Lipsett, M. N. (1963) *Biophys. Res. Commun. 11*, 224.
Marck, C., & Thiele, D. (1978) *Nucleic Acids Res. 5*, 1017-1028.
Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499-560.
Michelson, A. M., Massoulié, J., & Guschlbauer, W. (1967) *Prog. Nucleic Acid Res. Mol. Biol. 6*, 83.
Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A., & Frank-Kamenetskii, M. D. (1987) *Nature 330*, 495-497.
Morgan, A. R., & Wells, R. D. (1968) *J. Mol. Biol. 37*, 63-80.
Moser, H., & Dervan, P. B. (1987) *Science 238*, 645-650.
Praseuth, D., Perroualt, L., Le Doan, T., Chassignol, M., Thuong, N., & Hélène, C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1349-1353.
Puglisi, J D., Wyatt, J. R., & Tinoco, I., Jr. (1988) *Nature 331*, 283-286.
Rajagopal, P., & Feigon, J. (1989) *Nature 239*, 637-640.
Rubin, C. M., & Schmid, C. W. (1980) *Nucleic Acids Res. 8*, 4613-4619.
Schimmel, P. (1989) *Cell 58*, 9-12.
Sklerer, V., & Feigon, J. (1990) *Nature 345*, 836-838.
Uhlenbeck, O. C. (1987) *Nature 328*, 596-600.
Voloshin, O. N., Mirkin, S. M., Lyamichev, V. I., Belotserkovskii, B. P., & Frank-Kamenetskii, M. D. (1988) *Nature 333*, 475-476.
Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M., & Wohlrab, F. (1988) *FASEB J. 2*, 2939-2949.

Reduced Tendency To Form a β Turn in Peptides from the P22 Tailspike Protein Correlates with a Temperature-Sensitive Folding Defect†

Adam N. Stroup‡ and Lila M. Gierasch*,‡,§
Departments of Biochemistry and Pharmacology, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Boulevard, Dallas, Texas 75235-9041

Received July 11, 1990; Revised Manuscript Received August 23, 1990

ABSTRACT: A family of mutants of the P22 bacteriophage tailspike protein has been characterized as temperature sensitive for folding (tsf) by King and co-workers [King, J. (1986) *Bio/Technology 4*, 297-303]. There is substantial evidence that the tsf mutations alter the folding pathway but not the stability of the final folded protein. Several point mutations are known to cause the tsf phenotype; most of these occur in regions of the tailspike sequence likely to take up reverse turns. Hence, it has been hypothesized that the correct folding of the P22 tailspike protein requires formation of turns and that the mutations causing tsf phenotypes interfere at this critical stage. We have tested this hypothesis by study of isolated peptides corresponding to a region of the P22 tailspike harboring a tsf mutation. Comparison of the tendencies of wild-type and tsf sequences to adopt turn conformations was achieved by the synthesis of peptides with flanking cysteine residues and the use of a thiol–disulfide exchange assay. We find that the wild-type sequence, either as a decapeptide (Ac-CVKFPGIETC-$CONH_2$) or as a dodecapeptide (Ac-CYVKFPGIETLC-$CONH_2$), has a 3–5-fold greater tendency for its termini to approach closely enough to form the intramolecular disulfide than do the peptide sequences corresponding to the tsf mutant sequences, which have a Gly → Arg substitution (Ac-CVKFPRIETC-$CONH_2$ or Ac-CYVKFPRIETLC-$CONH_2$). A peptide with a D-Arg substituted for the Gly has a slightly higher turn propensity than does the wild type. Together with data from nuclear magnetic resonance analysis of the oxidized peptides, this suggests that a type II β turn is favored by the wild-type sequence. Our results on isolated peptides from the P22 tailspike protein support the model for its folding that includes reverse turn formation as a critical step.

Much research effort is currently directed toward a better understanding of the pathway of protein folding. While well established that amino acid sequence determines the ultimate three-dimensional structure of a native protein, the mechanism by which the linear polymer of amino acids achieves its complex folding pattern remains a puzzle. Most would agree that the ability to predict protein structures from amino acid sequences will be realized only concomitant with an understanding of the protein folding process. Recent evidence that protein folding in vivo often involves factors apart from the nascent polypeptide chain adds another layer of complexity to elucidating the pathway of folding in a cell. It is likely, nonetheless, that the additional factors that facilitate folding in vivo do not alter the basic process but instead ensure that competing processes (e.g., aggregation) are minimized or prevented.

Characterization of folding intermediates has emerged as a key advance in our understanding of the folding pathway. Two recent studies illustrate this approach. Oas and Kim (1988) have modeled an intermediate in the folding of bovine pancreatic trypsin inhibitor (BPTI)[1] by synthesis of two peptide fragments that form a domain likely to be present early in the

† This work was supported by a grant from the National Institutes of Health (GM 27616) and by the Robert A. Welch Foundation.
‡ Department of Biochemistry.
§ Department of Pharmacology.

[1] Abbreviations: Ac, acetyl; BPTI, bovine pancreatic trypsin inhibitor; DMF, dimethylformamide; DIEA, diisopropylethylamine; EDTA, ethylenediaminetetraacetic acid; GSH, reduced glutathione; GSSG, oxidized glutathione; HObt, 1-hydroxybenzotriazole; IAcm, iodoacetamide; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; RP-HPLC, reverse-phase high-performance liquid chromatography; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid; Tris, tris(hydroxymethyl)aminomethane; tsf, temperature-sensitive folding; 2-D, two dimensional; WT, wild type.

0006-2960/90/0429-9765$02.50/0

folding process. Roder et al. (1988) have used pulse labeling of NH's by deuterium exchange in cytochrome *c* during its folding and NMR characterization of the pattern of labeling to deduce which parts of the protein are folded first.

An alternative to these physical approaches is offered by genetic strategies, which have the advantage that they can reveal aspects of folding in vivo. King and co-workers used genetic methods to find a family of mutants of the P22 bacteriophage that are temperature-sensitive for production of functional tailspike protein (Smith & King, 1981; Goldenberg et al., 1983; Yu & King, 1988; King, 1990). At the permissive temperature, the 666 amino acid long tailspike monomers form native, trimeric tailspikes which are indistinguishable from wild type by a variety of criteria. At the nonpermissive temperature, irreversible aggregates of the tailspike protein are formed. It was concluded that these mutants were temperature-sensitive for folding (tsf) of the tailspike monomer and that a defect in their folding left them vulnerable to the competing process in the cell, viz., aggregation. Several point mutations in the protein could cause this phenotype. Most of the mutations causing the tsf phenotype occurred in regions of the wild-type sequence that have a high probability to take up a reverse turn, and the mutations led to a lower predicted tendency for the region to form a turn (Villafane & King, 1988). These results and independent evidence from Raman spectroscopy (Sargent et al., 1988) that the P22 tailspike is largely β suggested that a cross-β motif may exist in the folded protein. The genetic observations thus lead to a model for the in vivo folding pathway of the P22 tailspike protein in which an early event is formation of the cross-β motif, followed by coalescence of monomers into protrimers (still thermolabile) and eventually native trimers. The aggregation process appears to compete with the productive folding pathway when the monomers are involved in initial steps of folding (Haase-Pettingill & King, 1988).

On the basis of this model, it can be inferred that the local sequence harboring the point mutations in the tsf strains must influence the process of folding, most likely when the cross-β structure is forming. Since the point mutations appear in turn regions and lead to reduced predicted turn formation, the roles of these sites in the pathway may be to facilitate coalescence of the β-sheet structure. Since there is a phenotype associated with folding which arises from a very modest change in the sequence of the tailspike, we decided to test the effect of the tsf point mutation in its local sequence, removed from the rest of the tailspike protein. Hence, we have synthesized deca- and dodecapeptides which comprise the sequences found at the tsH304 site in the P22 tailspike. The wild-type decapeptide sequence, YVKFPGIETL, contains Pro-Gly, highly likely to adopt a reverse turn, and in the tsf version, Gly is replaced by Arg (Yu & King, 1984). In order to address the question of tendency to adopt a turn conformation *experimentally*, we needed a method that sampled the equilibrium conformational distribution of these flexible peptides. We chose to incorporate flanking cysteine residues and to use the intramolecular disulfide formation tendency as a measure of turn likelihood. The assay is based on one originally described by Creighton (1978) in studies of the cooperativity of folding and disulfide formation in BPTI. The same or similar approaches have since been used by several laboratories to ask about turn formation (Snyder, 1987; Milburn et al., 1987; Lin & Kim, 1989) or to explore orientation of monomers in a leucine-zipper model (O'Shea et al., 1989).

We find that the peptides corresponding to the wild-type (WT) and tsf mutant sequences in the tsH304 region differ significantly in their tendencies to form folded turn conformations. The tsf mutation causes a 3–5-fold decrease in the turn likelihood. A control peptide with the wild-type Gly substituted by D-Arg tends to fold slightly more readily than the wild type. Together with NMR data on the oxidized dodecapeptides, this result argues for presence of a type II β turn in the folded forms of these peptides. Extrapolating our results on isolated peptides to the P22 tailspike folding process suggests that a relatively small shift in an equilibrium may perturb a folding step sufficiently to cause off-pathway processes such as aggregation to dominate.

Experimental Procedures

Materials. Protected amino acids were obtained from Peptides International (Louisville, KY) and *p*-methylbenzhydrylamine resin–1% divinylbenzene cross-linked polystyrene (0.96 mequiv of nitrogen/g) was from U.S. Biochemical (Cleveland, OH). Glutathione was purchased from Sigma (St. Louis, MO). Urea was Schwarz/Mann ultrapure grade. Other reagents, including 5,5′-dithiobis(2-nitrobenzoic acid) and iodoacetamide, were of the highest grade available (Aldrich). Nitrogen was prepurified grade, minimum of 99.99% purity.

Peptide Synthesis. Peptides were synthesized by standard solid-phase strategy (Stewart & Young, 1984) with *p*-methylbenzhydrylamine resin and *tert*-butyloxycarbonyl- (Boc) protected amino acids and the following side chain protected residues: Cys(4-methylbenzyl), Tyr(*O*-benzyl), Lys(N^{ϵ}-2-chlorobenzyloxycarbonyl), Arg(N^{g}-tosyl), Glu(γ-benzyl ester), and Thr(*O*-benzyl). A Peptides International PI-2000 peptide synthesizer was used for decapeptides with semiautomated methodology. Decapeptides were synthesized on an Applied Biosystems 430A peptide synthesizer with all reagents obtained from Applied Biosystems (Foster City, CA).

Peptides were cleaved from the resin with 90:10 anhydrous HF/anisole at 0 °C for 45–60 min. Resin thus treated was extensively washed with anhydrous ether followed by water and acetonitrile/water mixtures. The aqueous washes were combined and lyophilized, giving peptides with free thiols. Purification of reduced peptides was by HPLC using either Vydac C18 or phenyl (25 × 2.5 cm) columns and acetonitrile/water gradients with 0.1% TFA as an ion pairing agent.

Air oxidation of the peptides was carried out at pH 8 in dilute (<1 mg/mL water/methanol or water/acetonitrile, stirring vigorously for typically 18 h. Completion was judged by lack of thiol reactive with Ellman's reagent [5,5′-dithiobis(2-nitrobenzoic acid)] (Ellman, 1959). The pH was adjusted to approximately 2 with 25% TFA followed by preparative RP-HPLC. The G → R dodecapeptide was oxidized using GSSG in 0.10 M Tris, 0.20 M KCl, and 1.0 mM EDTA, pH 8.70, at approximately 25 μM peptide and purified by preparative RP-HPLC. An air oxidation reaction of the G → R dodecapeptide resulted in predominantly intermolecular (dimeric) oxidized product.

Carboxyamidomethylation was by addition of iodoacetamide (4 equiv) to a fresh aqueous solution of the thiol form of the peptide adjusted to pH 8 with 1 N NaOH. After 15 min the solutions were no longer reactive with Ellman's reagent and were quenched by adjusting the pH to approximately 2 with 25% TFA. The crude reaction mixture was immediately separated by HPLC, giving predominantly the desired derivative. The purity and identity of all peptides were established by RP-HPLC, amino acid analysis, and FAB mass spectrometry.

Equilibrium Thiol–Disulfide Exchange. A mixture of oxidized and reduced glutathione, referred to here as a redox

buffer, allows measurement of an effective concentration (C_{eff}) of cysteine thiols for the intramolecular reaction:

$$P(SH)_2 \leftrightarrow PSS \qquad K_{intra} = [PSS]/[P(SH)_2]$$

$$2GSH \leftrightarrow GSSG \qquad K_{inter} = [GSSG]/[GSH]^2$$

$$C_{eff} = K_{intra}/K_{inter} = [PSS][GSH]^2/[P(SH)_2][GSSG]$$

The C_{eff} measure typically describes an intramolecular ring closure but was adopted by Creighton (1983) as a measure of conformational stability of protein folding intermediates. Addition of chaotropes to the thiol–disulfide exchange medium has been used to disrupt conformation-promoting interactions, giving thiol–disulfide exchange based on steric constraints of the unfolded polypeptide chains (Goto & Hamaguchi, 1981). Urea has been demonstrated not to influence the C_{eff} for Ac-Cys-$(Gly)_6$-Cys-$CONH_2$ used as a control random coil peptide (Lin & Kim, 1989).

All solutions were prepared in a stock buffer of 0.10 M Tris, 0.20 M KCl, and 1.0 mM EDTA, pH 8.70 (Saxena & Wetlaufer, 1970; Creighton, 1974), and degassed by sonication while under water aspiration followed by saturation with nitrogen gas. Transfers were performed in a glovebag (I^2R, Cheltenham, PA) which had been extensively purged with nitrogen. Disulfide exchange was initiated by addition of the glutathione redox buffer, 3:1 or 5:1 GSH/GSSG, to peptide solutions contained in glass autosampler vials (Alltech, Deerfield, IL), giving a final peptide concentration of approximately 25 μM and either 7.5 mM GSH/2.5 mM GSSG or 8.33 mM GSH/1.67 mM GSSG. The vials were sealed under nitrogen with a Teflon-lined silicone septum, and exchange proceeded at ambient (23 °C) temperature. Equilibration times greater than 1 h were found to be sufficient for the system to come to equilibrium. Results were independent of the initial redox state of the peptide under study. The exchange reaction was quenched by the addition via syringe under nitrogen atmosphere of either 100 μL of aqueous 25% TFA (degassed by brief sonication while under water aspiration followed by saturation with nitrogen) or 0.5 mL of 0.27 M IAcm in stock buffer. The carboxyamidomethylation reaction was carried out for 5 min, followed by the addition of 100 μL of aqueous 25% TFA (pH approximately 2). Samples were analyzed immediately or stored at −25 °C until analysis. The characteristics of the two methods for quenching the thiol–disulfide exchange are that acid is rapid but not irreversible while the IAcm reaction is irreversible but slow (Snyder, 1987; Lin & Kim, 1989).

Analysis by RP-HPLC. Analyses were performed on a C18 column (15 × 0.4 cm) (Beckman) with an acetonitrile gradient starting at 10% acetonitrile (with 0.1% TFA in all solvents) and increasing at 1% acetonitrile/min at a flow rate of 1 mL/min, except for the WT dodecapeptide where reduced and oxidized forms coeluted. Analysis of the WT dodecapeptide was on a Vydac phenyl column (25 × 1 cm) starting at 20% acetonitrile with a 1%/min increasing acetonitrile gradient and a flow rate of 2 mL/min. Data were analyzed with System Gold software (Beckman, version 3.00 or 3.1). Peak areas at 220 nm were calibrated by amino acid analysis. Concentration of reduced glutathione was determined by colorimetric assay with Ellman's reagent (Ellman, 1959; Riddles et al., 1979). The total glutathione (oxidized + reduced) concentration was determined by amino acid analysis. Glutathione ratios were also determined by quenching with iodoacetamide, the Cys(SAcm) and cystine ratios determined by amino acid analysis being used. The values obtained were within experimental error of the gravimetrically determined glutathione concentrations which are reported.

Nuclear Magnetic Resonance (NMR). Conformationally diagnostic NMR parameters were determined for the wild-type and tsf mutant deca- and dodecapeptides on a Varian VXR 500S spectrometer operating at a proton frequency of 500 MHz. Phase-sensitive 2-D NOE spectra were obtained by the phase-cycling method of States et al. (1982). Suppression of the water signal in aqueous samples was carried out by presaturation. Assignments were obtained with a sequential resonance assignment strategy (Wagner et al., 1981). The peptides examined included the methylcarboxyamidylated derivative of the decapeptide and both the oxidized deca- and dodecapeptides. All samples were prepared at concentrations of 6–10 mg/mL in 90% H_2O/10% D_2O. The oxidized dodecapeptide had poor solubility in aqueous solution and hence was examined in greater detail in dimethyl sulfoxide/chloroform (65/35). Processing of 2-D NMR spectra was performed on a Sun4/260 computer using the program FTNMR, kindly provided by Dennis Hare (Hare Research, Woodinville, WA).

Results

Disulfide Exchange Assay. A representative HPLC chromatogram of the equilibrium thiol–disulfide exchange assay for the WT and G → R decapeptides [Ac-CVKFP(G,R)-IETC-$CONH_2$] (Figure 1) illustrates the various equilibrating species, as quenched by either the acid or iodoacetamide methods. Determination of C_{eff} requires quantitation of the concentrations of the fully reduced (or fully alkylated) and intramolecularly oxidized peptide species, only. Comparison of the profiles for the wild-type sequence and the G → R mutant indicates a higher concentration of the intramolecularly oxidized species for the former and hence clearly demonstrates the greater tendency of the wild type to adopt folded conformations. Similar equilibria are obtained for WT and G → R dodecapeptides [Ac-CYVKFP(G,R)IETLC-$CONH_2$].

Table I gives the C_{eff} determined under a variety of conditions, including ratios of reduced to oxidized glutathione of 3:1 or 5:1 and the presence or absence of 4 M urea. Note that the WT peptide favors the oxidized species significantly more than the mutant G → R peptide under all conditions. Additionally, the dodecapeptide with a D-Arg residue shows a slightly stronger preference for the "folded" oxidized species than does wild type.

Addition of urea to the thiol–disulfide exchange analysis results in equilibrium constants that reflect the probability of collision of the cysteine thiols on the basis of the steric constraints of an amino acid sequence. Urea addition resulted in only small differences in the measured effective concentrations, suggesting that specific interactions (hydrogen bonding and hydrophobic effects) do not contribute extensively to the conformational preferences of either the WT or the G → R deca- and dodecapeptides.

In order to dissect the interactions that may stabilize the folded form of these peptides, the side chain of Lys in the wild-type peptide was acetylated. In a hairpin conformation, this residue would be within salt-bridge distance of the Glu. The result (Table I) shows that the modified peptide has essentially the same tendency to fold (within experimental error) as does the wild type, indicating that there is no stabilizing effect of a Lys/Glu salt bridge under the conditions of the assay. The presence of 0.2 M KCl in the disulfide assay may deemphasize the importance of electrostatic interactions in the folding equilibria.

Relatively small differences in the measured effective concentrations were observed between reactions quenched with acid and reactions quenched with IAcm, indicating the disadvantages associated with either quenching method are cir-

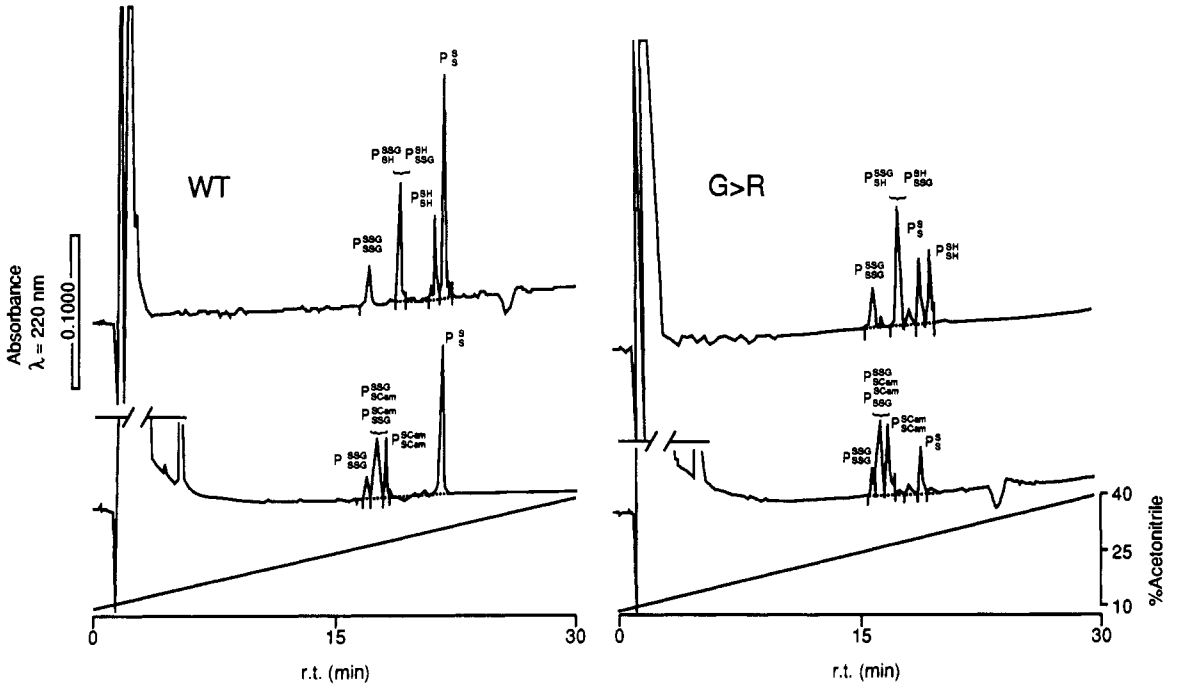

FIGURE 1: Representative HPLC chromatograms of thiol–disulfide exchange assay performed on P22 tailspike peptides. Shown are the results of equilibrating either the wild-type decapeptide, Ac-CVKRPGIETC-$CONH_2$ (left), or the tsH304 G → R mutant decapeptide, Ac-CVKRPRIETC-$CONH_2$ (right), in 3:1 GSH:GSSG exchange buffer. The exchange was quenched with acid in the top trace and with IAcm in the bottom trace. Other experimental details are given in the text. Note the relative intensities of the intramolecularly oxidized and fully reduced peptide peaks for the two sequences.

Table I: Effective Thiol Concentrations[a] for Deca- [Ac-CVKFP(G,R)IETLC-$CONH_2$] and Dodecapeptides [Ac-CVKFP(G,R)IETC-$CONH_2$] Corresponding to Sequences from P22 Tailspike Protein and tsH304 G_{244} → R Temperature-Sensitive Folding Mutant

	GSH:GSSG			
			+4 M urea	
	3:1	5:1	3:1	5:1
	Acid Quench (Decapeptide)			
wild type	12 ± 2 (3)	21.1 ± 0.8 (4)	11.2 ± 0.4 (1)	19 ± 2 (3)
G → R mutant	3.5 ± 1 (3)	5 ± 2 (4)	3.8 ± 0.9 (2)	5 ± 1 (3)
wild type, Lys(Ac)[b]	14.4 ± 0.1 (1)	ND[c]	ND	ND
	IAcm Quench (Decapeptide)			
wild type	21 ± 4 (3)	28 ± 4 (3)	19 ± 2 (2)	26.4 ± 0.6 (2)
G → R mutant	3 ± 1 (3)	5 ± 2 (4)	4.0 ± 0.2 (2)	6.9 ± 0.8 (3)
wild type, Lys(Ac)[b]	23.6 ± 0.1 (1)	ND	ND	ND
	Acid Quench (Dodecapeptide)			
wild type	11.0 ± 0.6 (1)	19 ± 3 (2)	ND	12.6 ± 0.1 (1)
G → R mutant	4.6 ± 0.9 (4)	7.2 ± 0.8 (5)	2.8 ± 0.5 (3)	ND
D-Arg	15.9 ± 0.5 (1)	22 ± 2 (4)	ND	20 ± 2 (2)
	IAcm Quench (Dodecapeptide)			
wild type	ND	27.7 ± 0.6 (1)	ND	ND
G → R mutant	3.0 ± 0.3 (1)	2.0 ± 0.1 (1)	6.7 ± 0.5 (1)	ND
D-Arg	ND	ND	ND	ND

[a] Values reported are C_{eff} (mM) determined in triplicate ± one standard deviation with the number of times the experiment was repeated reported in parentheses. [b] Ratio of peak areas per absorbance unit (AU) for oxidized and reduced species of acetylated derivative assumed equal to ratios for WT peptide which are 18.4 and 11.6 nmol/AU, respectively. [c] ND = not determined.

cumvented by the techniques used in the assay (viz., rapid analysis of acid-quenched reactions and short reaction times required for carboxyamidomethylation).

The errors associated with the reported effective concentrations are quite small for reactions run simultaneously but increase to 20–50% when reactions are not run simultaneously. These errors are comparable to the error estimates found by others (Zhang & Snyder, 1989). The two different glutathione ratios (5:1 and 3:1) were used due to the large difference in equilibrium constants between the WT and G → R sequences. The effective concentration is a ratio of oxidized to reduced peptide concentrations as determined by HPLC analysis. A ratio of approximately 1 is desired to minimize the error associated with dividing a large number by a small one or vice versa. As can be seen in Figure 1, use of a GSH/GSSG ratio of 3:1 results in an approximately 1:1 ratio of reduced to oxidized G → R decapeptide derivatives while the oxidized species is greatly favored for the WT decapeptide.

Analyses not performed simultaneously suffer larger errors probably due to slight differences in pH, temperature, trace amounts of oxygen, or the glutathione ratios. While C_{eff} should be independent of the GSH:GSSG ratio, error in glutathione concentration probably accounts for the observed differences in C_{eff} for the same peptide in assays performed in 3:1 and 5:1 GSH:GSSG ratios. Sufficient sensitivity for distinguishing the effects of urea and the synthetic modification of the lysine

side chain by acetylation with the thiol–disulfide exchange assay is only obtained for analyses run simultaneously.

From the equal areas of the mixed one disulfide species (Figure 1), no indication of a preferential interaction of glutathione with the N-terminal cysteine over the C-terminal cysteine was observed as might be expected if there existed a pK_a difference between the two cysteines (Jocelyn, 1967). Differences in rates of thiol–disulfide exchange between glutathione and the individual cysteines may exist and are expected to be on the order of seconds (Snyder, 1987); however, only equilibrium concentrations were measured in the experiments described here.

No covalent dimerization was observed for any of the peptides under the conditions of the equilibrium thiol–disulfide exchange assay. However, air oxidation of the G → R dodecapeptide resulted in formation of a dimer which was sparingly soluble in water (polarity of the linkage in the dimers was not determined). A thiol–disulfide exchange reaction starting with the dimeric wild-type peptide gives the same equilibrium constants as analyses begun with either the intramolecularly cyclized disulfide or reduced peptide, with quantitative conversion of the dimer into the monomeric species. The dimeric species elutes from the HPLC column well resolved from the other species and was not present in samples subjected to the disulfide exchange assay conditions. Precipitation of the dimeric species (and thus its absence in the analysis) is unlikely since the use of 4 M urea, in which the solubility of the G → R dodecapeptide dimer is greatly improved, results in no detectable dimeric species.

Nuclear Magnetic Resonance. NMR results for the linear decapeptides blocked by methylcarboxyamidylation showed no indication from NOEs or NH temperature dependences of any preferred conformations in aqueous solution, even upon cooling to 1 °C. One of the most diagnostic NMR parameters for a population of β turn is the NOE between the NH of residue i and the NH of residue $i+1$ [e.g., Dyson et al. (1988)]. However, in a decapeptide where NOEs corresponding to distances larger than 4 Å are rarely observed, one can estimate that a population of turn conformer greater than 25% would be necessary to see the NH to NH NOE. The disulfide exchange assay can assess more sensitively a small bias toward a turn and, furthermore, a *difference* in tendency to adopt a turn for two sequences.

The analysis of the cyclized tailspike peptides by NMR is more straightforward. A summary of the most informative NMR parameters for the wild-type and mutant oxidized decapeptides is given in Figure 2. The large H^{α} to NH NOEs from Cys to Gly (or Arg) and from Ile to Cys are supportive of an extended conformation (as opposed to local conformation in the helical region of ϕ,ψ space). The strong NOEs between the Pro H^{α} and the Gly NH and between the NH Gly and the NH Ile support the presence of a type II β turn; interestingly, the *same* feature appears to be present in the Arg-containing tsf sequence. Quantitative analysis of the buildup rate of the pro H^{α} to Gly(or Arg) NH NOE was carried out by variation of mixing time and integration of the cross-peak volume in order to estimate the population of this turn type. In both WT and G → R decapeptides, the type II form is clearly predominant (≥90%). Additionally, no significant difference is found between the conformations adopted by the WT and G → R peptides. Consistent with an overall hairpin conformation are the NH accessibilities determined either by temperature dependence or by deuterium exchange: the NHs of Ile, Thr, Phe, and Val are of reduced accessibility to solvent by one or both criteria. $J_{N\alpha}$ coupling constants are also consistent with the deduced conformation. The decapeptides yielded analogous results in the dimethyl sulfoxide/chloroform solvent mixture [data not shown; for further details see Stroup (1989)]. No NOEs other than those from adjacent residues were observed for any of the peptides studied.

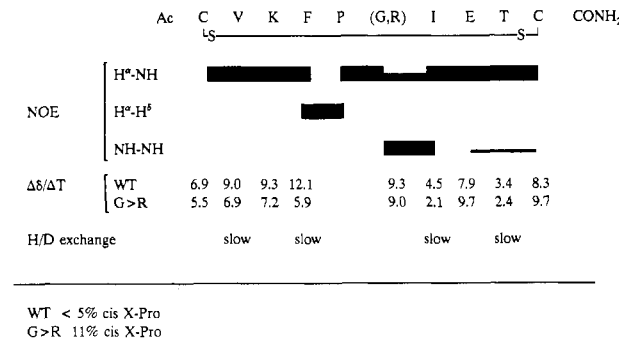

FIGURE 2: Summary of conformationally diagnostic NMR data for the oxidized WT and G → R decapeptide tailspike peptides. The filled bars represent interproton NOEs, with the height of the bar indicating the relative magnitude of the NOE. H–D exchange rates (pH 3, freshly lyophilized peptide dissolved in 100% D_2O) are considered slow if $t_{1/2} > 1$ h. Temperature dependences are given in ppb/deg.

Discussion

We have tested the hypothesis that the formation of turns in local sequences of the P22 tailspike is critical to its productive folding by "excising" regions of the protein for study as isolated peptides. By flanking the sequences of interest from wild-type protein and from a tsf mutant with cysteines, we were able to use a thiol–disulfide exchange assay to determine the tendencies of the P22 tailspike peptides to adopt reverse turn conformations. This assay yields a sensitive measure of the fraction of the conformers that bring the two thiols close in space. Indeed, both the deca- and dodecapeptide wild-type sequences favor the oxidized, "folded" species significantly more strongly (3–5-fold) than do the corresponding peptides harboring the tsf G → R mutation. The C_{eff}'s of the wild-type peptide were approximately 20 mM. A peptide synthesized with a D-Arg in place of the wild-type Gly showed a folding tendency slightly greater than that of wild type, as would be expected if a Pro-X reverse turn leads to the approach of the chain ends. The observation that addition of urea did not markedly alter the relative or absolute magnitudes of the wild-type and mutant C_{eff}'s (vide infra) argues persuasively that the origin of the different folding behavior of these two sequences is largely steric. Consistent with this conclusion was the finding that the contribution of one potential specific interaction, a salt bridge between Lys and Glu, was negligible (based on study of an acetylated Lys-containing peptide). Study of peptides corresponding to other regions of the P22 tailspike protein may reveal alternative sources of folding defects from point mutations.

Sequence considerations and model building argue that the β-hairpin conformation favored by the intramolecular disulfide bond would place the Pro-Gly or Pro-Arg residues in the corner positions of a β turn. NMR analysis indeed indicates that both the WT and G → R oxidized deca- and dodecapeptides adopt β turns with type II geometry. While a type II β turn is favored for the Pro-Gly sequence, Pro-Arg would be expected to favor type I β-turn geometry on the basis of the chirality of the amino acids in the putative turn (Venkatachalam, 1968). The difference in the tendency for the WT and G → R sequences to fold into the turn conformation may therefore reflect constraints from the flanking sequence that force the Pro-Arg corner residues to adopt the less likely turn type. The

overall conformation of this excised fragment of the P22 tailspike places the hydrophobic residues on one face of the molecule and the hydrophilic residues on the opposite side. The amphiphilic arrangement of residues in the resulting β hairpin may be important in the oligomerization process and thermostability of the tailspike protein.

The present study has thus enabled an assessment of the likelihood of a highly flexible peptide to adopt folded conformations that may be important to a protein folding pathway. Wright et al. (1988) have previously suggested from NMR results on small flexible peptides that biases toward turn conformations may predispose a polypeptide to fold into secondary structures, including helices. It is noteworthy that the nuclear Overhauser effects they observed revealed *local* structural tendencies (within neighboring residues), while the C_{eff} parameter measures a tendency for proximity of more distant parts of the chain. Furthermore, in the present family of peptides, no NOEs indicating preferred conformations were observed for the linear peptides. Yet, the thiol–disulfide assay enabled comparable sampling of the tendency toward folding into a reverse turn by both sequences. Most importantly, results from both the NMR and disulfide exchange approaches suggest that highly flexible small peptides may reveal conformational preferences that are critical to the folding of longer polypeptides.

It is striking that the tendency to form the folded, turn-containing conformation is quite small in these sequences and that the difference between wild-type and tsf mutant sequences is only a factor of 3–5, or about 1 kcal/mol. Nonetheless, this residue substitution influences a critical folding step in the intact protein, and the local perturbation is likely to be the source of its effect. C_{eff} of cysteine pairs in proteins can vary over several orders of magnitude. The effective concentrations observed in this study (approximately 20 mM for the WT deca- or dodecapeptides or D-Arg dodecapeptide and 3–5 mM for the G → R deca- or dodecapeptides) are quite small when compared to the effective concentration of the 5–55 disulfide bond in BPTI, which is 1×10^7 M (Creighton, 1983). Under denaturing conditions, the average effective concentration of the first 15 random disulfides formed for BPTI is 0.3 M (Creighton, 1977) while the C_{eff} for Ac-Cys-$(Gly)_6$-Cys-$CONH_2$ in either the presence or absence of urea is 60 mM (Lin & Kim, 1989). The C_{eff}'s for the P22 tailspike peptides are on the order of those observed for peptides used to model small disulfide loops in proteins, which ranged from 1.5 to 20 mM (Zhang & Snyder, 1989).

Equilibrium constants for formation of the intramolecular disulfide bond may depend on the number of residues separating the cysteines. Since similar C_{eff} values were found for both the deca- and dodecapeptides, the additional residues in this case exert no detectable stabilization of the β hairpin, nor does the difference in cycle size appear to alter significantly the probability of the intramolecular collision. These findings add further support to our interpretation that it is the residues involved in turn formation that are at the root of the difference between the folding tendencies of the WT and mutant sequences.

The tsf phenotype is manifest only at high temperature in vivo. An obvious extension of the present studies, in addition to analysis of other tsf sites, is the assessment of the difference in turn tendency at various temperatures. Some modification of the disulfide exchange assay will be required. Another approach that will be explored is the analysis of kinetics of formation of the disulfide, as opposed to equilibrium determination (Snyder, 1987). This approach may be more applicable to the study of temperature dependence.

Conclusions

Our peptide models of the P22 tailspike protein and the tsH304 G → R mutant argue that the temperature-sensitive folding defect is a result of a lowered propensity to adopt a reverse turn due to the steric constraints of the amino acid substitution. NMR studies on the peptide models show that, once a turn conformation is formed, both wild-type and mutant sequences adopt the same type of β hairpin with a Pro-Gly or Pro-Arg type II β turn. As in the P22 tailspike protein, the folding defect is only manifested in the pathway toward turn formation and not in the final "folded" conformation of the peptide models. The local amino acid sequence identified by genetic selection for temperature sensitivity is critically important in the folding of the 666 residue long tailspike protein. Our biophysical analysis also suggests that small differences in an equilibrium distribution over conformational states can make a major impact on the productive adoption of native three-dimensional structure in vivo. Furthermore, our results support the use of "excised" fragments of a large protein to examine critical early steps in the folding pathway.

Acknowledgments

Mass spectrometry results were provided by B. Pramanik and C. Slaughter, amino acid analysis was performed by K. K. Ng, and synthesis of decapeptides was performed by S. J. Stradley. We thank P. Kim for technical guidance on the thiol–disulfide exchange assay and J. King and co-workers for stimulating discussions on P22 tailspike folding.

References

Creighton, T. E. (1977) *J. Mol. Biol. 113*, 313–328.
Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol. 33*, 231–297.
Creighton, T. E. (1983) *Biopolymers 22*, 49–58.
Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., & Lerner, R. A. (1988) *J. Mol. Biol. 201*, 161–200.
Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70–77.
Goldenberg, D. P., Smith, D. H., & King, J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 7060–7064.
Goto, Y., & Hamaguchi, K. (1981) *J. Mol. Biol. 146*, 321–340.
Haase-Pettingill, C. A., & King, J. (1988) *J. Biol. Chem. 263*, 4977–4983.
Jocelyn, P. C. (1967) *Eur. J. Biochem. 2*, 327–331.
Kaiser, E., Colescott, R. L., Bossinger, C. D., & Cook, P. I. (1970) *Anal. Biochem. 34*, 595–598.
Kaiser, E., Bossinger, C. D., Colescott, R. L., & Olsen, D. B. (1980) *Anal. Chim. Acta 118*, 149–151.
Kelley, R., Wilson, J., Bryant, C., Shalongo, W., Ledger, R., Lowery, D., & Stellwagen, E. (1986) in *Folding of Escherichia coli Thioredoxin into Its Native Conformation* (Holmgren, A., et al., Eds.) pp 77–85, Raven Press, NY.
King, J. (1986) *Bio/Technology 4*, 297–303.
King, J. (1990) in *Protein Folding: Deciphering the Second Half of the Genetic Code* (Gierasch, L. M., & King, J., Eds.) pp 211–219, AAAS Publications, Washington, DC.
Lin, T.-Y., & Kim, P. S. (1989) *Biochemistry 28*, 5282–5287.
Milburn, P. J., Konishi, Y., Meinwald, Y. C., & Scheraga, H. A. (1987) *J. Am. Chem. Soc. 109*, 4486–4496.
Oas, T. G., & Kim, P. S. (1988) *Nature (London) 336*, 42–48.
O'Shea, E. K., Rutkowski, R., & Kim, P. S. (1989) *Science 243*, 538–542.
Riddles, P. W., Blakeley, R. L., & Zerner, B. (1979) *Anal. Biochem. 94*, 75–81.

Roder, H., Elove, G., & Englander, S. W. (1988) *Nature (London) 335*, 700–704.
Sargent, D., Benevides, J. M., Yu, M.-H., King, J., & Thomas, G. J., Jr. (1988) *J. Mol. Biol. 199*, 491–502.
Saxena, V. P., & Wetlaufer, D. B. (1970) *Biochemistry 9*, 5015–5022.
Smith, D. H., & King, J. (1981) *J. Mol. Biol. 145*, 653–676.
Snyder, G. H. (1987) *Biochemistry 26*, 688–694.
Stewart, J. M., & Young, J. D. (1984) in *Solid Phase Peptide Synthesis*, 2nd ed., Pierce Chemical Co., Rockford, IL.
Stroup, A. N. (1989) Ph.D. Dissertation, University of Texas Southwestern Medical Center.
Venkatachalam, C. M. (1968) *Biopolymers 6*, 1425–1436.
Villafane, R., & King, J. (1988) *J. Mol. Biol. 204*, 607–619.
Weber, U., & Hartter, P. (1974) *Hoppe-Seyler's Z. Physiol. Chem. 355*, 189–199.
Wright, P. E., Dyson, H. J., & Lerner, R. A. (1988) *Biochemistry 27*, 7167–7175.
Yu, M.-H., & King, J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6584–6588.
Yu, M.-H., & King, J. (1988) *J. Biol. Chem. 263*, 1424–1431.
Zhang, R., & Snyder, G. H. (1988) *Biochemistry 27*, 3785–3794.

Articles

Direct and Indirect Pathways of Functional Coupling in Human Hemoglobin Are Revealed by Quantitative Low-Temperature Isoelectric Focusing of Mutant Hybrids†

Vince J. LiCata,‡ Phil C. Speros, Ermanna Rovida,§ and Gary K. Ackers*,‖

Department of Biology, The Johns Hopkins University, Baltimore, Maryland 21218
Received April 16, 1990; Revised Manuscript Received July 16, 1990

ABSTRACT: Functional energetic coupling within human hemoglobin has been explored by using quantitative analysis of asymmetric mutant hybrid equilibria. Previous work showed that the free energy of cooperativity is largely attributable to alterations in free energy that accompany changing interactions at the interface between $\alpha^1\beta^1$ and $\alpha^2\beta^2$ dimers [Pettigrew et al. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1849]. However, the issue of how cooperativity-linked sites in the molecule are energetically coupled in manifesting cooperative ligation is still not well delineated. In this paper we address the questions of what types of functional coupling pathways are operational in hemoglobin, what some of their characteristics are, and how they are related to one another. By constructing asymmetric mutant hybrid hemoglobins, we can assay how two structurally identical, symmetrically equivalent sites are energetically coupled in manifesting subunit assembly and/or cooperative ligation. Asymmetric hybrid hemoglobins, i.e., those containing a single modified site, cannot be isolated and must be studied in equilibrium with their symmetric parent molecules. In order to study these asymmetric hybrid equilibria, we have developed new theory and quantitation techniques to augment the low-temperature quenching and isoelectric focusing procedures of Perrella et al. [(1978) *Anal. Biochem. 88*, 212]. Studies of these mutant hybrid hemoglobins have provided evidence for three distinct types of energetic coupling within the hemoglobin tetramer. All $\alpha^1\beta^2$ interface sites examined are involved in coooperativity-linked indirect coupling. Within the context of this indirect "pathway" there exist two different types of direct long-range coupling. One of these classes of direct long-range pathways is linked to cooperative ligand binding while the other class is not.

A key issue in structure–function studies of allosteric proteins is how regulatory energy is communicated between relatively distant sites within a molecule or macromolecular assembly. When particular residue sites within a molecule have been identified as being on the pathway for manifesting a particular function, one may ask in what way do these sites particpate with one another in generation of that function. There are two possibilities for the nature of the pathways by which regulatory information is communicated energetically from site to site in a macromolecule:

One possibility is by *indirect pathways*, wherein energetic communication between two functionally important sites is mediated through some other part of the molecule but where the two sites themselves are not directly coupled to one another. Such an effect might involve a simple "triangulation pathway" through a third residue site or could be manifested by a global mechanism wherein overall changes in conformation or solvation are propagated throughout an interface, a domain, or the entire molecule.

† This work was supported by NIH Grants R37-GM24486 and HL-40453 and NSF Grant DMB86-15497.
* To whom correspondence should be addressed.
‡ Present address: Department of Chemistry, University of Wisconsin, Madison, WI 53706.
§ Present address: Dipartmento Scienze e Tecknologie Biomedicke, Ospedale San Raffaele, Milano, Italy.
‖ Present address: Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110.

0006-2960/90/0429-9771$02.50/0