Reduced BPTI is collapsed. A pulsed field gradient NMR study of unfolded and partially folded bovine pancreatic trypsin inhibitor

HONG PAN,¹ GEORGE BARANY,² AND CLARE WOODWARD¹

¹Department of Biochemistry, University of Minnesota, St. Paul, Minnesota 55108

²Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455

(RECEIVED February 24, 1997; ACCEPTED May 7, 1997)

Abstract

Pulsed field gradient NMR was used to measure the hydrodynamic behavior of unfolded variants of bovine pancreatic trypsin inhibitor (BPTI). The unfolded BPTI species studied were $[R]_{\text{Abu}}$, at pH 4.5 and pH 2.5, and unfolded $[14-38]_{\text{Abu}}$, at pH 2.5. These were prepared by chemical synthesis. $[R]_{\text{Abu}}$ is a model for reduced BPTI; all cysteine residues are replaced by α -amino-*n*-butyric acid (Abu). $[14-38]_{\text{Abu}}$ retains cysteines 14 and 38, which form a disulfide bond, while the other cysteine residues are replaced by Abu. In the PFG experiments, the diffusion coefficient is measured as a function of protein concentration, and the value of D° —the diffusion coefficient extrapolated to infinite dilution—is determined. From D° , a value of the hydrodynamic radius, R_h , is computed from the Stokes-Einstein relationship. At pH 4.5, $[R]_{\text{Abu}}$ has an R_h value significantly less than the value calculated for a random coil, while at pH 2.5 the experimental R_h value is the same as for a random coil. In view of the changes in NMR-detected structure of $[R]_{\text{Abu}}$ at pH 4.5 versus pH 2.5 (Pan H, Barbar E, Barany G, Woodward C. 1995. Extensive non-random structure in reduced and unfolded bovine pancreatic trypsin inhibitor. *Biochemistry* 34:13974–13981), the collapse of reduced BPTI at pH 4.5 may be associated with the formation of non-native hydrophobic clusters of pairs of side chains one to three amino acids apart in sequence. The diffusion constant of $[14-38]_{\text{Abu}}$ was also measured at pH 4.5, where the protein is partially folded. An increase in hydrodynamic radius of partially folded $[14-38]_{\text{Abu}}$, relative to native BPTI, is similar to the increase in radius of gyration measured for other proteins under “molten globule” conditions.

Keywords: bovine pancreatic trypsin inhibitor; protein folding; protein hydrodynamics; pulsed field gradient NMR; unfolded proteins

The structure of polypeptides at the onset of folding is a fundamental issue in protein folding. Two questions concerning unfolded proteins are of special interest: The tendency of the unfolded chain to form non-random conformations, and the extent of collapse of the polypeptide prior to formation of secondary structure.

Initial collapse of the polypeptide would reduce the conformational space to be searched in a thermodynamic model of folding, thereby providing a partial solution to the “Levinthal paradox” (Dill, 1985; Sali et al., 1995); in addition, hydrophobic collapse may facilitate the formation of secondary structure (Dill et al., 1995; Dill & Chan, 1997). Non-random conformations favored in the unfolded ensemble may serve as folding initiation sites, e.g., in hydrophobic zippering (Dill et al., 1993).

We have undertaken structural and hydrodynamic characterization of unfolded BPTI variants. These BPTI species, summarized in the diagrams and legend of Figure 1, are particularly useful models for unfolded proteins because they are obtained without introducing extremes of pH, temperature, and/or chemical denaturants. The tendency of reduced, unfolded BPTI chains to sample non-random conformations (Fig. 1) was demonstrated in previous NMR studies (Pan et al., 1995). The NMR data provide a likely candidate for the site of a non-random conformation favored in the reduced BPTI ensemble which may serve

Reprint requests to: Clare Woodward, Department of Biochemistry, College of Biological Science, University of Minnesota, 1479 Gortner Ave., St. Paul, Minnesota 55108; e-mail: clare@biosci.cbs.umn.edu.

Abbreviations: Abu, α -amino-*n*-butyric acid; BPTI, bovine pancreatic trypsin inhibitor; $[R]_{\text{Abu}}$, BPTI in which all cysteines are replaced with Abu; $[14-38]_{\text{Abu}}$, BPTI in which cysteines 5, 30, 51, and 55 are replaced by Abu; SNase, staphylococcal nuclease; RNase A, ribonuclease A; BLG, bovine β -lactoglobulin; PGK, phosphoglycerate kinase; BCAB, bovine carbonic anhydrase B; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; PFG, pulsed-field gradient; CD, circular dichroism; HPLC, high performance liquid chromatography; R_h , hydrodynamic radius; R_g , radius of gyration; D , diffusion constant; D° , diffusion constant extrapolated to infinite dilution.

Fig. 1. Unfolded and partially folded variants of BPTI. The diagrams are schematic representations of dynamic ensembles of interconverting conformations detected by NMR and CD (Barbar et al., 1995, 1996, 1997; Ferrer et al., 1995; Pan et al., 1995). The native structure at lower right is shown for reference. (Top left) $[R]_{\text{Abu}}$ is reduced BPTI with no disulfide bonds and all six cysteines replaced by α -amino-*n*-butyric acid (Abu). $[R]_{\text{Abu}}$ at pH 4.5 and 1–5° lacks regular secondary structure; nevertheless, it samples non-random conformations reported by numerous, non-native, backbone–backbone and side-chain–side-chain NOEs (Pan et al., 1995). Non-native, turn-like backbone conformations are indicated especially in a stretch of residues (29–35), which in native BPTI fold into a strand of β -sheet. In addition, non-native aliphatic–aromatic interactions are indicated by ($i, i \pm 3$) NOEs between side chains in stretches of residues (18–24 and 29–35), which in native BPTI fold into the central anti-parallel β -sheet. This suggests that the non-native turn-like conformations are stabilized by non-native hydrophobic interactions; hydrophobic clustering is consistent with ANS fluorescence in the presence of $[R]_{\text{Abu}}$ (Ferrer et al., 1995). $[R]_{\text{Abu}}$ at pH 2.5 and 1 °C has diminished non-native aromatic–aliphatic NOEs and behaves more like a random coil than at pH 4.5. (Top right) $[14-38]_{\text{Abu}}$ is BPTI with one disulfide bond, between 14–38, but with the other four cysteines replaced by Abu. At pH 2.5, $[14-38]_{\text{Abu}}$ is unfolded. Like $[R]_{\text{Abu}}$, $[14-38]_{\text{Abu}}$ at pH 2.5 has a CD spectrum very similar to random coil (Ferrer et al., 1995); however, there are significant differences between the NMR behavior of an unfolded $[14-38]_{\text{Abu}}$ variant and $[R]_{\text{Abu}}$ (Barbar et al., 1996; summarized in text). (Bottom left) When the pH is raised from 2.5 to 4.5 at 1–10 °C, unfolded $[14-38]_{\text{Abu}}$ undergoes a transition to an ensemble of partially folded conformations undergoing segmental motions (Barbar et al., 1995; summarized in text). For partially folded $[14-38]_{\text{Abu}}$ thermal unfolding is cooperative, but is clearly not two-state (Barbar et al., 1997). Like $[R]_{\text{Abu}}$, partially folded $[14-38]_{\text{Abu}}$ has solvent-exposed, clustered hydrophobes indicated by ANS fluorescence (Ferrer et al., 1995). A strong NOE associated with the reverse turn between β -strands 18–24 and 29–35 is the only native-like NOE observed in $[R]_{\text{Abu}}$; further, it is observed in NMR spectra of $[R]_{\text{Abu}}$ at pH 2.5, of unfolded $[14-38]_{\text{Abu}}$ and of partially folded $[14-38]_{\text{Abu}}$, making it a likely candidate for an initiation site of folding.

to initiate folding, namely, the reverse turn between the anti-parallel strands of the central β -sheet (Barbar et al., 1995; Pan et al., 1995). Here we examine the extent of collapse of reduced BPTI using a hydrodynamic method.

Hydrodynamic studies of denatured states have been reported recently for several proteins using solution small-angle X-ray scattering (SAXS) methods. Size and shape parameters of molten globule and/or unfolded states have been determined by SAXS for

myoglobin, streptomyces subtilisin inhibitor, cytochrome *c*, truncated staphylococcal nuclease, α -lactalbumin, hen egg lysozyme and others (Izumi et al., 1983; Dolgikh et al., 1985; Flanagan et al., 1992; Hamada et al., 1993; Kataoka et al., 1993, 1995; Eliezer et al., 1995; Konno et al., 1995; Chen et al., 1996; Semisotnov et al., 1996). Recent developments in time-resolved SAXS techniques, employing a high flux synchrotron source and rapid mixing, make it possible to assess the compactness of proteins during folding (Eliezer et al., 1995; Semisotnov et al., 1996).

We apply a different method, pulsed-field gradient (PFG) NMR, to monitor the hydrodynamic behavior of denatured states. PFG NMR experiments measure the translational diffusion coefficient of a molecule in solution (Altieri et al., 1995). PFG NMR has been applied previously to protein self-association and protein-ligand complex formation (Lennon et al., 1994; Dingley et al., 1995; Mayo et al., 1996). A full study of the condition dependence of native BPTI self-association by PFG NMR was also reported recently (Ilyina et al., 1997). From the concentration dependence of the diffusion coefficient obtained from PFG NMR data, we determine the diffusion coefficient at infinite dilution, D° , and estimate the related hydrodynamic radius, R_h . At pH 4.5, reduced BPTI is not as extended as a statistical random coil of the same number of residues; it is collapsed and its average compactness is about the same as unfolded BPTI constrained by a crosslink between residues 14 and 38. These hydrodynamic results enrich the descriptions, based on NMR analysis, of the dynamic structures of unfolded and partially folded BPTI variants.

Results

For diffusion of a molecule in an isotropic medium, the signal intensity, A , recorded in a PFG NMR experiment, normalized to the signal in the absence of the magnetic field gradient pulses, is given in Stejskal and Tanner (1965).

$$A = \exp(-\gamma^2 G^2 D \delta^2 (\Delta - \delta/3)) \quad (1)$$

where G and δ are the magnitude and duration of the magnetic field gradient pulses, Δ is the time between field gradient pulses, and γ is the gyromagnetic ratio of the observed nucleus. Plots of A as a function of $\gamma^2 G^2 \delta^2 (\Delta - \delta/3)$ are fit to Equation 1 by nonlinear least-square regression analysis to give the apparent translational diffusion coefficient, D , and its standard deviation.

Diffusion coefficients of three unfolded states, and one partially folded state, were measured at 1 °C. Two of the unfolded states are $[R]_{\text{Abu}}$ at pH 4.5 and pH 2.5, and the third is $[14-38]_{\text{Abu}}$ at pH 2.5 (see diagrams and legend to Fig. 1). The partially folded state is $[14-38]_{\text{Abu}}$ at pH ≥ 4.5 .

Representative data at a single protein concentration and pH are given in Figure 2; the curve shows the fit of the data to Equation 1, from which the value of D is obtained. For each variant, D was measured at 1 °C as a function of protein concentration at pH 4.5 and pH 2.5. Figure 3 shows the protein concentration dependence of D for partially folded $[14-38]_{\text{Abu}}$ and unfolded $[R]_{\text{Abu}}$ at pH 4.5; data for native BPTI are also shown. Figure 4 shows the protein concentration dependence of D for unfolded $[14-38]_{\text{Abu}}$ and unfolded $[R]_{\text{Abu}}$ at pH 2.5. First-order dependence of the diffusion coefficient on concentration is observed for all cases in the concentration range covered. For $[R]_{\text{Abu}}$ and $[14-38]_{\text{Abu}}$, the concentration dependence of D is more pronounced at pH 4.5, as compared to pH 2.5. At both pH values, the protein concentration dependence of D is larger for $[R]_{\text{Abu}}$ than for $[14-38]_{\text{Abu}}$.

Fig. 2. Normalized peak intensity, A , of $[14-38]_{\text{Abu}}$ in a PFG NMR experiment plotted as a function of $\gamma^2 G^2 \delta^2 (\Delta - \delta/3)$. Data were fit to the exponential decay function in Equation 1 by nonlinear regression analysis. The 0.5 mM protein sample was maintained at pH 4.5 and 1 °C, conditions under which $[14-38]_{\text{Abu}}$ is partially folded.

The diffusion coefficient at infinite dilution, D° , is estimated by linear least-square fits to the data in Figures 3 and 4. The hydrodynamic radius, R_h , is derived from D° by the Stokes-Einstein relationship:

$$D^\circ = k_B T / 6 \pi \eta R_h \quad (2)$$

where k_B is the Boltzmann constant, and η the viscosity of the solvent. Values of D° and hydrodynamic radii (R_h) are summarized in Table 1. Other methods are available for measurement of diffusion coefficients, e.g., ultracentrifugation and dynamic light scat-

Fig. 3. Diffusion coefficient as a function of protein concentration for native BPTI (squares), partially folded $[14-38]_{\text{Abu}}$ (triangles), and $[R]_{\text{Abu}}$ (circles) at pH 4.5. Each value of D was derived from a PFG NMR experiment conducted at 1 °C. Error bars denote ± 1 SD; for data points without error bars, the error is within the size of the symbol.

Fig. 4. Diffusion coefficient as a function of protein concentration for unfolded $[14-38]_{\text{Abu}}$ (triangles), and $[R]_{\text{Abu}}$ (circles) at pH 2.5 and 1 °C. Error bars denote ± 1 SD; for data points without error bars, the error is within the size of the symbol. To facilitate comparison of $[R]_{\text{Abu}}$ and $[14-38]_{\text{Abu}}$ at pH 2.5 versus 4.5, curves from Figure 3 are reproduced for $[R]_{\text{Abu}}$ at pH 4.5 (lower dotted line) and for $[14-38]_{\text{Abu}}$ at pH 4.5 (upper line).

tering, but the PFG method is particularly useful for small proteins, such as BPTI, at relatively low concentrations.

Discussion

Partially folded and unfolded $[14-38]_{\text{Abu}}$

Partially folded and unfolded BPTI variants have values of D° smaller than native BPTI, as expected for more disordered conformations. For the native protein, D° obtained from extrapolation of the concentration dependence in Figure 3 corresponds to a hy-

drodynamic radius, R_h , of 15.2 ± 0.2 Å (Table 1). This is reasonable for monomeric, native BPTI with a monolayer of hydration, and is in agreement with R_h values obtained from light scattering, 14.9 Å and 13.9 Å (Wills & Georgalis, 1981; Gallagher & Woodward, 1989). Recent PFG NMR studies of native BPTI (Ilyina et al., 1997), indicate that native BPTI begins to dimerize at concentrations above 3 mg/mL and below 7 °C. In Figure 3, the three lowest concentrations of native BPTI are below 3 mg/mL (the very lowest concentration is 0.7 mg/mL), and the extrapolated D° is considered to be the monomer value. In Table 1, the extent to which each variant is more expanded than native is expressed by the percent increase in R_h , i.e., the difference in R_h between variant and native divided by R_h of native.

For partially folded $[14-38]_{\text{Abu}}$, an expansion of 19% over native is consistent with its NMR- and CD-detected structure. At pH ≥ 4.5 , $[14-38]_{\text{Abu}}$ is an ensemble of partially folded conformations undergoing *segmental* motion on a time scale of msec or longer (Fig. 1, lower left). NMR reporter groups in all parts of the molecule indicate that there are two to three conformations in slow exchange, and that conformer populations and interconversion rates vary for different segments of the molecule. The central anti-parallel β -sheet strands are native-like about 80% of the time; multiple, strong non-local NOEs to the side chains of Y21 and Y23 imply that these residues are involved in tertiary interactions stabilizing the native-like core. Apart from the anti-parallel strands, the rest of the protein is much more flexible. *N*- and *C*-terminal segments are the most disordered: They have two conformations in slow exchange but neither conformation is native-like. It is highly significant that the stable, native-like structure is not in the vicinity of the only disulfide crosslink, 14–38; this, along with other data, leads us to conclude that the order of disulfide bond formation is not important in folding from reduced to native BPTI (Barbar et al., 1995). This is because the presence of any single disulfide bond results in the same type of intermediate, one in which the core is collapsed and native-like and the rest of the molecule is more disordered.

Partially folded $[14-38]_{\text{Abu}}$ may be unfolded by lowering the pH to 2.5, by heating past the cooperative thermal unfolding tran-

Table 1. Diffusion coefficients at infinite dilution, D° , and hydrodynamic radii, R_h , for various conformational states of BPTI^a

Conformation			D° ($\text{cm}^2/\text{s}) \times 10^7$	R_h (Å)	% increase ^b
BPTI	pH 4.5	Native	7.7 ± 0.1	15.2 ± 0.2	
$[14-38]_{\text{Abu}}$	pH 4.5	Partially folded, 1 S–S	6.45 ± 0.08	18.1 ± 0.2	19
$[14-38]_{\text{Abu}}$	pH 2.5	Unfolded, 1 S–S	5.63 ± 0.03	20.7 ± 0.1	36
$[R]_{\text{Abu}}$	pH 4.5	Unfolded, no S–S	5.65 ± 0.06	20.6 ± 0.3	36
$[R]_{\text{Abu}}$	pH 2.5	Unfolded, no S–S	5.13 ± 0.02	22.7 ± 0.1	49
$[R]_{\text{Abu}}$		Random coil ^c		23.6 ^c	55

^aValues of D° are estimated from PFG NMR experiments at 1 °C. R_h is calculated from the Stokes-Einstein relationship, Equation 2, as described in the text.

^bThe percent increase in R_h is relative to native.

^cThe hydrodynamic radius, R_h , for a BPTI random coil is obtained in the following way. The radius of gyration, R_g , for the random coil conformation, with no disulfide bonds, is estimated from the relationship $6R_g^2 = 130n$, where n is the number of residues; a further correction is made for a relatively short polypeptide chain as described in Tanford (1968). For a fully "random coil" polypeptide chain of 58 amino acids, the calculated R_g is 31 Å, and the corresponding R_h is 20.8 Å, computed from the relationship $R_h = 0.655(R_g)^{1/2}$ (Bloomfield et al., 1974). A monolayer of hydration is accounted for by adding 2.8 Å, giving a final calculated value of R_h for "random coil" BPTI of 23.6 Å.

sition, or by replacing Y21 or Y23 with alanine. Unfolded [14–38]_{Abu} (Fig. 1, upper right) obtained by any of these methods loses all CD- and NMR-detected secondary and tertiary structure, and gives NMR spectra with chemical shift dispersion similar to a disordered peptide. The large decrease in D° for unfolded [14–38]_{Abu} (pH 2.5), compared to that of partially folded [14–38]_{Abu} (pH 4.5), is concordant with their structural differences. The most complete NMR characterization of unfolded [14–38]_{Abu} is for Y21A and Y23A mutants of [14–38]_{Abu} at pH 4.5 (Barbar et al., 1996). These mutants lack the striking, non-native, amide–amide NOEs observed in residues 29–35 in reduced BPTI; apparently backbone conformations in unfolded [14–38]_{Abu} are, on average, more extended than in reduced BPTI. In addition, a large number of non-native, aromatic–aliphatic NOEs observed for reduced BPTI are not observed for unfolded [14–38]_{Abu}. Disordered, unfolded [14–38]_{Abu} retains a crosslink, and as expected, this is reflected in its D° and the related hydrodynamic radius, which is less than calculated for a random coil of the same length (Table 1).

Reduced BPTI is collapsed at pH 4.5

At pH above the carboxyl titrations, reduced BPTI is collapsed. Conformations sampled by [R]_{Abu} at pH 4.5 have, on average, the same effective hydrodynamic radius as acid-unfolded BPTI constrained with a 14–38 crosslink. [R]_{Abu} at pH 4.5, and acid-unfolded [14–38]_{Abu} at pH 2.5, have the same D° , corresponding to an R_h that is increased 36% relative to native, and is 87% of an equivalent random coil (Table 1). Given the inherent uncertainties in calculated random coil values of R_h , one might question the significance of an experimental R_h that differs from the calculated random coil R_h by only 15%. However, in addition to D° (and the related R_h) of [R]_{Abu} at pH 4.5, other experimental evidence indicates that [R]_{Abu} is indeed collapsed at pH 4.5: At pH 2.5, [R]_{Abu} expands further, and the measured D° corresponds to an R_h equivalent to the calculated random coil R_h (Table 1). An expansion of [R]_{Abu} at pH 2.5, compared to pH 4.5, is consistent with the widespread loss of NOEs in NMR spectra of [R]_{Abu} at pH 2.5 compared to spectra at pH 4.5. The value of D° for [R]_{Abu} at pH 4.5 is the same as for unfolded BPTI crosslinked between residues 14 and 38; unfolded [14–38]_{Abu} is undoubtedly more collapsed than a fully extended random coil.

Our hydrodynamic data are consistent with conclusions from fluorescence energy transfer experiments of Haas and associates, who measured the intramolecular segmental end-to-end distance distributions of unfolded BPTI species by non-radiative excitation energy transfer measurements (Amir & Haas, 1987; Gottfried & Haas, 1992; Ittah & Haas, 1995). Their results imply that even in GdmCl, reduced BPTI is not a condensed statistical coil but an ensemble of non-randomly folded conformations. Results in Table 1 are also consistent with previous native gel electrophoresis experiments (Ferrer et al., 1995), indicating that both [14–38]_{Abu} and [R]_{Abu} are expanded relative to the native state at pH 4.5, 7 °C. However, the hydrodynamic difference observed here between partially folded [14–38]_{Abu} and [R]_{Abu} at pH > 4.5 was not evident in their gel mobilities.

Collapsed conformations in [R]_{Abu} may be stabilized by non-native hydrophobic interactions

In considering the likely driving force for the collapse of reduced BPTI at pH 4.5, it is useful to consider the hydrodynamic results in light of the NMR structural data. The observed diffusion coef-

ficient of a denatured state is the ensemble average; deviations from the random coil D° may indicate the presence of either a small population of very compact molecules interconverting rapidly with fully disordered conformers or a large population of moderately compact conformers. NMR parameters such as chemical shifts and coupling constants are also average properties of all conformers present, while NOEs may report small subpopulations.

[R]_{Abu} at pH 4.5 samples numerous conformations in fast exchange (interconverting on a time scale of $\leq$ msec). Residues throughout the entire protein report $\text{CaH}_i\text{--NH}_{i+1}$ NOEs, indicating the presence of extended conformations. In addition, 18 peptide units also report $\text{NH}_i\text{--NH}_{i+1}$ NOEs indicating the presence of turn-like conformers in equilibrium with extended conformers. The NMR-detected sampling of turn-like conformations at numerous sites along the backbone implies that the protein is to some extent collapsed, the conclusion also drawn from this hydrodynamic study. The only *contiguous* stretch of residues with strong, *non-native* $\text{NH}_i\text{--NH}_{i+1}$ NOEs are 29–35, which in native BPTI fold into one of the anti-parallel β -strands. Residues that, in native BPTI, fold into both strands of β -sheet show, in [R]_{Abu} at pH 4.5, numerous, non-native, aromatic–aliphatic, side-chain–side-chain NOEs between residues that are one to three residues apart. These non-native hydrophobic clusters may stabilize more collapsed conformers of reduced BPTI at pH 4.5. This possibility is supported by the observations that most aromatic–aliphatic side-chain NOEs observed at pH 4.5 are absent at pH 2.5, and that R_h of reduced BPTI at pH 2.5 is equivalent to the value calculated for a random coil (Table 1). The larger effective hydrodynamic radius of [R]_{Abu} at pH 2.5, compared to pH 4.5, is probably due to increased intramolecular charge repulsion of the unfolded polypeptide resulting from neutralization of the four carboxyl groups in a small protein with six arginine and four lysine residues.

Aggregation

There is no indication from NMR studies that [14–38]_{Abu} and [R]_{Abu} aggregate at concentrations ≤ 0.6 mM; below this level, no changes in line widths, chemical shift, or NOEs are observed (Barbar et al., 1995; Pan et al., 1995). Slopes of the curves in Figures 3 and 4 vary with the protein and with pH; the concentration dependence is most pronounced with [R]_{Abu} at pH 4.5. Since the value of D° is obtained by extrapolation of the data in Figures 3 and 4 to infinite dilution, it is important to assess the possible contribution from protein aggregation to D° determined in this way. To do this, consider a simple monomer–dimer equilibrium ($M + M \leftrightarrow D$), with dissociation constant K_d . If x is the mole fraction of total monomer converted to dimer, then

$$K_d = x/(2C(1 - x)^2) \quad (3)$$

where C is the protein concentration. The observed D_{app} is the weighted average of diffusion coefficients of monomer and dimer:

$$D_{app} = D_1(1 - x) + D_2x/2 \quad (4)$$

where D_1 is the diffusion coefficient of monomer, and D_2 is the diffusion coefficient of dimer. For the case of two monomer spheres at a separation of $2r$ (r is the radius of one monomer), $D_2 \sim 0.719D_1$ (Wills & Georgalis, 1981). Fits of D_{app} versus protein concentration to Equations 3 and 4 give approximate K_d values of 20 mM,

6.7 mM, and 2.9 mM, respectively, for native, [14–38]_{Abu}, and [R]_{Abu} at pH 4.5 and 1 °C. These values of K_d mean that, at a protein concentration of 0.5 mM, the total monomer converted to dimer is 12% for partially folded [14–38]_{Abu} and 22% for [R]_{Abu}. For higher order association, an even smaller percentage of total monomer in aggregates would be expected. We conclude that there is not a significant contribution to D° from aggregation at protein concentrations below 0.5 mM.

At both pH 2.5 and pH 4.5, [R]_{Abu} shows a greater concentration dependence of diffusion coefficient than does [14–38]_{Abu}, and both proteins have lower concentration dependencies at pH 2.5. The net intermolecular interaction is smaller at pH 2.5, presumably from the neutralization of the carboxyl groups, and the associated rise in net positive charge, as discussed above for intramolecular interactions. Further, the presence of numerous aliphatic–aromatic NOEs in [R]_{Abu} at pH 4.5, but not pH 2.5, is consistent with the possibility that decreased solubility at the higher pH is due in part to intermolecular interactions of hydrophobic clusters.

Comparison to denatured states of other proteins

The radius of gyration, R_g , has been measured for native and denatured states of several other proteins; these are listed in Table 2 along with literature citations. Most studies focus on molten globules, thought by some to be specific states of proteins corresponding to folding intermediates with extensive secondary structure and little tertiary structure (Ptitsyn, 1992) but this concept has recently been called into question (Privalov, 1996). Under “molten globule” conditions, R_g values of cytochrome *c* and apomyoglobin are increased 26–32% relative to native. The SNase fragment, truncated 13 residues from the C-terminus, also lacks disulfide bonds, and its R_g value is increased 31%. For α -lactalbumin under molten globule conditions, the increase is about 10%, much less than for other molten globules, and probably due to the presence of four disulfide bonds (Kataoka et al., 1995, 1997). Molten globules listed in Table 2, as well as truncated SNase, have significant CD-detected secondary structure and some NMR-detected native-like secondary structure.

Table 2. Literature values of the radius of gyration, R_g , of protein native and non-native states

Protein	No. of residues	R_g (Å) ^a						Random coil ^b
		Native	Molten globule	Truncated	Heat-denatured	Acid-denatured	Urea/GdmCl denatured	
Cytochrome <i>c</i> ^c	104	13.5	17.0, 17.4			24.2	32.4	40
SSI ^d	113	20.1			25.8 (2 S–S)		35 (2 S–S)	43
α -Lactalbumin ^e	123	15.7	17.2 (4 S–S)				24.30 (4 S–S)	45
RNase A ^f	124	15.0			19.3 (4 S–S)			45
Reduced					28			
Lysozyme ^g	129	15.3					21.8 (4 S–S)	47
SNase ^h	149	16.2		21.2 ⁱ			35	51
Myoglobin ^j								
holo	153	17.5						52
apo		19.7	23.1			30.2	34.2–35.8	
BLG ^k	162	18					35 (2 S–S)	54
BCAB ^k	259	19					59	71
Yeast PGK ^k	415	21					66	95
% of increase ^l								
With S–S			10%		28–29%		42–94%	
No S–S			26–32%	31%	87%	73–79%	95–214%	114–352%
% of r. c. ^m								
With S–S			38%		43–60%		46–81%	
No S–S			44–46%	42% ⁿ	62%	58–60%	66–83%	

^aProteins with disulfide bonds are indicated by (S–S), with the number of disulfide bonds included in the parentheses. All other proteins have no disulfide bonds.

^bThe radius of gyration for the random coil conformation, with no disulfide bonds, is calculated from the relationship $6R_g^2 = 130n$, where n is the number of residues, and includes the correction factor for chain length, as described in Tanford (1968).

^cData are for holocytochrome *c* (Kataoka et al., 1993).

^dStreptomyces subtilisin inhibitor (SSI) data are reported in Konno et al. (1995); for cold denatured SSI, $R_g = 20.7$ Å.

^eData for α -lactalbumin are in Kataoka et al. (1996, 1997), Dolgikh et al. (1985), and Semisotnov et al. (1996).

^fRibonuclease A data are in Sosnick et al. (1992).

^gLysozyme data are in Chen et al. (1996).

^hStaphylococcal nuclease (SNase) data are in Flanagan et al. (1992).

ⁱ R_g value for the SNase fragment missing 13 amino acid residues from the C-terminus.

^jMyoglobin data are in Kataoka et al. (1995).

^kBovine β -lactoglobulin (BLG), bovine carbonic anhydrase B (BCAB), and yeast phosphoglycerate kinase (PGK) are reported in Semisotnov et al. (1996).

^lThe percent increase in R_g is relative to native. The range of percent increase is listed separately for proteins with disulfide bonds and for proteins without disulfides.

^mThe percent of random coil is the R_g of non-native divided by R_g of random coil. The range of percent random coil is listed separately for proteins with disulfide bonds and for proteins without disulfides.

ⁿThe % of random coil is based on the R_g calculated for the SNase fragment (47 Å).

For unfolded states obtained by acid, temperature, or chemical denaturation, R_g values are much larger than for molten globules (Table 2, columns 6, 7, and 8) but still significantly less than the random coil radius of gyration computed for an equivalent peptide of the same number of residues and no disulfide bonds (Table 2, column 9). An effect of disulfide bonds on unfolded proteins is illustrated by the data for heat-denatured RNase A (Sosnik & Trewella, 1992). The R_g of heat-denatured RNase with four disulfide bonds is 43% of the random coil value, as compared to 62% for heat-denatured RNase without disulfide bonds (Table 2).

Comparison of the BPTI results in Table 1 to hydrodynamic data for other proteins in Table 2 is complicated somewhat by the fact that the parameter measured for BPTI is hydrodynamic radius while for the others it is radius of gyration. For native proteins, modeled as solid spheres, $R_h > R_g$; for random coils, $R_h < R_g$ (Bloomfield et al., 1974; van Holde, 1985). Nevertheless, a qualitative comparison of the percentage change relative to native, or of the percentage of random coil, is reasonable.

Partially folded BPTI ([14–38]_{Abu} at pH 4.5) is expanded by 19% relative to native, based on R_h values. This is in the same range as other proteins under molten globule conditions, and for truncated SNase, based on R_g values. The NMR structure of partially folded [14–38]_{Abu} is entirely consistent with this.

R_g values for denatured proteins having no disulfide bonds and no covalent prosthetic groups (reduced RNase A, apomyoglobin, SNase, BCAB, PKG) are 62–83% of the calculated random coil value (Table 2). Comparable to this, the R_h of [R]_{Abu} is 87% of the calculated random coil value (Table 1). Residual secondary structure does not explain why either [R]_{Abu} or the denatured proteins in Table 2 are less expanded than a random coil. A small amount of α -helical and β -sheet structure is observed for heat-denatured reduced RNase (Sosnik & Trewella, 1992) but this is lost upon addition of high concentrations of chemical denaturants. Acid and chemically unfolded apomyoglobin CD spectra indicate, respectively, 5% and zero helical structure (Nishii et al., 1994). Urea-unfolded, wild-type SNase has a small residual CD signal at 220 nm (Sugawara et al., 1991; Wang & Shortle, 1995). [R]_{Abu} at pH 4.5 has a CD spectrum similar to chemically unfolded apomyoglobin and SNase, and NMR data indicate no secondary structure. Although [R]_{Abu} at pH 4.5 is less expanded than it is at pH 2.5, and less expanded than an equivalent random coil, it is still more expanded relative to a random coil than other proteins denatured by chemicals and/or temperature. This suggests another possible facet of the functional role of disulfide bonds: To be an effective protease inhibitor, BPTI needs to be very stable, and three disulfide bonds are required to sufficiently decrease the entropic stability of the unfolded state.

At pH 2.5, the R_h value of [R]_{Abu} is very close to the calculated random coil value, and among the protein species listed in Tables 1 and 2, is the only example of a protein approaching random coil hydrodynamic behavior.

Conclusions

Reduced BPTI ([R]_{Abu}) at pH 4.5 is more compact than at pH 2.5, where it behaves as a statistical random coil. Since BPTI contains no groups that titrate between pH 4.5 and 7.5, [R]_{Abu} at pH 4.5 is presumably similar to the species present at the onset of in vitro folding at neutral pH. At pH 4.5, the hydrodynamic radius derived from the measured diffusion coefficient is about 87% that of an equivalent random coil. The similarity of D° values for [R]_{Abu} at

pH 4.5, and for unfolded [14–38]_{Abu} at pH 2.5 implies that the extent of collapse of reduced BPTI at pH 4.5 is about the same, on average, as the equivalent random coil with a crosslink between residues 14 and 38.

Structural data on reduced BPTI at pH 4.5 and 2.5, from NMR and from ANS binding, suggest that the collapse reported by its hydrodynamic behavior at pH 4.5 arises from non-native hydrophobic interactions between side chains one to three amino acids apart in residues that, in native BPTI, fold into the core of the protein.

Partially folded BPTI with a single disulfide crosslink ([14–38]_{Abu} at pH 4.5) is fairly compact; relative to native BPTI it is expanded by about 19%, similar to other reports for non-disulfide containing proteins under “molten globule” conditions.

Materials and methods

Sample preparation

Naturally occurring BPTI was purchased from Novo Industries, Copenhagen (aprotinin Novo, ultrapure grade). [R]_{Abu} and [14–38]_{Abu} were chemically synthesized by automated Fmoc solid-phase synthesis and purified as described in Ferrer et al. (1995) and Barany et al. (1996). Protein purities were confirmed by analytical HPLC and capillary zone electrophoresis. Ion electrospray mass spectrometry of purified proteins gave mass values in good agreement with calculated values. NMR protein samples were dialyzed against ¹H₂O at pH 4.5 adjusted with HCl, lyophilized, and then dissolved in 99.96% ²H₂O with 10 mM sodium phosphate. Protein concentrations were determined by quantitative amino acid analysis.

Pulsed-field gradient NMR self-diffusion experiments

As described in Results, the apparent translational diffusion coefficient, D , may be obtained from the signal intensity, A , recorded in a PFG NMR experiment, normalized to the signal in the absence of the magnetic field-gradient pulses. Plots of A as a function of $\gamma^2 G^2 \delta^2 (\Delta - \delta/3)$ are fit to Equation 1 by nonlinear least-square regression analysis to give D and its standard deviation. In this study, $\Delta = 34.2$ ms, $\delta = 4$ ms, $G = 1$ –65 G/cm. A PFG longitudinal eddy-current delay (LED) pulse sequence (Gibbs & Johnson, 1991) was used to measure the diffusion coefficient, D , of proteins. NMR self-diffusion measurements were carried out on a Bruker AMX-600 equipped with a GRASP gradient unit, using a 5 mm triple-resonance probe with an actively shielded Z-gradient coil. The maximum magnitude of the gradient was calibrated to be 100 G/cm (Mayo et al., 1996). For native BPTI, amplitudes of resolved resonances of individual ¹H were measured, while for [R]_{Abu} or [14–38]_{Abu}, amplitudes of composite resonances were measured. All spectra were derived from 8k data points and a spectral width of 7,246 Hz. The number of scans ranges from 512 to 2k.

Acknowledgments

We thank Dr. Victor Bloomfield for help and discussions. This work is supported by NIH grants GM26242 (C.W.) and 51628 (G.B. & C.W.).

References

- Amir D, Haas E. 1987. Estimation of intramolecular distribution in bovine pancreatic trypsin inhibitor by site-specific labeling and nonradiative excitation energy-transfer measurements. *Biochemistry* 26:2162–2175.

- Altieri AS, Kinton DP, Byrd RA. 1995. Association of biomolecular systems via pulses field gradient NMR self-diffusion measurements. *J Am Chem Soc* 117:7566–7567.
- Barany G, Gross C, Ferrer M, Barbar E, Pan H, Woodward C. 1996. Optimized methods for chemical synthesis of bovine pancreatic trypsin inhibitor analogues. In: Marshak D, ed. *Techniques in protein chemistry VII*. San Diego: Academic Press. pp 503–514.
- Barbar E, Barany G, Woodward C. 1995. Dynamic structure of a highly ordered β -sheet molten globule: Multiple conformations with a stable core. *Biochemistry* 34:11423–11434.
- Barbar E, Barany G, Woodward C. 1996. Unfolded BPTI variants with a single disulfide bond have diminished non-native structure distant from the cross-link. *Folding & Design* 1:65–76.
- Barbar E, Barany G, Woodward C. 1997. Local fluctuations and global unfolding of partially folded BPTI detected by NMR. *Biophysical Chemistry* 64:45–57.
- Bloomfield V, Crothers D, Tinoco I Jr. 1974. Transport phenomena. In: *Physical chemistry of nucleic acids*. New York: Harper & Row, Publishers. pp 215–254.
- Chen L, Hodgson KO, Doniach S. 1996. A lysozyme folding intermediate revealed by solution X-ray scattering. *J Mol Biol* 261:658–671.
- Dill KA. 1985. Theory for the folding and stability of globular proteins. *Biochemistry* 24:1501–1509.
- Dill KA, Bromberg S, Yue K, Fiebig K, Yee D, Thomas P, Chan H-S. 1995. Principles of protein folding—A perspective from simple exact models. *Protein Sci* 4:561–602.
- Dill KA, Chan H-S. 1997. From Levinthal to pathways to funnels. *Nature Struct Biol* 4:10–19.
- Dill KA, Fiebig KM, Chan HS. 1993. Cooperativity in protein folding kinetics. *Proc Natl Acad Sci USA* 90:1942–1946.
- Dingley AJ, Mackay JP, Chapman BE, Morris M, Kuchel PW, Hambly BD, King GF. 1995. Measuring protein self-association to myosin light chain 2. *J Biomol NMR* 6:321–328.
- Dolgikh DA, Abaturon LV, Bolotina IA, Brazhnikov EV, Bychkova VE, Gilmanishin RI, Lebedev YO, Semisotnov GV, Tikopulo EI, Ptitsyn OB. 1985. Compact state of a protein molecule with pronounced small-scale mobility: bovine α -lactalbumin. *Eur Biophys J* 13:109–121.
- Eliczer D, Jennings PA, Wright PE, Doniach S, Hodgson K, Tsuruta H. 1995. The radius of gyration of an apomyoglobin folding intermediate. *Science* 270:487–488.
- Ferrer M, Barany G, Woodward C. 1995. Partially folded, molten globule and molten coil states of bovine pancreatic trypsin inhibition. *Nature Struct Biol* 2:211–217.
- Flanagan JM, Kataoka M, Shortle D, Engelman DM. 1992. Truncated staphylococcal nuclease is compact but disordered. *Proc Natl Acad Sci* 89:748–752.
- Gallagher WH, Woodward C. 1989. The concentration dependence of the diffusion coefficient for bovine pancreatic trypsin inhibitor: A dynamic light scattering study of a small protein. *Biopolymers* 28:2001–2024.
- Gibbs SJ, Johnson CS. 1991. A PFG NMR experiment for accurate diffusion and flow studies in the presence of eddy currents. *J Magn Reson* 93:395–402.
- Gottfried DD, Haas E. 1992. Nonlocal interactions stabilize compact folding intermediates in reduced unfolded bovine pancreatic trypsin inhibitor. *Biochemistry* 31:12353–12362.
- Hamada D, Hoshino M, Kataoka M, Fink AL, Goto Y. 1993. Intermediate conformational states of apocytochrome *c*. *Biochemistry* 32:10351–10358.
- Ilyina E, Roongta V, Pan H, Woodward C, Mayo K. 1997. A pulsed field gradient NMR study of BPTI self-association. *Biochemistry* 36:3383–3388.
- Ittah V, Haas E. 1995. Nonlocal interactions stabilize long range loops in the initial folding intermediates of reduced bovine pancreatic trypsin inhibitor. *Biochemistry* 34:4493–4506.
- Izumi Y, Miyake Y, Kuwajima K, Sugai S, Inoue K, Iizumi M, Katano S. 1983. Folding-unfolding of α -lactalbumin. *Physica* 120b:444–448.
- Kataoka M, Goto Y. 1996. X-ray solution scattering studies of protein folding. *Folding & Design* 1:R107–R114.
- Kataoka M, Hagihara Y, Mihara K, Goto Y. 1993. Molten globule of cytochrome *c* studied by small angle X-ray scattering. *J Mol Biol* 229:591–596.
- Kataoka M, Kuwajima K, Tokunaga F, Goto Y. 1997. Structural characterization of the molten globule of α -lactalbumin by solution X-ray scattering. *Protein Sci* 6:422–430.
- Kataoka M, Nishii I, Fujisawa T, Ueki T, Tokunaga F, Goto Y. 1995. Structural characterization of the molten globule and native states of apomyoglobin by solution X-ray scattering. *J Mol Biol* 249:215–228.
- Konno T, Kataoka M, Kamatari Y, Kanaori K, Nosaka A, Akasaka K. 1995. Solution X-ray scattering analysis of cold-heat-, and urea-denatured states in a protein, streptomyces subtilisin inhibitor. *J Mol Biol* 251:95–103.
- Lennon AJ, Scott NR, Chapman BE, Kuchel PW. 1994. Hemoglobin affinity for 2,3-bisphosphoglycerate in solution and intact erythrocytes: Studies using pulsed-field gradient nuclear magnetic resonance and Monte Carlo simulations. *Biophys J* 67:2096–2109.
- Mayo KH, Ilyina E, Park H. 1996. A recipe for designing water-soluble, β -sheet-forming peptides. *Protein Sci* 5:1301–1315.
- Nishii I, Kataoka M, Tokunaga F, Goto Y. 1994. Cold denaturation of the molten globule states of apomyoglobin and a profile for protein folding. *Biochemistry* 33:4903–4909.
- Pan H, Barbar E, Barany G, Woodward C. 1995. Extensive nonrandom structure in reduced and unfolded bovine pancreatic trypsin inhibitor. *Biochemistry* 34:13974–13981.
- Privalov PL. 1996. Intermediate states in protein folding. *J Mol Biol* 258:707–725.
- Ptitsyn OB. 1992. The molten globule state. In: Creighton TE, ed. *Protein folding*. New York: W.H. Freeman and Company. pp 243–300.
- Sali A, Shakhnovich E, Karplus M. 1995. How does a protein fold? *Nature* 369:248–251.
- Semisotnov GV, Kihara H, Kotova NV, Kimura K, Amemiya Y, Wakabayashi K, Serdyuk IN, Timchenko AA, Chiba K, Nikaido K, Ikura T, Kuwajima K. 1996. Protein globularization during folding: A study by synchrotron small angle X-ray scattering. *J Mol Biol* 262:559–574.
- Sosnik TR, Trehwella J. 1992. Denatured states of ribonuclease A have compact dimensions and residual secondary structure. *Biochemistry* 31:8329–8335.
- Stejskal EO, Tanner JE. 1965. Spin diffusion measurements: Spin echoes in the presence of a time dependent field gradient. *J Chem Phys* 42:288–292.
- Sugawara T, Kuwajima K, Sugai S. 1991. Folding of staphylococcal nuclease A studied by equilibrium and kinetic circular dichroism spectra. *Biochemistry* 30:2698–2706.
- Tanford C. 1968. Protein denaturation. *Adv Protein Chem* 23:121–282.
- van Holde KE. 1985. Scattering. In: *Physical biochemistry*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc. pp 209–234.
- Wang Y, Shortle D. 1995. The equilibrium folding pathway of staphylococcal nuclease: Identification of the most stable chain-chain interactions by NMR and CD spectroscopy. *Biochemistry* 34:15895–15905.
- Wills PR, Georgialis Y. 1981. Concentration dependence of the diffusion coefficient of a dimerizing protein: Bovine pancreatic trypsin inhibitor. *J Phys Chem* 85:3978–3984.