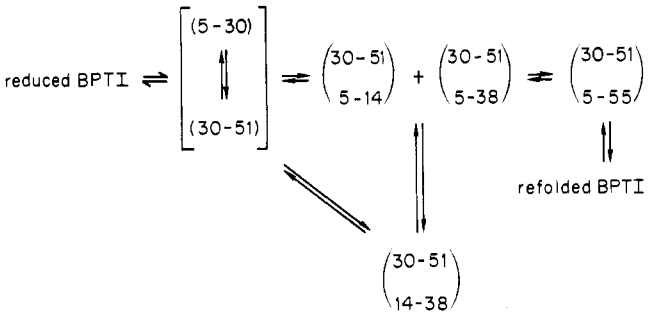
Circular Dichroism Spectroscopy of the Intermediates That Precede the Rate-Limiting Step of the Refolding Pathway of Bovine Pancreatic Trypsin Inhibitor. Relationship of Conformation and the Refolding Pathway†

Phyllis Anne Kosen,‡ Thomas E. Creighton, and Elkan R. Blout*

ABSTRACT: Circular dichroism spectra of the partially folded trapped intermediates were measured in order to aid in the elucidation of the conformational forces which determine a nonrandom, nonsequential pathway of disulfide bond formation upon refolding of bovine pancreatic trypsin inhibitor. Whatever conformation was responsible for the kinetic rates of the intermediates should be stabilized by the presence of their trapped disulfide bonds. The near-ultraviolet spectra provide considerable information about the environments of the aromatic and disulfide side chains. The predominant single-disulfide intermediate has significant nonrandom conformation not present in the fully reduced protein, with aromatic rings and the disulfide bond in stabilized asymmetric environments. Forming either of the two nonnative, but kinetically important, second disulfides in this intermediate does not produce unequivocably different conformations. Forming a second native, but kinetically unproductive, disulfide produces a substantial decrease in randomness, which may hinder formation of the third disulfide. The largest conformational changes occur upon disulfide rearrangement to the stable, correctly refolded, two- and three-disulfide species. Interpretation of the far-ultraviolet spectra in terms of the secondary structure of the intermediates is uncertain, due to the atypical spectra of the folded forms of the protein. Consequently, we are unable to determine unambiguously the secondary structure of the intermediates. However, all the spectra show that nonrandom conformations of the polypeptide chain gradually appear as disulfide bond formation progresses, as expected from the nonrandom pathway of the latter.

Bovine pancreatic trypsin inhibitor (BPTI)[1] is the only protein for which the pathway of refolding has been determined experimentally. It was elucidated by trapping and isolating the small number of intermediates with one or two disulfide bonds which accumulate kinetically during unfolding and refolding [reviewed by Creighton (1978, 1980)]. When the various species are designated by the residue numbers of the cysteine residues paired in disulfides, the major pathway may be summarized as



The parentheses around the single-disulfide intermediates indicate that they are in rapid equilibrium by intramolecular disulfide interchange; the two predominant, most stable intermediates are indicated. The "+" between two of the two-disulfide species indicates that both have the same kinetic role.

The conformational forces which direct this nonrandom, but nonsequential, pathway of disulfide bond formation and breakage are not yet known. They should be elucidated most readily from the conformational properties of the trapped intermediates, since whatever conformation favors formation of a particular disulfide should be stabilized to the same extent by the presence of the trapped disulfide (see Discussion). The conformations of the folding intermediates have been examined previously by UV difference spectroscopy (Kosen et al., 1980) and immunochemical measurements (Creighton et al., 1978). We were particularly interested in their circular dichroism spectra to determine at what stage of folding the secondary structure appears. However, calculations of the secondary structure by using the very similar far-UV CD spectra of the native protein and the nativelike species (30–51, 5–55) were found (Kosen et al., 1981) to be significantly different from those expected from their known secondary structure (Huber et al., 1971; Deisenhofer & Steigemann, 1975), so interpretation of the far-UV CD spectra of the intermediates is perilous. Nevertheless, the near-UV CD spectra provide substantial information about the environments of the aromatic and disulfide side chains.

Materials and Methods

The materials and the spectral procedures used were those described previously (Kosen et al., 1981). The CD spectra

† From the Department of Biological Chemistry, Harvard Medical School, Boston, Massachusetts 02115, and the Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, England. *Received January 10, 1983.* This work was supported by U.S. Public Health Service Grant AM07300 and by Predoctoral National Institutes of Health Training Grant GM-0045. This study is from the Ph.D. Dissertation of P.A.K.

* Address correspondence to this author at the Department of Biological Chemistry, Harvard Medical School.

‡ Present address: Department of Chemistry, Reed College, Portland, OR 97202.

[1] BPTI is the abbreviation used for bovine pancreatic trypsin inhibitor, irrespective of the states of the disulfide bonds. The protein, prior to unfolding and reduction, is referred to as native BPTI. The protein after reoxidation and refolding is referred to as refolded BPTI. R is the fully denatured protein with the six cysteines carboxamidomethylated. Other forms of BPTI are designated by the residue numbers which are involved in disulfide bonds. For example, a one-disulfide-bond species with the 30–51 disulfide is denoted as (30–51). All thiol groups have been carboxymethylated. Other abbreviations: CD, circular dichroism; Pipes, piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); UV, ultraviolet; [θ], mean residue ellipticity (deg·cm^2/dmol); θ_0, λ_0, and Δ_0, the maximal ellipticity, the wavelength position, and the band half-width of a resolved Gaussian band, respectively.

0006-2960/83/0422-2433$01.50/0

Table I: Tentative Assignments and Parameters of the Resolved Gaussian Bands in the Spectra of (5-30), (30-51), (30-51, 5-14) plus (30-51, 5-38), and (30-51, 14-38)[a]

	(5-30)			(30-51)			(30-51, 5-14) plus (30-51, 5-38)			(30-51, 14-38)		
tentative assignment	λ_0	Δ_0	θ_0	λ_0	Δ_0	θ_0	λ_0	Δ_0	θ_0	λ_0	Δ_0	θ_0
Tyr												
0-0 (1)[b]		*c*		281.5	6.8	−48	281.5	6.8	−48	281.8	6.3	−125
0 + 800 (2)		*c*		274.9	3.6	−15	274.9	3.6	−15	275.4	3.7	−69
0 + 2(800) (3)		*c*	0	270.4	4.5	−21	270.5	4.8	−20	269.4	5.9	−83
-S-S-												
30-51 (4)				264.6	24.9	−106	264.6	24.9	−106	264.1	27.3	−102
5-30 (5)	272.5	17.1	−61									
5-14 plus 5-38 (6)				269.7	25.7	−46						
14-38 (7)										273.0	19.5	−85
Phe												
0 + 930 (8)		*d*			*d*			*d*		260.1	2.0	−11

[a] Assignments are based on studies by Horwitz et al. (1969, 1970), Kahn (1972), Holladay et al. (1974, 1976), and Holladay & Puett (1976). λ_0 and Δ_0 are expressed in nanometers; θ_0 is expressed in deg·cm²/dmol on a mean residue basis. [b] Values in parentheses correspond to band numbers in Figures 2-5. [c] There was no evidence for the presence of tyrosine transitions in the spectrum of (5-30). [d] There was no evidence for the presence of phenylalanine transitions in the spectra of (30-51), (30-51, 5-14) plus (30-51, 5-38), or (5-30).

were expressed as mean residue ellipticity (deg·cm²/dmol), and mean residue concentrations were calculated by using the following molecular weights and mean residue weights, respectively: one-disulfide intermediates, 6741, 116.2; two-disulfide intermediates, 6627, 114.3.

The technique of curve resolution and its limitations have been discussed in detail previously (Kosen et al., 1981). Reiterating briefly, Gaussian curves were added, subtracted, and altered in position, intensity, and half-width until their sum visually agreed with an experimental spectrum displayed on a graphics terminal. Then a computerized nonlinear least-squares fitting operation was performed to minimize the differences between the experimental and theoretical spectra. The parameters (Tables I–III) of the Gaussian curves, assumed to represent specific electronic transitions of the final theoretical spectra (Figures 2–6), conform with those obtained when model tyrosine, phenylalanine, or disulfide CD spectra were fitted (Horwitz et al., 1969, 1970; Kahn, 1972; Holladay et al., 1974, 1976; Strickland, 1974; Holladay & Puett, 1976). Indeed, the curve parameters in this study were chosen to conform with those of the aforementioned model studies. Since only a minimum number of curves fully fitting the experimental spectra were included finally, (1) not all the spectra have the same number of bands, (2) in many of the spectra, phenylalanine transitions were not included (their contribution to the total ellipticity would be predicted to be small), and (3) each tyrosine electronic–vibronic transition and, in some cases, the disulfide electronic transitions were represented as an averaged transition. These resolved spectra are undoubtedly not unique; nevertheless, they are internally self-consistent and consistent with other theoretical and experimental results. Thus, curve resolution has been an aid in the interpretation of the CD spectra reported here and previously (Kosen et al., 1981).

Results

The CD spectra of the fully reduced and carboxamidomethylated protein (R) of the refolded-like intermediate (30–51, 5–55) and of refolded BPTI (Kosen et al., 1981) are presented in Figure 1 for reference. That of R is characteristic of a fully unfolded protein, with no near-UV CD intensity and only negative ellipticity in the far-UV with a single minimum at 200 nm. The far-UV minimum is shifted to 202.5 nm in the spectrum of refolded BPTI, with a substantial shoulder between 215 and 225 nm; the near-UV CD spectrum is very intense and has previously been resolved into contributions by

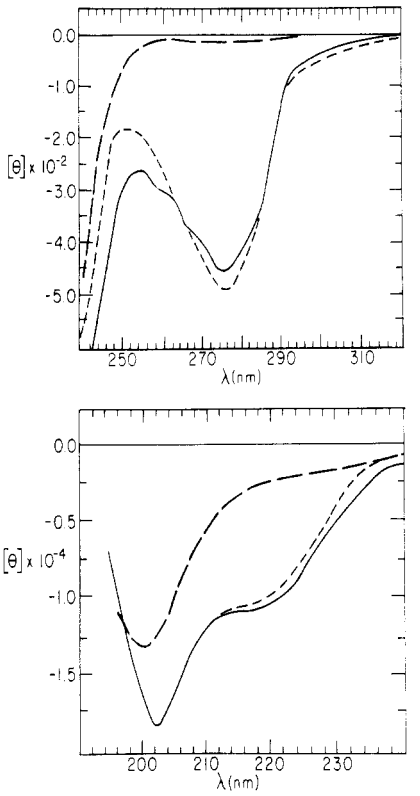


FIGURE 1: CD spectra of R (– –), (30–51, 5–55) (—), and refolded BPTI (- - -) in 6.0 mM Pipes, pH 6.8. (Top panel) Near-UV portion of the spectrum. (Bottom panel) Far-UV portion of the spectrum. [Taken in modified form from Kosen et al. (1981).]

disulfide, tyrosine, and phenylalanine side-chain electronic transitions (Kosen et al., 1981). The spectrum of intermediate (30–51, 5–55) is very similar to that of the refolded protein, differing primarily in the absence of one disulfide transition and some consequent small conformationally induced transition perturbations.

We present here the CD spectra of the partially folded intermediates. The near-UV spectra have been analyzed in the same way as those presented previously (Kosen et al., 1981). The tentative assignments and parameters of the resolved bands are presented in Table I; their rotational strengths are in Table II.

Intermediate (5–30). This intermediate is the second most stable single-disulfide intermediate, even though it has a

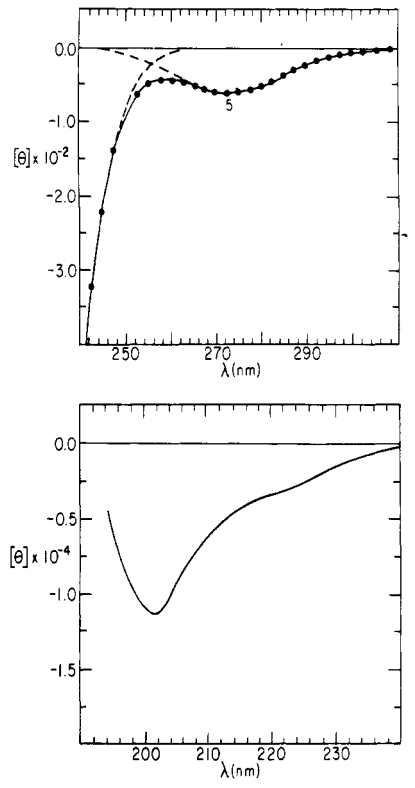


FIGURE 2: CD spectrum of (5–30) in 6.0 mM Pipes, pH 6.8. (Top panel) Near-UV portion of the spectrum: experimental values (●); resolved Gaussian bands (---); sum of the Gaussian bands (—). The band at lowest wavelengths was included only to completely fit the spectrum; no assignment was made for this band. The other band is numbered 5 and corresponds to the band number in Table I. $[\theta]_{272.5}$ = −60 ± 2 deg·cm²/dmol. (Bottom panel) Far-UV portion of the spectrum. $[\theta]_{202.5}$ = −11 150 ± 300 deg·cm²/dmol.

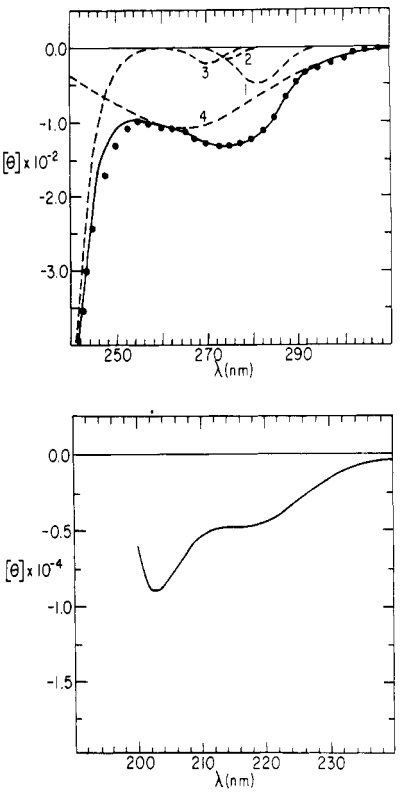


FIGURE 3: CD spectrum of (30–51) in 6.0 mM Pipes, pH 6.8. (Top panel) Near-UV portion of the spectrum: experimental values (●); resolved Gaussian bands (---); sum of the Gaussian bands (—). The band at lowest wavelengths was included only to completely fit the spectrum; no assignment was made for this band. All other bands are numbered and correspond to the band numbers in Table I and in the text. $[\theta]_{272.5}$ = −130 ± 2 deg·cm²/dmol. (Bottom panel) Far-UV portion of the spectrum. $[\theta]_{202.5}$ = −8900 ± 200 deg·cm²/dmol.

nonnative disulfide bond. Its near-UV CD spectrum (Figure 2) consists of a single, weak, and broad minimum characteristic of a transition due to a rather flexible disulfide bond (Kahn, 1972; Takagi & Ito, 1972) (Tables I and II; band 5). There was no fine structure between 275 and 285 nm, typical of tyrosine transitions, suggesting that these residues are not involved in nonrandom conformations. The far-UV CD spectrum is similar to that of R, with no significant shoulder between 215 and 225 nm, but the single minimum was shifted to 202.5 nm like that of the refolded BPTI spectrum.

Intermediate (30–51). This is the most stable [by 0.6 kcal/mol relative to (5–30)] single-disulfide intermediate, with a nativelike disulfide bond which is retained in further productive folding. Its far-UV CD spectrum (Figure 3) has only half the intensity of that of refolded BPTI and less than that of R, but it has the nativelike position at 202.5 nm and a substantial shoulder between 215 and 225 nm.

The near-UV CD portion of the (30–51) spectrum is characterized by substantial negative ellipticity, with a minimum in the region of 275 nm, but ellipticity also above 290 nm and especially between 250 and 260 nm. These latter aspects are unlikely to be due to anything but a broad disulfide transition (Horwitz et al., 1970; Kahn, 1972; Strickland, 1974; Holladay & Puett, 1976; Holladay et al., 1976). The spectrum has been fitted with Gaussian bands representing the (30–51) disulfide transition (band 4) and three bands each representing the average of four possible 0–0 (band 1), 0 + 800 (band 2), and 0 + 2(800) cm^{-1} (band 3) tyrosine vibronic transitions associated with the L_b electronic transition (Tables I and II).

The disulfide bond parameters assigned in the spectrum of (30–51) are similar to those assigned previously to disulfides (30–51) and (5–55) in the spectrum of (30–51, 5–55): λ_0 = 262.3 nm; Δ_0 = 26.3 nm; θ_0 = −270 deg·cm²/dmol, with an average rotational strength of -970×10^{-42} cgsu per disulfide (Kosen et al., 1981). The rotational strength per disulfide approaches the maximum value expected for a disulfide (Kahn, 1972; Takagi & Ito, 1972; Strickland, 1974), suggesting that the disulfide in this early intermediate shows greatly hindered mobility and possibly has a conformation comparable to that in the folded forms of BPTI.

Tyrosine residues also contribute to the near-UV CD spectrum of (30–51), as there is fine structure between 275 and 285 nm. The rotational strength of these transitions (Table II) is about that found for small model tyrosine compound spectra (Horwitz et al., 1970; Strickland, 1974; Holladay & Puett, 1976), which suggests that the tyrosines of (30–51) retain a significant degree of conformational mobility but that their environments differ from those of R, as its near-UV CD spectrum is nearly zero at all wavelengths.

Intermediates (30–51, 5–14) plus (30–51, 5–38). These two intermediates, with the same kinetic role, are always obtained as a mixture, with comparable amounts of the two (Creighton, 1975a, 1977a). The far-UV portion of this spectrum (Figure 4) is characterized by a minimum at 202.5 nm and a shoulder at higher wavelengths, both of which are strikingly like that of (30–51). This suggests that there is little alteration in the secondary structure of (30–51) on formation of either the (5–14) or the (5–38) disulfide.

The near-UV CD spectrum of (30–51, 5–14) plus (30–51, 5–38) was fitted by using the bands of the (30–51) spectrum plus one additional band for the (5–14) and (5–38) disulfide transitions (band 6) (Tables I and II). The magnitude of band

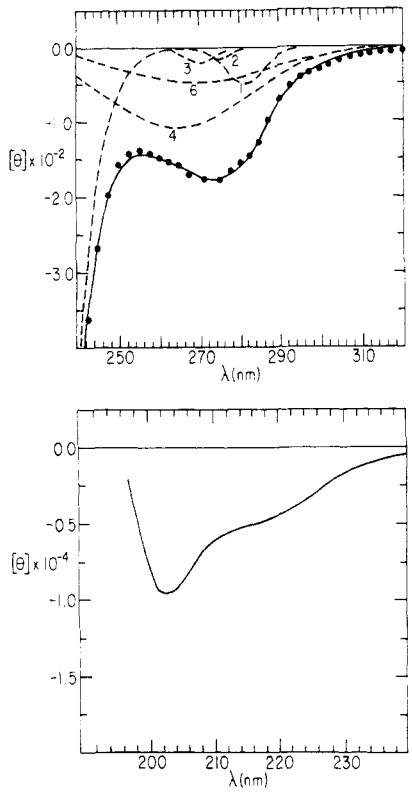


FIGURE 4: CD spectrum of (30–51, 5–14) plus (30–51, 5–38) in 6.0 mM Pipes, pH 6.8. (Top panel) Near-UV portion of the spectrum: experimental values (●); resolved Gaussian bands (---); sum of the Gaussian bands (—). The band at lowest wavelengths was included only to completely fit the spectrum; no assignment was made for this band. All other bands are numbered and correspond to the band numbers in Table I and in the text. $[\theta]_{275} = -175 \pm 2$ deg·cm²/dmol. (Bottom panel) Far-UV portion of the spectrum. $[\theta]_{202.5} = -9450 \pm 350$ deg·cm²/dmol.

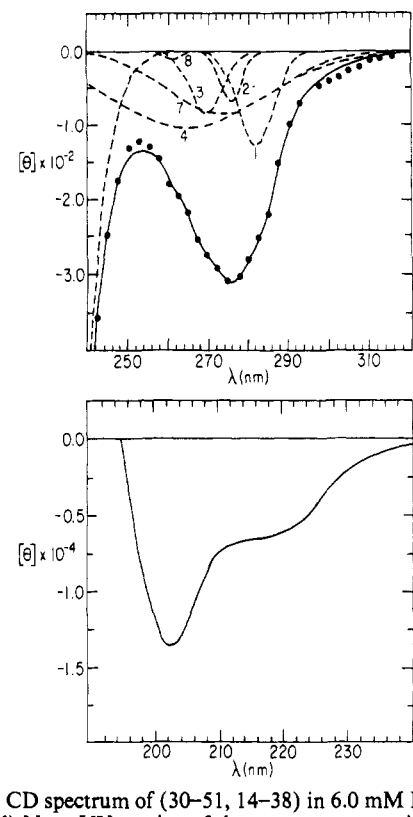


FIGURE 5: CD spectrum of (30–51, 14–38) in 6.0 mM Pipes, pH 6.8. (Top panel) Near-UV portion of the spectrum: experimental values (●); resolved Gaussian bands (---); sum of the Gaussian bands (—). The band at lowest wavelengths was included only to completely fit the spectrum; no assignment was made for this band. All other bands are numbered and correspond to the band numbers in Table I and in the text. $[\theta]_{275} = -310 \pm 8$ deg·cm²/dmol. (Bottom panel) Far-UV portion of the spectrum. $[\theta]_{202.5} = -13\,550 \pm 320$ deg·cm²/dmol.

6 in Figure 4 suggests that the 5–14 and 5–38 disulfides have significant flexibility. The ease of fitting the spectrum of (30–51, 5–14) plus (30–51, 5–38) with minimal alteration of the tyrosine transitions is consistent with the previous finding by difference spectroscopy (Kosen et al., 1980) that the solvent accessibilities, and thus probably the environments, of the tyrosines of (30–51) are nearly unaltered when either the (5–14) or the (5–38) disulfide is made. All available spectral data are consistent with the overall conformations of (30–51), (30–51, 5–14), and (30–51, 5–38) being nearly equivalent.

Intermediate (30–51, 14–38). The far-UV CD spectrum of (30–51, 14–38) shown in Figure 5 is similar in shape to that of native BPTI, but only about 75% as intense. Nevertheless, it is substantially different from (30–51); so forming the 14–38 nativelike disulfide produces substantial changes in conformation.

Resolution of the near-UV portion of the spectrum was achieved by addition of a disulfide transition and alteration of the bands in the (30–51) resolved spectrum. A phenylalanine transition (band 8) was also included in the resolution of this spectrum, but its contribution is small, as expected (Horwitz et al., 1969; Holladay et al., 1974; Strickland, 1974).

The rotational strengths of the tyrosine transitions (Table II) in the spectrum of (30–51, 14–38) are more than 3 times as large as those in the spectrum of (30–51). This and a comparable decrease in solvent accessibility (Kosen et al., 1980) imply a significant alteration in the tyrosine environments when the 14–38 disulfide is formed, which contrasts with the relatively small changes found if disulfides 5–14 or 5–38 are formed instead.

In fitting the (30–51, 14–38) spectrum, only slight alterations were required in the curve (band 4) previously assigned to the 30–51 disulfide; this may imply that the 30–51 disulfide and the surrounding region of the protein are not significantly altered on formation of the 14–38 disulfide. The resolved curve (band 7) assigned to the 14–38 transition in the spectrum of (30–51, 14–38) has a lesser intensity than that of 30–51 but is approximately double that assigned to the same disulfide band in the spectrum of the native or the refolded protein (Kosen et al., 1981). Therefore, the 14–38 disulfide need not be in equivalent environments in (30–51, 14–38) and native BPTI or refolded BPTI; the disulfide in (30–51, 14–38) should be less flexible than that in the latter two.

CD Difference Spectrum of (30–51, 5–55) minus (30–51, 5–14) plus (30–51, 5–38). The most significant alterations in the peptide backbone conformation likely occur at this step of refolding, (30–51, 5–14) plus (30–51, 5–38) to (30–51, 5–55), as the changes throughout the far-UV CD spectrum are of the greatest magnitude (Figure 6).

The near-UV portion of the CD difference spectrum was resolved into bands assigned to changes in the tyrosine (bands 1–3) and disulfide transitions (band 4) (Figure 6 and Table III). Substantial contributions by tyrosine transitions are indicated by the narrow minimum at 277.5 nm, the shoulder at longer wavelengths, and the sharp onset of ellipticity between 290 and 300 nm. The resolved disulfide envelope is simply due to the increased rotational strength of the 5–55 transition over those of 5–14 and 5–38.

Discussion

Relevance of the Conformational Properties of the Trapped Intermediates of BPTI. The fully reduced form of BPTI is

Table II: Molar Rotational Strengths (cgs Units) and Comparison with the Molar Rotational Strengths from the Resolutions of Some Model Compound Spectra[a]

	(5-30)		(30-51)		(30-51, 5-14) plus (30-51, 5-38)		(30-51, 14-38)	
tentative assignment[b]	$R_0 \times 10^{42}$	R_0/R_m	$R_0 \times 10^{42}$	R_0/R_m	$R_0 \times 10^{42}$	R_0/R_m	$R_0 \times 10^{42}$	R_0/R_m
Tyr[c]								
0-0			−20.9	1.1	−20.9	1.1	−50.1	2.8
0 + 800			−3.6	0.65	−3.6	0.65	−16.5	3.0
0 + 2(800)			−6.2	0.62	−6.2	0.62	−32.3	3.2
				(0.79)[f]		(0.79)[f]		(3.0)[f]
-S-S[d]								
30-51			−716		−716		−755	
				4.1		4.1		4.4
				2.8		2.8		3.0
				0.80		0.80		0.84
5-30	−274							
		1.6						
		1.1						
		0.31						
5-14 plus 5-38					−312			
						1.8		
						1.2		
						0.35		
14-38							−434	
								2.5
								1.7
								0.48
Phe[e]								
0 + 930							−1.44	1.24

[a] $R_0 = 58/[n(1.234 \times 10^{-42})(\theta_0\Delta_0/\lambda_0)]$ cgsu where n is the number of chromophores, i.e., four tyrosines, four phenylalanines, or one disulfide. [b] Tentative assignments are as in Table 1. [c] R_m is an averaged molar rotational strength for the resolved CD transitions of glycyl-L-tyrosylglycine in 0.02 M Tris-HCl, pH 7.0; 6.0 M guanidine hydrochloride and 0.02 M Tris-HCl, pH 7.0; and 80% ethylene glycol and 0.02 M Tris-HCl, pH 7.0 (Holladay & Puett, 1976). The R_m values of the three model spectra are of similar intensities. [d] The R_0/R_m values were calculated by using the molar rotational strength of the longest wavelength resolved band for three different model spectra. From top to bottom: L-cystine, 1.0 KF and 0.02 M Tris-HCl, pH 7.5 (Holladay et al., 1976); L-cystine, pH 5.06-5.9 (Kahn, 1972); hexagonal L-cystine KBr disk (Kahn, 1972). [e] R_m is an averaged molar rotational strength for the resolved CD transitions of glycyl-L-phenylalanylglycine in 0.01 M KCl and 0.002 M sodium phosphate, pH 7.5, and in 80% ethylene glycol, 0.01 M KCl, and 0.002 M sodium phosphate, pH 7.5 (Holladay et al., 1974). The R_m values of the two model compound spectra are of similar intensities. [f] The average R_0/R_m value for the three tyrosine transitions.

Table III: Tentative Assignments, Parameters, and Molar Rotational Strength of the Resolved Gaussian Bands in the CD Difference Spectrum (30-51, 5-55) minus (30-51, 5-14) plus (30-51, 5-38)[a]

tentative assignment	λ_0	Δ_0	θ_0	$R_0 \times 10^{42}$
Tyr				
0-0 (1)[b]	282.5	6.4	−171	−69.3
0 + 800 (2)	276.3	3.4	−76	−16.6
0 + 2(800) (3)	270.9	6.1	−100	−40.1
-S-S- (4)	265.0	19.8	−142	−759.0

[a] Assignments are based on studies by Horwitz et al. (1969, 1970), Kahn (1972), Holladay et al. (1974, 1976), and Holladay & Puett (1976). λ_0 and Δ_0 are expressed in nanometers; θ_0 is expressed in deg·cm²/dmol on a mean residue basis. $R_0 = 58/[n(1.234 \times 10^{-42})(\theta_0\Delta_0/\lambda_0)]$ cgsu where n is the number of chromophores, i.e., four tyrosines, four phenylalanines, or one disulfide. R_0 is the molar rotational strength. [b] Values in parentheses correspond to band numbers in Figure 6.

a highly disorganized polypeptide chain, with a multitude of possible conformations of comparable energies, yet it is able to refold rapidly to a conformation like that of native BPTI. The nonrandom pathway of disulfide bond formation that occurs during this process must arise from the presence of substantial conformational interactions at the various stages of refolding. The nature of these forces should be apparent from the conformational properties of the trapped intermediates, so this spectral study was undertaken.

Partially folded intermediates of small single-domain proteins are known to be thermodynamically unstable (Privalov, 1979), as are the one- and two-disulfide intermediates of BPTI relative to the fully reduced and the fully folded states (Creighton, 1977b). Consequently, they are difficult, but not impossible, to detect experimentally at equilibrium. However, trapping intermediates overcome this difficulty by using their disulfide bonds, even though only the disulfide bonds are trapped, and the rest of the protein is free to equilibrate to many conformations. This is due to the thermodynamic necessity that whatever conformation contributes to the stability of a particular disulfide bond, that conformation should be stabilized to an equal extent by the presence of that disulfide bond, as in the trapped intermediates.

This may be illustrated in the case of the one-disulfide intermediates of BPTI, which accumulate during refolding in proportions determined by their relative stabilities, as they are rapidly equilibrated by intramolecular disulfide rearrangement. Consider a two-state conformational equilibrium between $(30\text{–}51)_F$, with the nonrandom conformation (or family of conformations) responsible for its greater stability, and $(30\text{–}51)_U$, with all other possible unfolded conformations. The value of the equilibrium constant, K_1, for this transition is

$$(30\text{–}51)_U \overset{K_1}{\rightleftharpoons} (30\text{–}51)_F \qquad K_1 = \frac{[(30\text{–}51)_F]}{[(30\text{–}51)_U]}$$

As $(30\text{–}51)_U$, the 30–51 disulfide will have no particular stability and during folding will tend to equilibrate with the 14 other one-disulfide species, designated collectively as I, by intramolecular disulfide rearrangement.

$$(30\text{–}51)_U \overset{K_U}{\rightleftharpoons} I \qquad K_U = \frac{I}{[(30\text{–}51)_U]}$$

To a first approximation, with all disulfides considered to be

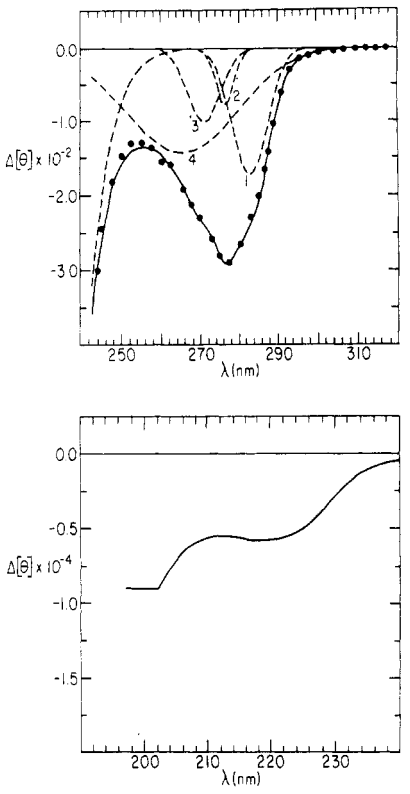


FIGURE 6: Calculated CD difference spectrum: (30–51, 5–55) minus (30–51, 5–14) plus (30–51, 5–38). Values were calculated every 2.5 nm. (Top panel) Near-UV portion of the spectrum: experimental values (●); resolved Gaussian bands (---); sum of the Gaussian bands (—). The band at lowest wavelengths was included only to completely fit the spectrum; no assignment was made for this band. All other bands are numbered and correspond to the band numbers in Table I and in the text. (Bottom panel) Far-UV portion of the spectrum.

equally probable, the value of K_U should be 14. The ratio of intermediate (30–51) with all possible conformations to other one-disulfide species observed during refolding is about 2, i.e.

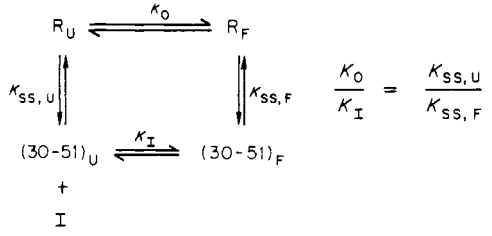
$$2 = \frac{[(30\text{–}51)_F] + [(30\text{–}51)_U]}{I} = \frac{[(30\text{–}51)_U](K_I + 1)}{I} = \frac{K_I + 1}{K_U}$$

Therefore, the value of K_I is approximately 27; that is, intermediate (30–51) should be in the nonrandom conformation about 96% of the time. This value of K_I is an underestimation, since the other one-disulfide intermediates present are not random but consist primarily of (5–30). This intermediate must also have a stabilizing conformation, although it is less stable and less populated.

The same conformational equilibrium may be considered in the reduced protein with no disulfides:

$$R_U \overset{K_0}{\rightleftharpoons} R_F \qquad K_0 = \frac{R_F}{R_U}$$

The value of K_0 may be estimated from its linkage to the one-disulfide equilibria during refolding:



The vertical equilibria involve disulfide bond breakage and formation. The value of $K_{SS,F}$, when the cysteine-30 and -51 residues are in correct proximity for forming a disulfide bond, should be similar to that for the 14–38 disulfide bond in native BPTI, where the folded conformation keeps the cysteine residues in proximity (Creighton, 1975b). The value of $K_{SS,U}$, in the absence of nonrandom conformation, may be taken as that measured for the average of the 15 initial disulfides in 8 M urea (Creighton, 1977c). The value of $K_{SS,U}/K_{SS,F}$ from those measurements is 8.8×10^{-5}. The observed value of K_I at 25 °C is about 2, so the value of K_0 must be about 1.8×10^{-4}. Therefore, the nonrandom conformation responsible for the stability of intermediate (30–51) would be present only 0.02% of the time in fully reduced BPTI, in contrast to >96% of the time in the trapped intermediate (30–51). Such specific conformations would be undetectable from the macroscopic properties of the reduced protein, but comparable values have been measured immunochemically for the acquisition of specific conformations in reduced BPTI (Creighton et al., 1978) and in staphylococcal nuclease fragments (Sachs et al., 1972).

These simple and approximate calculations illustrate how trapping the disulfide bond of a favorable intermediate state should also stabilize the nonrandom conformation responsible for it. The more favorable is the particular intermediate, the greater the stability of the folded conformation in the trapped intermediate. Whereas such nonrandom conformation may be negligible in the initial reduced protein, it should predominate in trapped intermediates such as (30–51) and those following it in the pathway. In contrast, reduced ribonuclease shows little evidence for nonrandom intermediates (Creighton, 1977d, 1979), and the trapped intermediates have little nonrandom conformation (Galat et al., 1981). However, it is essential to block irreversibly all free thiol groups in order to trap the intermediate, and carboxymethyl groups must be introduced in order to separate the trapped intermediates; these groups may alter the conformational properties of the protein. After the intermediates are trapped, slow conformational changes to stable states could conceivably occur, but there has been no indication of this. Such slow conformational transitions do not appear to occur during refolding, as the same pathway is observed over a wide range of disulfide bond formation.

Relevance of the CD Spectra to the Pathway of Unfolding and Refolding of BPTI. Comparison of CD spectra of (30–51) and (5–30) shows that (5–30) is the more disordered polypeptide. Intermediate (30–51) is believed to be slightly more stable and more compact than (5–30) (Creighton, 1974). The aromatic amino acids of (30–51) are also less exposed to solvent than are those of (5–30) (Kosen et al., 1980). Intermediate (30–51) thus has more stabilizing conformational interactions than (5–30), and CD spectroscopy reaffirms this conclusion.

The size of the band assigned to the disulfide transition in the spectrum of (30–51) (Table II) suggests that its disulfide has a single-screw sense and is relatively immobile; there should then be a region of the polypeptide chain around the 30–51 disulfide which acquires a relatively fixed conformation at an early stage of refolding. Whether it is nativelike or whether this region of the polypeptide chain is involved in further productive refolding remains to be seen. It has been proposed by Nagano (1974) that 30–51 might be the first disulfide formed during refolding due to stabilizing interactions between the potentially α-helical and β-sheet regions of the peptide chain, in a manner similar to that of the diffusion–collision theory of Karplus & Weaver (1976). For this reason, we were

Table IV: Secondary Structure Estimates for the BPTI Conformers[a]

BPTI conformer	calcd amount of secondary structure from the model CD spectra of: Greenfield and Fasman, unconstrained			Greenfield and Fasman, constrained[b]			Chen et al., unconstrained			Chen et al., constrained[b]		
	f_α	f_β	f_ρ	f_α	f_β	f_ρ	f_α	f_β	f_ρ	f_α	f_β	f_ρ
native BPTI[c]	21	40	97	33	2	65	30	43	122	35	−19	84
(30–51, 5–55)	21	36	86	30	7	63	30	31	109	34	−15	81
R	11	7	50	4	29	68	10	18	74	10	17	73
(30–51)	11	12	37	2	39	59	13	19	52	12	29	59
(5–30)	12	8	45	4	32	64	12	22	67	12	21	67
(30–51, 5–14) plus (30–51, 5–38)	13	9	38	4	36	60	14	22	57	14	26	60
(30–51, 14–38)	15	17	57	13	24	63	18	26	80	20	10	70

[a] The CD reference spectra of Greenfield & Fasman (1969) and Chen et al. (1974) were used. All estimates are expressed as percentages (f_α, f_β, f_ρ). [b] Under constrained conditions, the sum of the amounts of secondary structure must equal 100. Under unconstrained conditions, the sums can have any value. [c] Crystallographic studies by Huber et al. (1971) and Deisenhofer & Steigemann (1975) showed the presence of an α helix between residues 47–56 (17% of the total secondary structure) and a β sheet between residues 16–36. The two strands of the β sheet are connected by a tight turn at residues 24–27. The percentage of β sheet ranges from 33 to 40%, depending on whether those residues involved in the tight turn are excluded or included.

particularly interested in the far-UV CD spectra of these intermediates. They have been analyzed in terms of secondary structure by the procedures of Greenfield & Fasman (1969) and Chen et al. (1974) (Table IV), even though the spectrum of native BPTI is atypical, giving values of secondary structure that are physically impossible unless the sum of secondary structures is constrained to equal 100%. When constrained, the calculations give values substantially different from the known values (Table IV). Nevertheless, whatever perturbing factors are responsible for the native spectrum might not be present in the partially folded intermediates. However, their spectra give few and inconsistent indications of stable secondary structure (Table IV). If intermediate (30–51) were stabilized by interaction between an α helix containing Cys-51 and the β sheet containing Cys-30, like those in native BPTI, it might be expected to have nearly the secondary structure content of the native protein. Immunochemical measurements also indicate little nativelike structure in (30–51). These techniques may not be adequate to detect this conformation, or it may have been disrupted by the carboxymethyl groups required to block the free cysteine thiol groups. Consequently, the role of the secondary structure in refolding of BPTI still is not known.

To all appearances, (30–51), (30–51, 5–14), and (30–51, 5–38) have very similar conformations. Their CD spectra are very similar in the far-UV regions, and differences in the near-UV regions are explained most simply by the addition of a second disulfide transition for the latter two. The relative solvent exposures of the aromatic residues of these three intermediates are also very similar (Kosen et al., 1980). Unless the similarities in all of the spectral data are all coincidental, formation of either the 5–14 or the 5–38 disulfide appears to involve minimal conformational rearrangements.

In contrast, formation of the 14–38 disulfide in (30–51), to give (30–51, 14–38), produces larger spectral changes and therefore conformational rearrangements. It is possible that (30–51, 14–38) contains more regions of nativelike structure, since its CD spectrum appears to be converging on that of refolded BPTI. Immunochemical studies indicated that (30–51, 14–38) has a probability of 0.25 of being in a nativelike conformation (Creighton et al., 1978). Its tyrosine residues are involved in more tertiary interactions than are those of (30–51), (30–51, 5–14), or (30–51, 5–38), but whether these interactions are nativelike is not known; substantial differences are suggested by the CD spectra. The resolved band, assigned to the 14–38 disulfide transition in the spectrum of (30–51, 14–38), differs in intensity from that in the spectra of native or refolded BPTI (Tables I and II; Kosen et al., 1981); the immediate environments of the two 14–38 disulfides in the different folding states are unlikely to be similar.

Even though (30–51, 14–38) may contain regions of nativelike structure, it cannot form the third disulfide bond readily. This is probably due to the high energy of the transition state, which must go through a distorted version of the nativelike conformation (Creighton, 1980).

Most of the structural elements found in refolded BPTI appear to be formed productively only when (30–51, 5–55) is formed. The CD difference spectrum of Figure 6 shows that the transformations of (30–51, 5–14) to (30–51, 5–55) and (30–51, 5–38) to (30–51, 5–55) involve large changes in the CD spectrum. Indeed, overall, these are the largest optical changes that can be found for any two intermediates adjacent on the refolding pathway.

We previously showed that both the solvent exposure of the aromatic amino acid residues and the lack of affinity for antibodies recognizing native BPTI follow the order R > (5–30) > (30–51) ~ (30–51, 5–14) plus (30–51, 5–38) > (30–51, 14–38) > (30–51, 5–55) ~ refolded BPTI (Creighton et al., 1978; Kosen et al., 1980). The CD spectra qualitatively gave the same order, particularly the trends in increasing negative ellipticity for the tyrosine transitions and in increasing negative ellipticity in the far-UV portions of the spectra. The order given above differs from that which occurs in refolding and in recognition by antibodies against reduced BPTI in that (30–51, 14–38) precedes the other two-disulfide intermediates. So it must be concluded that the tendency to adopt less disordered or possibly nativelike conformations is not the sole determinant of the refolding pathway.

Acknowledgments

We thank Bayer AG for the generous gift of the inhibitor, Drs. L. A. Holladay, D. Puett, and R. G. Hammonds, Jr., of Vanderbilt University, Nashville, TN, for their hospitality and the introduction to the technique of curve resolution, and Dr. J. Brunelle for the preparation of the computer program used in the final resolutions of the CD spectra reported herein.

Registry No. Basic pancreatic trypsin inhibitor, 9087-70-1.

References

Chen, Y.-H., Yang, J. T., & Chau, K. H. (1974) *Biochemistry 13*, 3350–3359.

Creighton, T. E. (1974) *J. Mol. Biol. 87*, 603–624.

Creighton, T. E. (1975a) *J. Mol. Biol. 95*, 167–199.

Creighton, T. E. (1975b) *J. Mol. Biol. 96*, 767–776.
Creighton, T. E. (1977a) *J. Mol. Biol. 113*, 275–293.
Creighton, T. E. (1977b) *J. Mol. Biol. 113*, 295–312.
Creighton, T. E. (1977c) *J. Mol. Biol. 113*, 313–328.
Creighton, T. E. (1977d) *J. Mol. Biol. 113*, 329–341.
Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol. 33*, 231–297.
Creighton, T. E. (1979) *J. Mol. Biol. 129*, 411–431.
Creighton, T. E. (1980) in *Protein Folding* (Jaenicke, R., Ed.) pp 427–446, Elsevier/North-Holland, Amsterdam.
Creighton, T. E., Kalef, E., & Arnon, R. (1978) *J. Mol. Biol. 123*, 129–147.
Deisenhofer, J., & Steigemann, W. (1975) *Acta Crystallogr., Sect. B B31*, 238–250.
Galat, A., Creighton, T. E., Lord, R. C., & Blout, E. R. (1981) *Biochemistry 20*, 594–601.
Greenfield, N., & Fasman, G. D. (1969) *Biochemistry 8*, 4108–4116.
Holladay, L. A., & Puett, D. (1976) *Biopolymers 15*, 43–59.
Holladay, L. A., Hammonds, R. G., Jr., & Puett, D. (1974) *Biochemistry 13*, 1653–1661.
Holladay, L. A., Savage, C. R., Jr., Coehn, S., & Puett, D. (1976) *Biochemistry 15*, 2624–2633.
Horwitz, J., Strickland, E. H., & Billups, C. (1969) *J. Am. Chem. Soc. 91*, 184–190.
Horwitz, J., Strickland, E. H., & Billups, C. (1970) *J. Am. Chem. Soc. 92*, 2119–2129.
Huber, R., Kukla, D., Rülhmann, A., & Steigemann, W. (1971) in *Proceedings of the International Research Conference on Proteinase Inhibitors* (Fritz, H., & Tscheshe, H., Eds.) pp 56–63, Walter de Gruyter, Berlin.
Kahn, P. C. (1972) Ph.D. Dissertation, Columbia University, New York, NY.
Karplus, M., & Weaver, D. L. (1976) *Nature (London) 260*, 404–406.
Kosen, P. A., Creighton, T. E., & Blout, E. R. (1980) *Biochemistry 19*, 4936–4944.
Kosen, P. A., Creighton, T. E., & Blout, E. R. (1981) *Biochemistry 20*, 5744–5754.
Nagano, K. (1974) *J. Mol. Biol. 84*, 337–372.
Privalov, P. L. (1979) *Adv. Protein Chem. 33*, 167–241.
Sachs, D. H., Schechter, A. N., Eastlake, A., & Anfinsen, C. B. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 3790–3794.
Strickland, E. H. (1974) *CRC Crit. Rev. Biochem. 2*, 113–174.
Takagi, T., & Ito, N. (1972) *Biochim. Biophys. Acta 257*, 1–10.

Location of Plasminogen-Binding Sites in Human Fibrin(ogen)†

András Váradi and László Patthy*

ABSTRACT: Affinity chromatography of various fibrinogen and fibrin fragments on Lys-plasminogen–Sepharose was used to localize the plasminogen-binding sites in human fibrin(ogen). The fragments studied in the present investigation were derived from the central (E) and the terminal (D) globular domains of fibrinogen and fibrin. Our results showed that these two different, sequentially nonidentical domains of fibrin(ogen) both carry plasminogen-binding sites. Competitive affinity chromatography of fragment D_1 and fragments derived from it by proteolytic modification of its D γ-chain revealed that this modification causes an 11-fold increase of the association constant of the interaction with Lys-plasminogen–Sepharose. This suggests that the carboxy-terminal region of the D γ-chain is involved in controlling the plasminogen-binding site of the D domain. In contrast with its fragments, intact fibrinogen is not retained by Lys-plasminogen–Sepharose, indicating that the plasminogen-binding sites present in the constituent E and D domains are not fully functional in the parent molecule. It seems possible that the plasminogen-binding sites are present but hidden in fibrinogen and proteolytic dissection of the molecule uncovers these sites in E and D fragments by removing peptides masking the plasminogen-binding regions.

It is generally accepted that the main intravascular function of plasmin is the dissolution of fibrin clots and removal of fibrin deposits from vessel walls. Plasmin eliminates fibrin deposits by cleaving the fibrin molecule at a large number of different peptide bonds with concomitant conversion of the insoluble fibrin polymer into soluble fragments. Despite its trypsin-like sequence specificity, in vivo plasmin normally does not attack other plasma proteins, not even fibrinogen, the soluble precursor of fibrin. A unified molecular model was proposed to explain why plasmin activity is restricted to fibrin (Wiman & Collen, 1978). According to this proposal, plasminogen and plasminogen activator are bound to the fibrin polymer via fibrin-specific binding sites, and plasminogen activation occurs only on the fibrin surface. Plasmin remains bound to fibrin through the same fibrin-binding sites, as well as its active site region, until dissolution of the clot.

The specific interaction of fibrin and plasminogen is thus crucial to the above regulatory mechanism. The subject of the present investigation is the localization of the plasminogen-binding sites in the fibrin(ogen) molecule.

Fibrinogen is an M_r 340 000 protein composed of two sets of three nonidentical chains (Aα, Bβ, γ)$_2$, connected by a network of disulfide bridges [for a review, see Doolittle et al. (1978)]. By electron microscopy, the molecule appears as an elongated trinodular structure (Price et al., 1981; Telford et al., 1980; Fowler et al., 1980); the central nodule contains the amino-terminal globular domain where the two symmetrical halves of the molecule are linked. Thrombin-catalyzed removal of fibrinopeptides A and B unmasks two pairs of polymerization sites in the central nodule (Laudano & Doolittle, 1978; Olexa & Budzynski, 1980). Through these sites, the bivalent central domain binds terminal globular domains of two other

† From the Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, Budapest, Hungary H-1502. *Received November 23, 1982.*

0006-2960/83/0422-2440$01.50/0