hence guanidinium has no effect on the association rate constant. The associated state free energy, however, increases relative to the transition state, causing the dissociation rate constant to increase. All of these effects are qualitatively consistent with the changes in the measured rate constants for insulin and anti-insulin (Table 1).

Using this model and an analogous model in which the proteins are approximated as planar surfaces (*15*), the range of association rate effects caused by arginine can be quantitated. The spherical and planar models give a range of 0.8−2.8 kcal mol^{-1} M^{-1} for the maximum increase in the free energy barrier to association. For 0.5 M arginine solution, this is 0.4−1.4 kcal/mol, or a rate effect of $k_a/k_{a0} = e^{-\Delta\Delta\mu_\ddagger^{tr}/RT} = 0.51-0.10$. This range covers the experimentally observed value for the association rate depression of insulin and anti-insulin at 0.5 M ArgHCl ($k_a/k_{a0} = 0.27$, Table 1).

Effect on Refolding of Carbonic Anhydrase. To assess whether the effects of arginine and guanidinium on globular protein association reactions carry over to a more complex aggregation situation, we examined the effects of equimolar amounts of NaCl, GuHCl, and ArgHCl on the refolding of carbonic anhydrase II (CA). CA is a natural enzyme that is known to aggregate during refolding.

In previous studies in our laboratory and others (*27*, *28*, *35*−*40*), carbonic anhydrase II was found to refold from a denatured state by sequential formation of a molten intermediate state (M), a near-native conformation that has no

biological activity (I), and finally the native state (N):

$$U \rightarrow M \rightarrow I \rightarrow N \quad (7)$$

Cleland showed that the molten intermediate (M) can aggregate to form dimers and higher mers (*38*):

$$M \rightarrow A_2 \rightarrow \text{etc.} \quad (8)$$

In 1.0 M GuHCl and at low concentration of carbonic anhydrase (less than 30 μM), the formation of small mers was reversible, leading to yields of native protein approaching 100%. At lower GuHCl concentrations, formation of large aggregates occurred, resulting in significant losses of CA. At long times (hours to days), the only aggregate species observed were small multimers and very large, micron-sized aggregates. These observations lead to the following two predictions about the performance of ArgHCl and GuHCl as solution additives:

(1) The reversibility of small multimer formation implies that early association reactions are at least partially equilibrium-controlled. Then, since ArgHCl and GuHCl shift equilibrium toward the smaller mers (Table 1), they both should promote formation of the native protein during refolding. This was probed experimentally by measuring the native protein concentration as a function of refolding buffer conditions.

(2) The absence of intermediate-sized aggregates at long times implies that CA aggregation proceeds via a nucleation-dependent polymerization mechanism where a small multimer is the nucleus. After formation of the nucleus, association is rapid and dissociation is negligible. Since ArgHCl deters association, arginine should decrease the average aggregate size and molecular weight in this regime. Conversely, since guanidinium chloride affects the association equilibrium by increasing the dissociation rate, it will have a negligible effect on this regime of aggregation. This was probed experimentally by measuring the multimer distribution as a function of refolding buffer conditions via size exclusion HPLC, as described below.

Yield of Native Protein. Esterase activity assays were performed as a function of initial unfolded protein concentration and buffer composition to determine how equimolar concentrations of NaCl, ArgHCl, and GuHCl each affected refolding yield (Figure 3).

It was observed that the yield of active protein as a function of buffer additive increased in the following order: NaCl $\ll$ ArgHCl $<$ GuHCl.

If association and aggregation can account for the majority of the loss of native protein, then it should be possible to model the yield of native protein as a function of the initial protein concentration and a parameter characterizing the competition between refolding and aggregation (*41*). Assuming the unfolded protein rapidly collapses to the molten intermediate when introduced into refolding conditions (*35*), refolding and aggregation from the molten state can be modeled as being in direct kinetic competition (*3*):

$$N \xleftarrow{k_r} M \xrightarrow{k_{agg}} \text{aggregates} \quad (9)$$

where k_r is the refolding rate constant and k_{agg} is the aggregation rate constant.

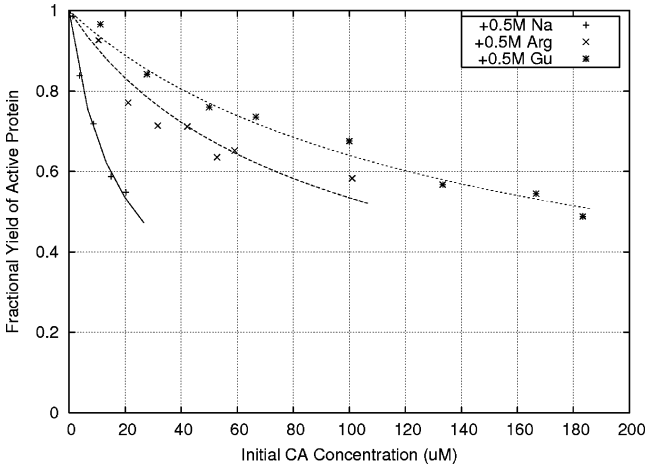


FIGURE 3: Effect of refolding buffer composition on carbonic anhydrase refolding yield. The points are experimental esterase activity data, and the lines are the best fit to a one-parameter, first- versus second-order kinetic model (eq 11).

Since refolding is a unimolecular reaction, it is expected that the refolding reaction is first order. The kinetic order of the macroscopic aggregation reaction, however, cannot be predicted in advance. In an earlier study of carbonic anhydrase refolding via dynamic light scattering, Cleland and Wang (*40*) proposed a 2.6-power relationship between initial protein concentration and monomer depletion rate at short times (30−60 s). Thus, we expect a reaction order of between 2 and 3 to be applicable in this case. Model cases for aggregation reaction orders of 2 and 3 were fit to the data and revealed that a macroscopic second-order aggregation reaction gave a much better fit for all three buffer conditions. The activity data with added 0.5 M GuHCl and 0.5 M ArgHCl are suggestive of slightly higher inactivation order than the added 0.5 M NaCl case, but because of the uncertainty ($\pm$5%) in the esterase activity data, it is not possible to determine the reaction order to better than about $\pm$0.5 by direct fitting.

For a second-order aggregation reaction, the yield of native protein is

$$\text{yield} = \frac{k_r}{k_{agg}[U]_0} \ln\left(1 + \frac{k_{agg}[U]_0}{k_r}\right) \quad (10)$$

where $[U]_0$ is the initial concentration of unfolded protein. Since the constants k_r and k_{agg} appear only as a quotient, they can be condensed to a single "refolding selectivity parameter", $\alpha \equiv k_r/k_{agg}$, having units of concentration and resulting in a working equation:

$$\text{yield} = \frac{\alpha}{[U]_0} \ln\left(1 + \frac{[U]_0}{\alpha}\right) \quad (11)$$

Each of the data sets in Figure 3 was fit to the above model equation, yielding the values of α shown in Figure 2. The functional forms of the model at these values of α are shown in Figure 3.

The parameter α is a direct measure of the performance of a refolding additive. It is equal to the concentration of unfolded protein at which the refolding yield will be ln(2), or about 70%.

Table 2: Refolding Selectivity Parameters (α) and Parameters Relative to 0.5 M NaCl (α/α_0) for Refolding of Carbonic Anhydrase with Three Different Buffer Additives[a]

additive	α (μM)	α/α_0
0.5 M NaCl	9.3	1
0.5 M ArgHCl	47	5.0
0.5 M GuHCl	77	8.2

[a] The base buffer composition was 0.5 M GuHCl.

Table 3: HPLC Analysis of Multimers Formed during Refolding of Carbonic Anhydrase in Different Buffers, Expressed as a Percentage of the Total Carbonic Anhydrase[a]

time (min)	M	A_2	A_{3-5}	A_{6-15}	large (%)
(a) Additive: 0.5 M NaCl; $[U]_0 = 5$ μM					
2	56	0	0	0	44
20	56	0	0	0	44
38	58	0	0	0	42
(b) Additive: 0.5 M ArgHCl; $[U]_0 = 20$ μM					
2	22	30	25	21	2
20	54	7	14	26	−1
38	62	4	11	24	−1
1500	80	0	0	19	1
(c) Additive: 0.5 M GuHCl; $[U]_0 = 20$ μM					
2	42	39	8	0	11
20	82	3	6	0	9
38	85	1	5	0	9
1500	89	0	2	0	9

[a] The time reported is the time between injection onto the HPLC column and dilution of the denatured carbonic anhydrase into the refolding buffer. The base refolding buffer contained 0.5 M GuHCl. M indicates monomer, and A_{i-j} indicates multimers of mer number i through j. The amount of "large" multimers which do not pass through the column is inferred from the difference between the amount of protein injected onto the column and the total chromatogram area. The reproducibility of any peak area determination from experiment to experiment is ±1%.

The relative refolding selectivity values (α/α_0) for ArgHCl and GuHCl indicate that both of these additives promote refolding. This supports the notion that formation of irreversible aggregates is at least partially equilibrium-controlled. The refolding selectivity values are also qualitatively consistent equilibrium shifts effects seen in globular protein association (Table 1).

Multimer Distribution. Size exclusion HPLC experiments were performed to analyze the distribution of multimers formed during refolding. CA was refolded with three different additives, 0.5 M NaCl, 0.5 M GuHCl, and 0.5 M ArgHCl, relative to a base refolding buffer of 0.5 M GuHCl, as done in the esterase activity assays above. The 0.5 M NaCl refolding experiment was performed at 4-fold lower concentration (5 μM) because visible aggregates were formed within seconds at concentrations comparable to the other two experiments (20 μM). Other than this protein concentration difference, these experiments allow direct comparison of how an additional 0.5 M of the three different cations affects refolding.

After initiating refolding by diluting denatured CA with an appropriate buffer, refolding was allowed to proceed for at least 2 h before performing HPLC. The samples were not filtered prior to introduction into the HPLC column. The molecular weight distributions observed are shown in Table 3.

In 0.5 M NaCl, the refolded carbonic anhydrase is partitioned entirely between monomers and large aggregates, with no significant mass observed in intermediate species. With 0.5 M ArgHCl or GuHCl added, the yield of monomeric protein is significantly increased, consistent with the observation of a larger native protein yield in the previous section.

In all three refolding buffers, significant amounts of large aggregates form which do not dissociate into monomeric protein. With longer refolding times, the average aggregate molecular weight and hydrodynamic radii continue to increase and monomer is slowly depleted (data not shown). This implies that the native protein and large aggregate states are separated by a large free energy barrier.

The average aggregate molecular weight (ignoring the monomer) is lowest in 0.5 M ArgHCl, despite the fact that 0.5 M GuHCl results in the highest yield of native protein. Since intermediate aggregates (A_{6-15}) are not observed in 0.5 M NaCl or 0.5 M GuHCl, but larger aggregates are observed, association must be rapid through the intermediate size range in these buffers. Because dissociation is negligible in such a regime, additives such as guanidinium that affect association equilibria through the dissociation rate cannot deter association here. In contrast, arginine, which slows association reactions, can deter formation of higher mers and ultimately leads to a lower average aggregate molecular weight than GuHCl or NaCl.

This type of difference may have important consequences when comparing the performance of different buffer additives via simple surrogate assays. As seen in the differences in yield and aggregate molecular weight distribution between the refolding buffer additives ArgHCl and GuHCl (Figure 3), a decrease in the average aggregate molecular weight may not be indicative of increased refolding yield. Thus, simple aggregation assays such as turbidity and dynamic light scattering, which roughly measure the amount of large particles in solution, will also not correlate with yield when comparing additives that affect association with those that affect dissociation.

CONCLUSIONS

The presence of arginine in solution was shown to slow protein−protein association reactions in two model systems: the association of insulin with a monoclonal antibody and the association of folding intermediates and aggregates of carbonic anhydrase II (CA). In CA refolding, arginine promoted formation of the native protein and decreased the average molecular weight of CA aggregates.

The denaturant guanidinium chloride (GuHCl), which is also used to dissolve aggregates and deter aggregation in certain situations, exhibited significantly different kinetic behavior than ArgHCl. GuHCl significantly increased the dissociation rate constant of insulin and anti-insulin and had a negligible effect on their association rate. GuHCl also significantly increased CA refolding yield, but because of the difference in kinetic effects, GuHCl had a smaller effect on reducing the average molecular weight of CA aggregates than ArgHCl.

The magnitudes of the observed effects were quantitatively consistent with gap effect theory (*15*). Arginine can be modeled as a "neutral crowder", an additive that is larger

than water but has a negligible effect on the free energy of isolated protein molecules.

The beneficial effect of arginine on protein refolding arises because it slows protein association reactions. Thus, in addition to being a useful refolding buffer additive, arginine should prevent aggregation in any application where aggregation exhibits second-order or higher order kinetics.

ACKNOWLEDGMENT

B.M.B. thanks Aaron Risinger and Andre Ditsch for expert technical assistance and Compound Therapeutics, Inc., and Prof. T. Alan Hatton for the use of equipment.

REFERENCES

1. Lilie, H., Schwarz, E., and Rudolph, R. (1998) Advances in refolding of proteins produced in *E. coli*, *Curr. Opin. Biotechnol. 9*, 497−501.
2. Dill, K. A. (1990) Dominant forces in protein folding, *Biochemistry 29*, 7133−7155.
3. Zettlmeissl, G., Rudolph, R., and Jaenicke, R. (1979) Reconstitution of lactic dehydrogenase. noncovalent aggregation vs. reactivation. 1. physical properties and kinetics of aggregation, *Biochemistry 18*, 5567−5571.
4. Wang, W. (1999) Instability, stabilization, and formulation of liquid protein pharmaceuticals, *Int. J. Pharmacol. 185*, 129−188.
5. Cleland, J. L., Powell, M. F., and Shire, S. J. (1993) The development of stable protein formulations: A close look at protein aggregation, deamidation, and oxidation, *Crit. Rev. Ther. Drug Carrier Syst. 10*, 307−377.
6. Arakawa, T., and Tsumoto, K. (2003) The effects of arginine on refolding of aggregated proteins: Not facilitate refolding, but suppress aggregation, *Biochem. Biophys. Res. Commun. 304*, 148−152.
7. Taneja, S., and Ahmad, F. (1994) Increased thermal stability of proteins in the presence of amino acids, *Biochem. J. 303*, 147−153.
8. Shiraki, K., Kudou, M., Fujiwara, S., Imanaka, T., and Takagi, M. (2002) Biophysical effect of amino acids on the prevention of protein aggregation, *J. Biochem. 132*, 591−595.
9. Rudolph, R., Fischer, S., and Mattes, R. (1985) Patent DE3537708, Process for the Activation of tPA after Expression in Prokaryotic Cells.
10. Arora, D., and Khanna, N. (1996) Method for increasing the yield of properly folded recombinant human gamma interferon from inclusion bodies, *J. Biotechnol. 52*, 127−133.
11. Armstrong, N., de Lencastre, A., and Gouaux, E. (1999) A new protein folding screen: Application to the ligand binding domains of a glutamate and kainate receptor and to lysozyme and carbonic anhydrase, *Protein Sci. 8*, 1475−1483.
12. Rinas, U., Risse, B., Jaenicke, R., Abel, K.-J., and Zettlmeissl, G. (1990) Denaturation-renaturation of the fibrin-stabilizing factor xiii a-chain isolated from human placenta, *Biol. Chem. Hoppe-Seyler 371*, 49−56.
13. Buchner, J., and Rudolph, R. (1991) Renaturation, purification and characterization of recombinant fab-fragments produced in *Escherichia coli*, *Biotechnology 9*, 157−162.
14. Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Philo, J. S., and Arakawa, T. (2004) Role of arginine in protein refolding, solubilization, and purification, *Biotechnol. Prog. 20*, 1301−1308.
15. Baynes, B. M., and Trout, B. L. (2004) Rational design of solution additives for the preventing of protein aggregation, *Biophys. J. 87*, 1631−1639.
16. Baynes, B. M., and Trout, B. L. (2003) Proteins in mixed solvents: A molecular-level perspective, *J. Phys. Chem. B 107*, 14058−14067.
17. Timasheff, S. N. (1998) Control of protein stability and reactions by weakly interacting cosolvents: The simplicity of the complicated, *Adv. Protein Chem. 51*, 355−431.
18. Colombo, M. F., Rau, D. C., and Parsegian, A. (1992) Protein solvation in allosteric regulation: A water effect on hemoglobin, *Science 256*, 655−659.
19. Kirkwood, J. G., and Buff, F. P. (1951) The statistical mechanical theory of solutions. i, *J. Chem. Phys. 19*, 774−777.
20. Shimizu, S. (2004) Estimating hydration changes upon biomolecular reactions from osmotic stress, high pressure, and preferential hydration experiments, *Proc. Natl. Acad. Sci. U.S.A. 101*, 1195−1199.
21. Shimizu, S., and Smith, D. J. (2004) Preferential hydration and the exclusion of cosolvents from protein surfaces, *J. Chem. Phys. 121*, 1148−1154.
22. Smith, P. E. (2004) Local chemical potential equalization model for cosolvent effects on biomolecular equilibria, *J. Phys. Chem. B 108*, 16271−16278.
23. Minton, A. P. (1997) Influence of excluded volume upon macromolecular structure and associations in crowded media, *Curr. Opin. Biotechnol. 8*, 65−69.
24. Linder, R., and Ralston, G. (1995) Effects of dextran on the self-association of human spectrin, *Biophys. Chem. 57*, 15−25.
25. Pocker, Y., and Stone, J. T. (1967) The catalytic versatility of erythrocyte carbonic anhydrase. iii. kinetic studies of the enzyme-catalyzed hydrolysis of *p*-nitrophenyl acetate, *Biochemistry 6*, 668−678.
26. Pocker, Y., and Biswas, S. B. (1981) Self-association of insulin and the role of hydrophobic bonding: A thermodynamic model of insulin dimerization, *Biochemistry 20*, 4354−4361.
27. Cleland, J. L., Hedgepeth, C., and Wang, D. I. C. (1992) Polyethylene glycol enhanced refolding of bovine carbonic anhydrase b, *J. Biol. Chem. 267*, 13327−13334.
28. Wetlaufer, D. B., and Xie, Y. (1995) Control of aggregation in protein refolding: A variety of surfactants promote renaturation of carbonic anhydrase ii, *Protein Sci. 4*, 1535−1543.
29. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant *m* values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding, *Protein Sci. 4*, 2138−2148.
30. Lee, J. C., and Timasheff, S. N. (1974) Partial specific volumes and interactions with solvent components of proteins in guanidine hydrochloride, *Biochemistry 13*, 257−265.
31. Gekko, K., and Timasheff, S. N. (1981) Mechanism of protein stabilization by glycerol: Preferential hydration in glycerol-water mixtures, *Biochemistry 20*, 4667−4676.
32. Arakawa, T., and Timasheff, S. N. (1985) The stabilization of proteins by osmolytes, *Biophys. J. 47*, 411−414.
33. Timasheff, S. N. (2002) Protein−solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components, *Proc. Natl. Acad. Sci. U.S.A. 99*, 9721−9726.
34. Davis-Searles, P. R., Saunders, A. J., Erie, D. A., Winzor, D. J., and Pielak, G. J. (2001) Interpreting the effects of small uncharged solutes on protein-folding equilibria, *Annu. Rev. Biophys. Biomol. Struct. 30*, 271−306.
35. Semisotnov, G., Rodionova, N. A., Kutyshenko, V. P., Ebert, B., Blanck, J., and Ptitsyn, O. B. (1987) Sequential mechanism of refolding of carbonic anhydrase b, *FEBS Lett. 224*, 9−13.
36. Semisotnov, G. V., Uversky, V. N., Sokolovsky, I. V., Gutin, A. M., Razgulyaev, O. I., and Rodionova, N. A. (1990) 2 slow stages in refolding of bovine carbonic anhydrase-b are due to proline isomerization, *J. Mol. Biol. 213*, 561−568.
37. Dolgikh, D. A., Kolomiets, A. P., Bolotina, I. A., and Ptitsyn, O. B. (1984) Molten-globule state accumulates in carbonic anhydrase folding, *FEBS Lett. 165*, 88−92.
38. Cleland, J. L. (1991) Mechanisms of Protein Aggregation and Refolding, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge.
39. Cleland, J. L., and Wang, D. I. C. (1992) Transient association of the first intermediate during the refolding of bovine carbonic anhydrase b, *Biotechnol. Prog. 6*, 97−103.
40. Cleland, J. L., and Wang, D. I. C. (1990) Refolding and aggregation of bovine carbonic anhydrase b: Quasi-elastic light scattering analysis, *Biochemistry 29*, 11072−11078.
41. Hevehan, D. L., and Clark, E. D. B. (1997) Oxidative renaturation of lysozyme at high concentrations, *Biotechnol. Bioeng. 54*, 221−230.

BI047528R