Protein Science (1996), 5:1108–1117. Cambridge University Press. Printed in the USA.

Intrinsic tryptophans of CRABPI as probes of structure and folding

PATRICIA L. CLARK,*[1,3] ZHI-PING LIU,*[2] JIANHUA ZHANG,[2,4] AND LILA M. GIERASCH[3]

[1] Molecular Biophysics Graduate Program, and [2] Departments of Pharmacology and Biochemistry, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75235
[3] Department of Chemistry, University of Massachusetts, Amherst, Massachusetts 01003

(RECEIVED September 12, 1995; ACCEPTED March 1, 1996)

Abstract

The native state fluorescence and CD spectra of the predominantly β-sheet cellular retinoic acid-binding protein I (CRABPI) include contributions from its three tryptophan residues and are influenced by the positions of these residues in the three-dimensional structure. Using a combination of spectroscopic approaches and single Trp-mutants of CRABPI, we have deconvoluted these spectra and uncovered several features that have aided in our analysis of the development of structure in the folding pathway of CRABPI. The emission spectrum of native CRABPI is dominated by Trp 7. Trp 109 is fluorescence-silent due to its interaction with the guanidino group of Arg 111. Although the far-UV CD spectrum of CRABPI is largely determined by the protein's secondary structure, aromatic clustering around Trp 87 and the aromatic–charge interaction between Arg 111 and Trp 109 give rise to a characteristic feature in the CD spectrum at 228 nm. The near-UV CD bands of CRABPI arise largely from additive contributions of the three tryptophan residues. Trp 7 and Trp 87 give a negative CD band at 275 nm. The near-UV CD band from Trp 109 is positive and shifted to longer wavelengths (to 302 nm) due to the charge–aromatic interaction between Arg 111 and Trp 109. Our deconvolution of the equilibrium spectra have been used to interpret kinetic folding experiments monitored by stopped-flow fluorescence. These dynamic experiments suggest the early evolution of a well-populated, hydrophobically collapsed intermediate, which undergoes global rearrangement to form the fully folded structure. The results presented here suggest several additional strategies for dissecting the folding pathway of CRABPI.

Keywords: aromatic–charge interaction; fluorescence; near- and far-UV CD; retinoic acid binding protein; tryptophan

Our laboratory has used cellular retinoic acid binding protein I (CRABPI) as a model system to analyze its mechanism of ligand binding and also to explore the folding mechanisms of β-sheet proteins (Zhang et al., 1992; Liu et al., 1994). CRABPI is a 15.5-kDa, cytoplasmic, hydrophobic ligand-binding protein whose fold is dominated by two nearly orthogonal five-stranded β-sheets, which form a clamshell-like structure and surround a central cavity (Fig. 1). Because this protein is made up largely of β-strands, it represents a class of proteins whose folding mechanism is not well understood. Additionally, the presence of the central cavity in CRABPI confounds the commonly held belief that protein folding is driven by formation of a hydrophobic core. In addition to these structural features, CRABPI has several other attributes that make it attractive for biophysical studies of protein folding: (1) It is one member of a large family of proteins that bind small hydrophobic ligands. The sequence homologies among family members vary from as high as 81% to as low as 15%, yet the proteins have remarkably similar three-dimensional structures (for review, see Banaszak et al., 1994). (2) Crystal structures for several family members have been solved, including those for ligand-bound CRABPI and CRABPII (Kleywegt et al., 1994). (3) CRABPI has no disulfide bonds, cofactors, or posttranslational modifications to complicate the folding process.

CRABPI has three tryptophan residues, at positions 7, 87, and 109. As seen from the crystal structure of ligand-bound CRABPI, each of these tryptophans is surrounded by a unique environment. Trp 7 is completely buried, located between the two β-sheets, and is protected from the solvent by the side chains of several nonpolar, nonaromatic residues, as well as by the hydrocarbon chain of Arg 135. Trp 87 is only 70% buried, and has

Reprint requests to: Lila M. Gierasch, Department of Chemistry, University of Massachusetts, Amherst, Massachusetts 01003-4510; e-mail: gierasch@chem.umass.edu.

*The first two authors contributed equally to this paper.

[4] Present address: Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142.

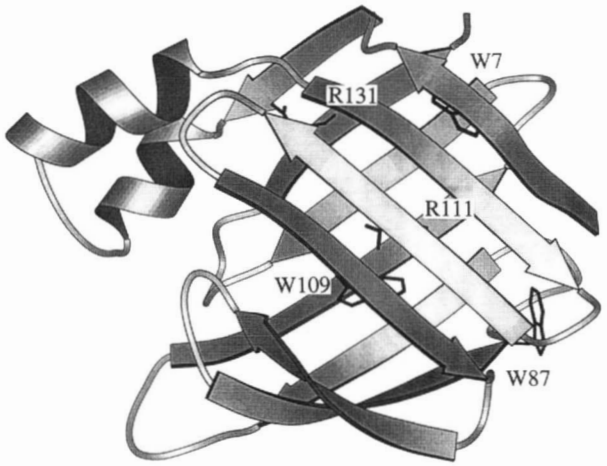


Fig. 1. Ribbon diagram of the crystal structure of native holo-CRABPI (from coordinates from Kleywegt et al., 1994), drawn using the program MOLSCRIPT (Kraulis, 1991). Relative locations of the tryptophan residues and Arg 111 and Arg 131 are indicated.

the peptide backbone and one side of the aromatic ring exposed to the solvent; however, this tryptophan is close to the aromatic rings of two phenylalanine residues and has the potential for a hydrogen bond from its imido proton to the carboxylate group of Asp 48. Trp 109 is almost completely buried and lines one side of the central ligand-binding cavity; it is in close proximity to three phenylalanine residues and the guanidino group of Arg 111.

In earlier work, we used the fluorescence spectrum of CRABPI to monitor ligand-binding specificity and affinity (Zhang et al., 1992). The emission fluorescence of CRABPI decreases due to fluorescence energy transfer from tryptophan to retinoic acid (RA) as the concentration of RA in solution increases (Cogan et al., 1976). We have also used fluorescence spectroscopy together with CD to perform unfolding studies under equilibrium conditions (Liu et al., 1994). Unfolding of CRABPI results in a strong increase of the emission fluorescence (125%) and a concomitant red shift of the maximum emission wavelength (approximately 25 nm). The intensity increase upon exposure of tryptophan residues to the more polar aqueous solvent indicates that the fluorescence of the native protein is quenched. Unfolding of the protein also results in a large change of the CD signals in both the near- and far-UV region.

The significant differences in the fluorescence and CD signals between the folded and unfolded states of CRABPI are potentially very useful probes for studying its folding mechanism. However, the presence of three tryptophan residues in the sequence of CRABPI makes it difficult to associate spectral changes with the evolution of residue-specific interactions. In an earlier mutagenesis study of ligand-binding residues in CRABPI (Zhang et al., 1992), we found that changing Arg 111 to glutamine removes an inflection in the far-UV CD spectrum at 228 nm. The location of the inflection suggested an interaction with an aromatic residue. Based on our model for the three-dimensional structure of CRABPI, we proposed that the inflection arose from a charge–ring interaction between the positively charged side chain of Arg 111 and the π electron cloud of Trp 109. Supporting evidence has been provided by the recently reported crystal structure of holo-CRABPI (Kleywegt et al., 1994), which places the guanidino group of Arg 111 only 3 Å away from the center of the Trp 109 aromatic ring.

Although equilibrium experiments help define the native and unfolded structures and their relative stability, they also provide a basis for kinetic experiments designed to define the evolution of structure during the transition from unfolded to native conformation. Populated intermediates have been observed during the folding of three other β-sheet proteins: interleukin-1β (Varley et al., 1993), intestinal fatty acid binding protein (Ropson et al., 1990) and β-lactoglobulin (Kuwajima et al., 1987); however, little is known concerning the nature of these intermediates. Because β-structure requires the close association of sequentially distant residues, an initial collapse step may be necessary to bring parts of the polypeptide chain into close proximity. For CRABPI and its family members, such a collapse would have to be partially undone in a later step to accommodate the formation of the central binding cavity. Alternatively, more extensive, native-like β-sheets may first arise from small hairpins, and later dock to form the final hydrogen bonding network. Knowledge of the specific interactions responsible for several CRABPI spectral signatures enables us to execute kinetic folding experiments designed to monitor the development of these interactions during the folding of CRABPI and differentiate between these two distinct models for β-structure formation.

In the present study, we have deconvoluted the contributions of Trp 7, Trp 87, and Trp 109 to the near- and far-UV CD and fluorescence spectra of CRABPI through site-directed mutagenesis. Mutant forms of CRABPI were constructed with one or two of the three tryptophan residues replaced with phenylalanine. We chose phenylalanine because of its low fluorescence relative to either tryptophan or tyrosine, and studies of several proteins have found that Trp → Phe substitutions can be tolerated with retention of overall conformation (Rule et al., 1987; Loewenthal et al., 1991; Locke et al., 1992). The fluorescence and CD spectra of the mutant proteins were compared to those of wild-type CRABPI and the R111Q and R131Q mutant proteins obtained previously (Zhang et al., 1992). We have confirmed the presence of a charge–aromatic interaction between Arg 111 and Trp 109, which quenches the fluorescence of Trp 109 in the native state. We find that the major contribution to the fluorescence emission spectrum in the native protein is from Trp 7.

The fluorescence contributions described here have been useful in interpreting the results of variable-emission wavelength stopped-flow fluorescence experiments. These experiments reveal that CRABPI folding occurs in several kinetic phases and is likely dominated by the fast formation of a well-populated intermediate. Together with the spectral deconvolutions, our results suggest that early in the folding process, Trp 7 and/or Trp 109 may experience a more hydrophobic environment than in the native state. In addition, these provocative findings demonstrate a clear need for kinetic experiments on stable, single-tryptophan-containing CRABPI variants, in order to understand the location and extent of structure formation during the folding process.

Results

Protein expression

All proteins, including wild-type CRABPI, were expressed as inclusion bodies in *Escherichia coli* (data not shown), and refolded

from the urea-solubilized inclusion bodies by dilution into refolding buffer (see Liu et al., 1994, for complete protocol). The protein expression levels were equivalent for all proteins (data not shown). However, only the W7F and W87F mutants had folding yields comparable to that of the wild-type protein (approximately 15 mg/L of cell culture). The folding yields of the W109F mutant and the double Trp-mutants were significantly lower than that of wild-type CRABPI (approximately 0.5 mg/L of cell culture). This is probably due to the reduced stability of these mutant proteins. Indeed, the double mutants that include the W109F mutation (W7F/W109F and W87F/W109F) were so unstable that aggregation occurred at protein concentrations necessary for biophysical studies; consequently, results described here only include double-Trp mutant data for W7F/W87F.

Fluorimetric titration

The fluorescence quenching of mutant CRABPI proteins caused by the addition of RA is similar to that of wild-type CRABPI; curves for mutant proteins show quasi-exponential shapes, and the quenching reaches a plateau when the amount of ligand is equal to the amount of the protein (Fig. 2). From the titration curves obtained with RA, it can be assessed that the affinities of the mutant proteins for RA are in the submicromolar range, as described for wild-type CRABPI (Zhang et al., 1992).

We also investigated the binding of RA to the W7F and W87F mutants by monitoring their stabilization against heat denaturation upon RA binding. The free energy of stabilization provided by bound ligand is similar to that of wild-type protein: the melting temperature (T_m) is increased by approximately 17 °C (data not shown). We were unable to monitor the heat-denaturation of the W109F mutant protein due to its aggregation upon heating. However, results from limited chymotrypsin digestion assays indicate that the W109F–RA complex is more resistant to protease than W109F alone, as has been observed for wild-type CRABPI (unpubl. results). Together with the fluorimetric titrations, these results argue that the substitution of tryptophan with phenylalanine at each of the three positions does not markedly change the RA-binding ability or the three-dimensional structure of the protein; the mutant proteins remain functional as RA-binding proteins (see also Fig. 6B, below).

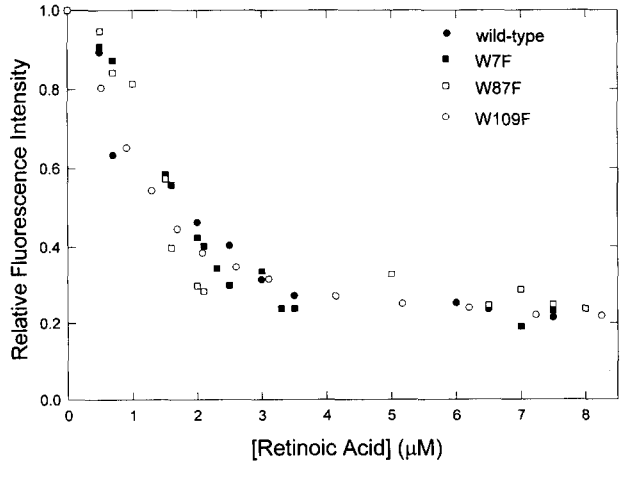


Fig. 2. Fluorimetric titration assay for ●, wild-type CRABPI and ■, mutant W7F; □, W87F; and ○, W109F. Each protein concentration is 3 μM.

Far-UV CD

The far-UV CD spectrum of wild-type CRABPI is typical of a β-sheet protein (Fig. 3A), with a CD minimum at 216 nm (Johnson, 1988, 1990). As described previously (Zhang et al., 1992), the CD spectrum of CRABPI also shows a well-defined inflection at 228 nm. Superpositions of the CD spectra of wild-type and mutant forms of CRABPI are shown in Figure 3. The CD spectrum of the W7F mutant is essentially the same as that of wild-type CRABPI, indicating that Trp 7 does not contribute to the far-UV CD (Fig. 3A). By contrast, removal of either Trp 109 or Trp 87 causes disappearance of the CD band at 228 nm (Fig. 3B). Interestingly, mutation of Arg 111 to glutamine also causes the disappearance of the CD band at 228 nm (Fig. 3B; Zhang et al., 1992), whereas mutation of Arg 131 to glutamine has no effect on the CD spectrum of the protein (Fig. 3A). These results indicate that the aromatic residues Trp 87 and Trp 109 are the determinants of the CD band at 228 nm. Because arginine itself does not contribute to far-UV CD spectra, the effect of Arg 111 on the 228 nm CD band is most likely due to its close proximity to Trp 109 (see the Discussion).

Near-UV CD

The CD spectra of wild-type CRABPI and its mutants in the near-UV region (Fig. 4A) have distinct differences. The wild-type protein displays one positive and one negative CD band at 274 nm and 303 nm, respectively. Mutation of either Trp 7 or Trp 87 to phenylalanine has a similar effect on the near-UV CD of the protein; it decreases the intensity of the negative CD band and increases the intensity of the positive CD band. The λ_{max} of the positive CD band also shifts to a lower wavelength (by 10 nm) in the W7F and W87F mutants. The effect of removal of either Arg 111 or Arg 131 on the near-UV CD spectrum of the protein is also illustrated in Figure 4A. In the R111Q mutant, the positive CD band at 303-nm disappears, and the 274-nm CD band becomes broader. In the R131Q mutant, the CD band at 303 nm remains unchanged and the intensity of the negative CD band at 274 nm increases and becomes broadened.

The difference spectra between wild-type and mutant forms of CRABPI are shown in Figure 4B. A difference spectrum of this sort can be expected to reveal the net contribution of the mutated tryptophan to the near-UV CD of wild-type protein. It should be noted that we assume that the mutation does not affect interactions with neighboring aromatic residues, causing their CD bands to be altered indirectly. Interpretation of the difference spectra for mutants W7F and W87F is straightforward. The contributions of Trp 7 and Trp 87 to the near-UV CD spectrum are similar, with values for $\Delta\epsilon$ consistent with those expected for constrained tryptophans ($-2.2\ M^{-1}\ cm^{-1}$; Strickland, 1974). The assignment of the near-UV contribution of Trp 109 is tempting, although ambiguous. Due to aggregation problems at the higher concentrations required to acquire near-UV CD spectra, we were unable to obtain the near-UV CD spectrum of W109F. However, we speculate that the contribution of Trp 109 to the near-UV signal of CRABPI is likely positive and centered around 295 nm, because mutation of Arg 111 to

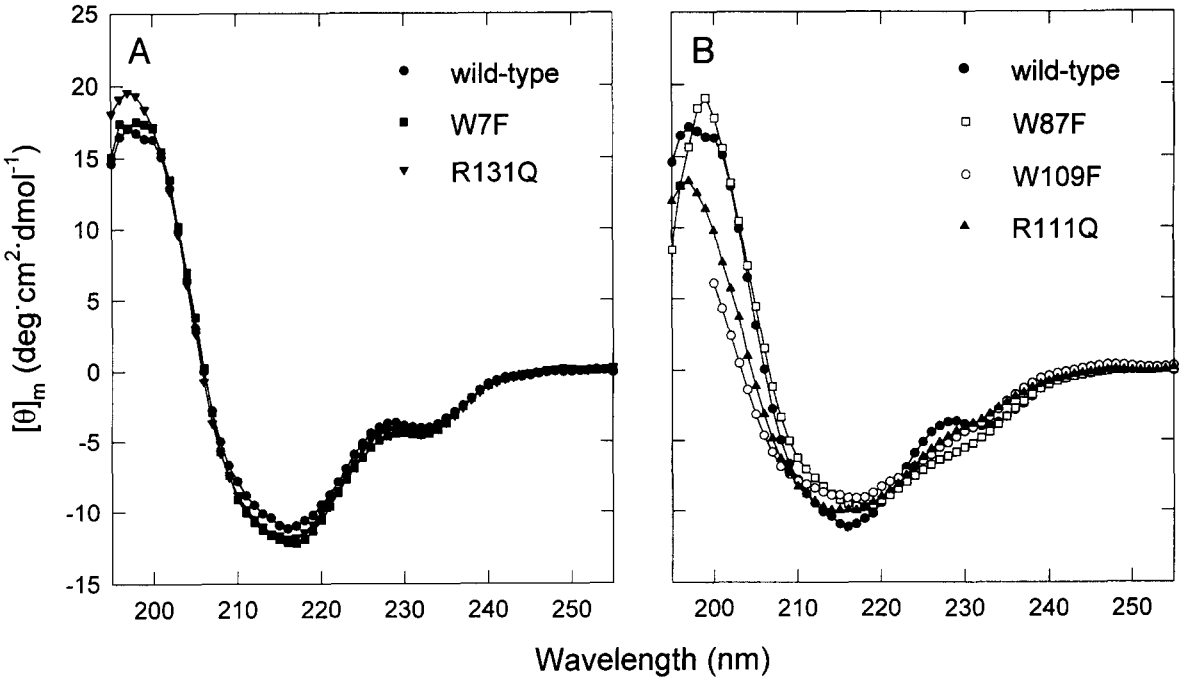


Fig. 3. Far-UV CD spectrum of (**A**) ●, wild-type CRABPI with ■, W7F and ▼, R131Q mutants; and (**B**) ●, wild-type CRABPI with □, W87F; ○, W109F; and ▲, R111Q mutants. Proteins (10 μM) were prepared in 10 mM Tris-HCl, pH 8.0.

glutamine abolishes the positive CD signal of wild-type CRABPI at 295 nm. This effect, like the loss of the far-UV CD band at 228 nm, cannot be from Arg 111 itself, but is likely to arise from its interaction with Trp 109.

Fluorescence emission spectra

The fluorescence emission spectra of native and denatured CRABPI as a function of excitation wavelength are shown in

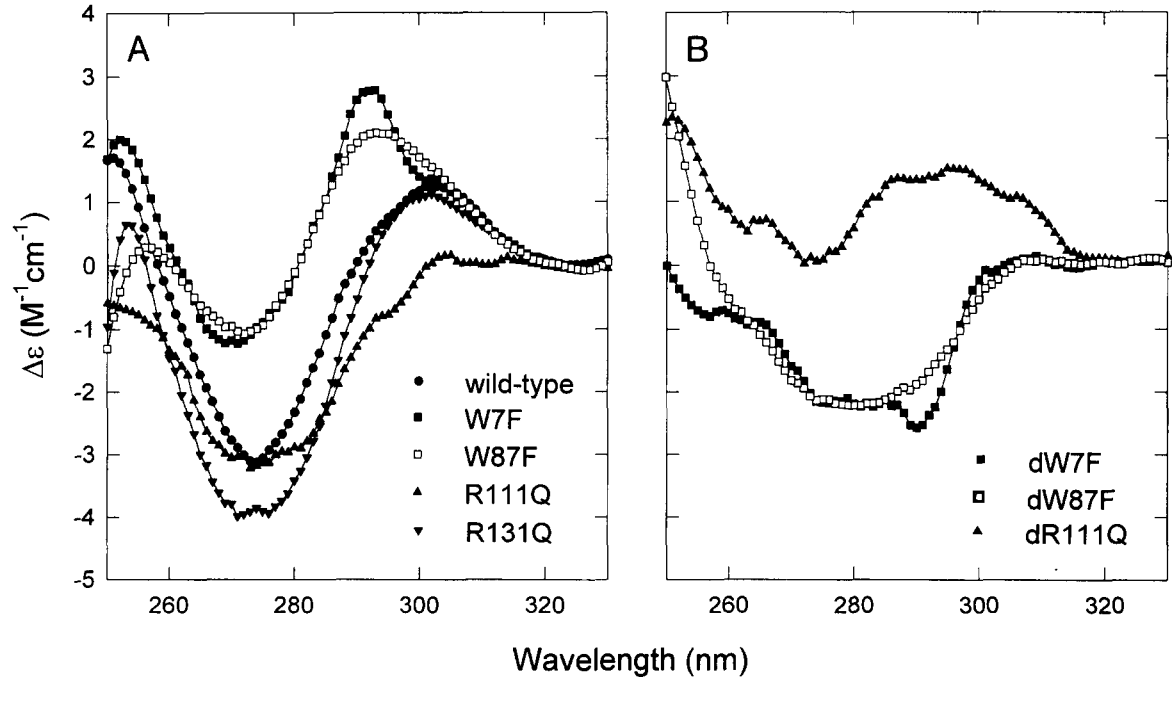


Fig. 4. **A:** Near-UV CD spectrum of ●, wild-type CRABPI; ■, mutant W7F; □, W87F; ▲, R111Q; and ▼, R131Q. **B:** Difference spectra (wild-type minus mutant): ■, dW7F; □, dW87F; ▲, dR111Q in the near-UV. The UV absorption at 280 nm of the protein solutions used here was 0.5, i.e., 24 μM for wild-type and Arg-mutants and 34 μM for single Trp-mutants.

Figure 5A and B, respectively. The spectra are intensity-scaled; unscaled spectra show differences in relative intensities as a function of the position of the excitation wavelength on the tryptophan absorbance spectrum (data not shown). The position of the spectral emission maximum of native CRABPI shows a clear dependency on the excitation wavelength, indicating that the tryptophan residues of CRABPI are experiencing different local environments. In contrast, the maximum emission wavelength of denatured CRABPI remains invariant upon variation of the excitation wavelength (Fig. 5B), indicating that the tryptophan residues in the denatured protein are equivalent and equally exposed to the rapidly relaxing aqueous solvent (Kwon & Churchich, 1994).

The fluorescence emission spectra of native and denatured wild-type and mutant forms of CRABPI are shown in Figure 6A and C, respectively, and their wavelengths of maximum emission (λ_{max}) and relative intensities at λ_{max} are listed in Table 1. In Figure 6A, it can be observed that mutation of Arg 111 to glutamine increases the fluorescence intensity of the protein relative to the wild-type by about 83%. In the R131Q mutant, the fluorescence spectrum is similar to the wild-type except in the region around 318 nm, where the fluorescence intensity is increased by about 20%. The W87F mutant exhibits a blue shift (of 8 nm) that can be associated with a loss of relative intensity (22%) at the longer-wavelength region with respect to the shorter-wavelength region. In contrast to the blue shift of the λ_{max} in the W87F mutant, mutation of Trp 7 to phenylalanine causes a red shift of λ_{max} (of 7 nm); the maximum fluorescence intensity decreases relative to the wild-type protein by approximately 45%. In the W7F/W87F double mutant, where only Trp 109 remains, the maximum fluorescence intensity decreases relative to the wild-type protein by 67%, and the λ_{max} is red shifted by about 9 nm. The sum of the native spectra of the three single Trp-mutant CRABPI variants, divided by two, yields a fluorescence spectrum that is virtually identical to that of native, wild-type CRABPI (Fig. 6B). The similarities between the two spectra indicate that the Trp → Phe mutations do not significantly alter the nature of the environment surrounding the remaining two tryptophan residues, because such perturbations would be likely to create anomalous fluorescence signals. The quality of the overlay reinforces the results from the RA fluorescence titrations and chymotrypsin digestions of the mutant proteins, providing strong evidence that the mutations used here do not greatly affect the fully folded CRABPI native state.[5]

The fluorescence emission spectra of the urea-denatured wild-type CRABPI and single Trp-mutant proteins are shown in Figure 6C. Because the protein concentrations and experimental conditions are identical in Figure 6C and A, intensity differences between the denatured and native proteins can be directly compared. All denatured proteins have λ_{max} at 353 nm, which is similar to the spectrum of free tryptophan in aqueous solution (data not shown). The fluorescence intensities of the wild-type and two arginine mutants are very similar (±7%). Mutation of tryptophan at position 7 or 87 decreases the denatured fluorescence intensity relative to the wild-type protein by about 30%, whereas mutation of Trp 109 to phenylalanine decreases the denatured fluorescence by ca. 50%. In mutant W7F/W87F, mutation of two tryptophan residues at once decreases the fluores-

[5] A similar comparison was attempted for the sum of the native W7F/W87F and W109F fluorescence spectra with the native spectrum of wild-type CRABPI. The overlaid spectra showed significant deviations (data not shown), indicating structural perturbations for W7F/W87F in the environment of W109.

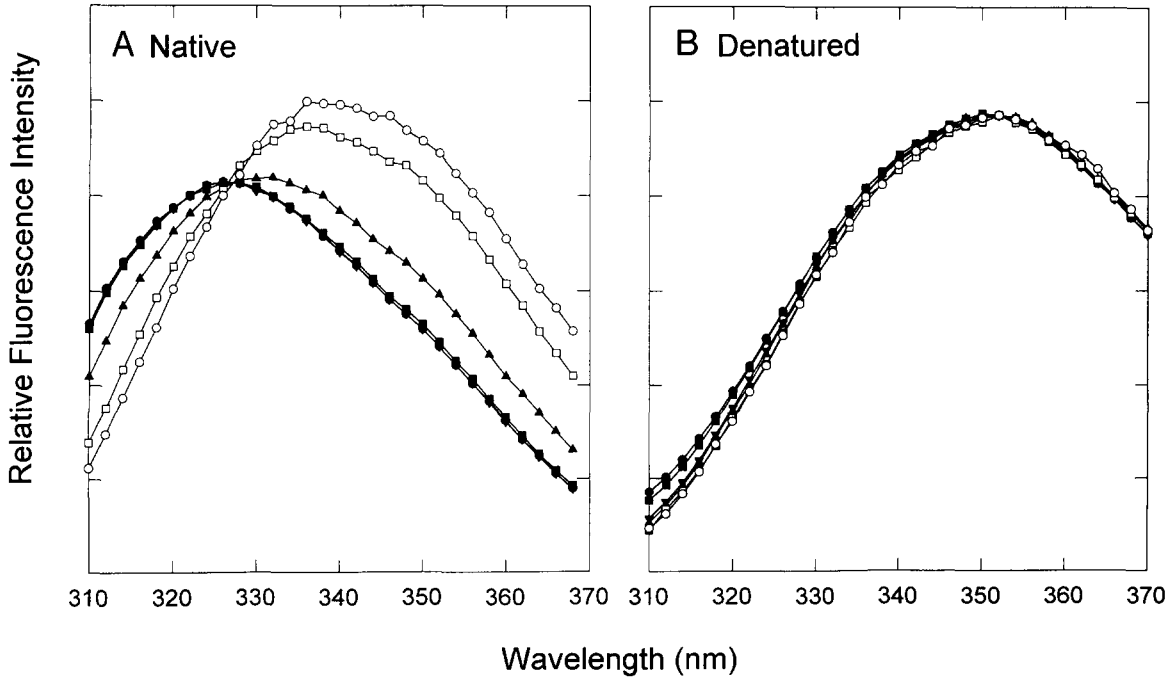


Fig. 5. Fluorescence emission spectra of **(A)** native and **(B)** denatured wild-type CRABPI (5 μM) as a function of excitation wavelength: ●, 280; ■, 285; ▼, 290; ▲, 295; □, 300; and ○, 303 nm. Spectra have been scaled to have similar intensities in order to emphasize the relative positions of the maximum emission wavelength.

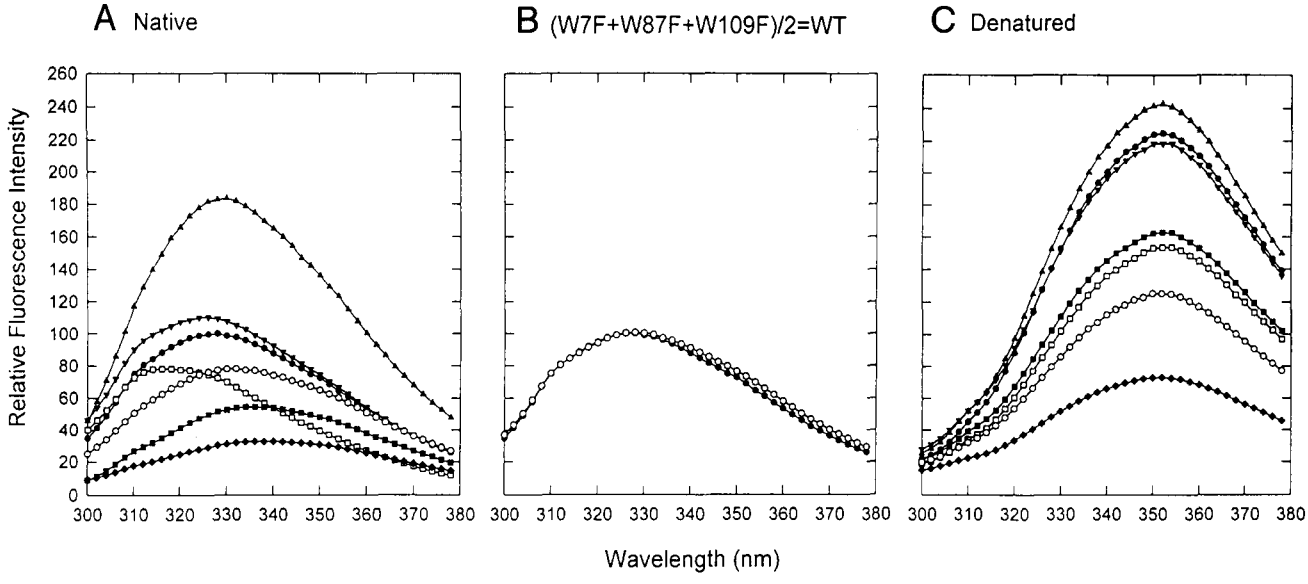


Fig. 6. A: Fluorescence emission spectra of native wild-type and mutant CRABPI proteins: ●, wild-type; ■, mutant W7F; □, W87F; ○, W109F; ◆, W7F/W87F; ▲, R111Q; and ▼, R131Q. All spectra were normalized for protein concentration. The maximum intensity of the wild-type CRABPI spectrum was set arbitrarily to 100. **B:** Comparison of the native, wild-type CRABPI spectrum (●) with a spectrum constructed from the native spectra of W7F, W87F, and W109F (○). The sum of the three mutant protein spectra contains fluorescence from six tryptophan residues; hence the sum is divided by two for the comparison. **C:** Fluorescence emission spectra for denatured wild-type and mutant CRABPI proteins (symbols are the same as in A). Concentrations of the solutions containing native and denatured proteins are identical.

cence by about 68%. The maximal fluorescence intensity of unfolded W7F/W87F is similar to that of an equal amount of free tryptophan in solution (data not shown).

Stopped-flow fluorescence

Stopped-flow fluorescence during wild-type CRABPI folding was measured at seven emission wavelengths; Figure 7A shows the kinetics of the development of native-state fluorescence at three of these wavelengths. Intriguingly, at two of the selected wavelengths (322 [data not shown] and 333 nm), there is a substantial increase in fluorescence (relative to the unfolded spectrum intensity), which occurs within the dead time of the folding experiment (<10 ms). Despite the different behaviors of the initial intensity, each curve fit best to a triple exponential equation, with similar time constants and amplitudes (see the Materials and methods). In Figure 7B, fluorescence spectra after ~10 ms, 100 ms, and 1 s of folding were constructed using the kinetic emission intensities at each wavelength. The spectrum observed at the earliest folding times is significantly different from the unfolded emission spectrum; in particular, the fluorescence intensity is enhanced in the 322–333 nm range. As folding proceeds, this intermediate spectrum gradually decays to create the native fluorescence spectrum.

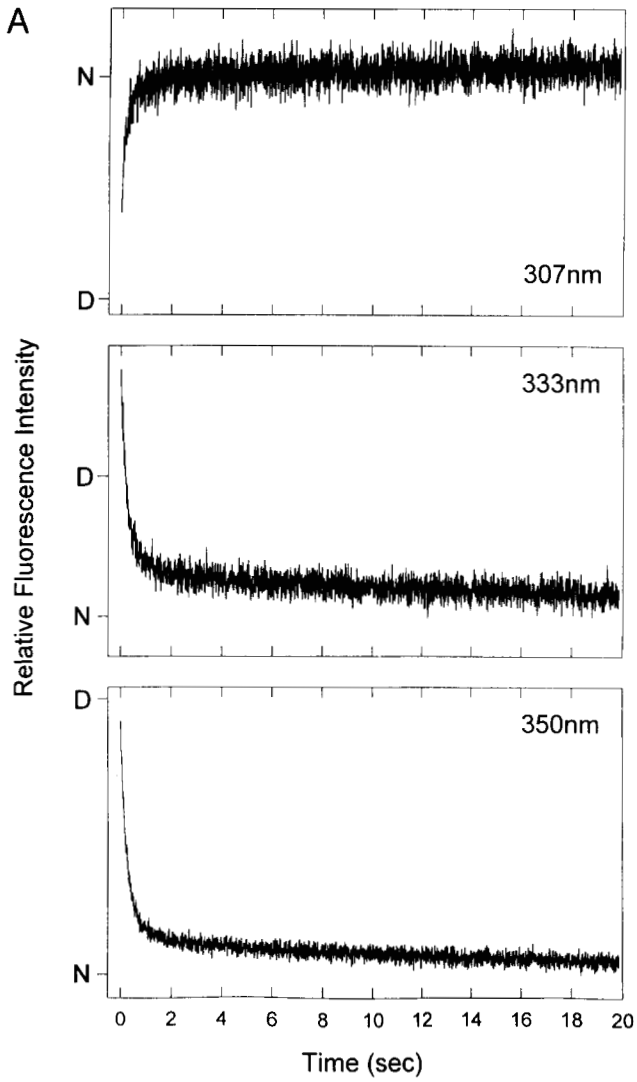
Table 1. *Relative fluorescence intensity and maximum emission wavelength (λ_{max}) of the single Trp-mutants relative to wild-type CRABPI*[a]

		Relative intensity	
	λ_{max} (nm)	Native	Denatured
Wild-type	328	100	225
W7F	335 (+7)	55	163
W87F	316 (−8)	78	154
W109F	332 (+4)	78	122
W7F/W87F	337 (+9)	33	73
R111Q	330 (+2)	183	242
R131Q	326 (−2)	110	218

[a] Values in parentheses indicate the blue (−) or red (+) shift of λ_{max} for the mutants relative to the wild-type value.

Discussion

The family of hydrophobic ligand-binding proteins that includes CRABPI has a wide range of sequence similarity between member proteins, yet shows a striking retention of three-dimensional structure (Banaszak et al., 1994). The results from our mutagenesis studies of CRABPI reinforce the conclusion that the β-barrel structure of these proteins can withstand amino acid substitution at many sites (Zhang et al., 1992; present work). Similar results have been obtained in studies of another family member, rat intestinal fatty acid binding protein (Ropson & Frieden, 1992). From numerous mutagenesis studies, the consensus for proteins as a whole is that amino acid substitutions are best accommodated at surface positions, whereas the residues most crucial to the retention of the native state are usually buried (Goldenberg, 1992). As expected, the three sites in CRABPI, which are not buried to the same extent in the crystal structure, do not equivalently tolerate the Trp → Phe changes. Two of our mutations (W7F and W87F) are easily tolerated by



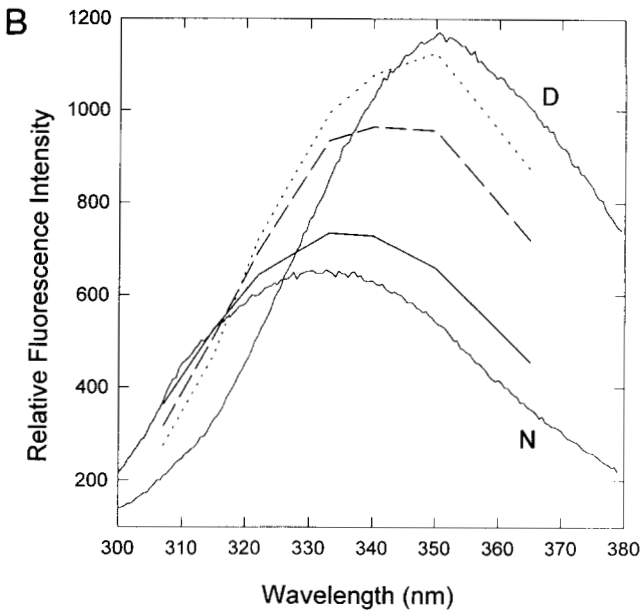


Fig. 7. A: Representative stopped-flow fluorescence traces of wild-type CRABPI during folding in 0.54 M urea with 10 mM Tris-HCl, pH 8.0. Data shown are for emission wavelengths of 307, 333, and 350 nm; data were also acquired at 315, 322, 340, and 365 nm (not shown). Relative fluorescence intensities for native (N) and denatured (D) CRABPI are indicated for each emission wavelength. **B:** Fluorescence spectra of CRABPI after ~10 ms (- - - -), 100 ms (—— ——), and 1 s (————) of folding, constructed from data in A. The equilibrium spectra of native (N) and denatured (D) CRABPI under identical experimental conditions are shown for comparison. Constructed spectra are defined by their line type, while equilibrium spectra are defined by "N" and "D" labels.

the CRABPI native structure, with yields of pure refolded protein comparable to that of the wild-type protein and with thermal stability reduced by less than 10°C (unpubl. results). The lack of change in stability is noteworthy given the location of the mutations in the crystal structure of CRABPI (Kleywegt et al., 1994), where Trp 7 is completely buried and Trp 87 partially so (69%). Comparisons between the stability of these mutant proteins and the location of the mutated residue in the crystal structure are, however, somewhat risky due to possible differences between the crystal structure and the ligand-free protein in solution. The solid-state ligand-bound structure may constrain regions that are flexible under other conditions; such constraints have been demonstrated for solution-state ligand-bound CRABPI (Rizo et al., 1994). The third mutation, W109F, is tolerated with respect to native-state conformation, as judged by far-UV CD characteristics and RA-binding ability, but its refolding yield is very low and the protein aggregates readily upon heating. Similar to Trp 7, the crystal structure shows Trp 109 to be almost completely buried; however, it is possible that Trp 109 is involved in specific contacts or packing that cannot be maintained upon its substitution with phenylalanine. One contact that is seen easily is the proximity of the guanidino group of Arg 111 (see below). In addition, it is the indole nitrogen of Trp 109 that has some exposure to the solvent within the ligand-binding cavity. Substitution with the more hydrophobic side chain of phenylalanine may eliminate polarity that is necessary for stability. Experiments are underway in our laboratory to construct a Trp → Tyr mutation at position 109 that is expected to result in a more stable CRABPI variant. Tyrosine is slightly bulkier than phenylalanine and also more polar, and it is anticipated that this will be a less perturbing substitution, although effects on the interaction with the side chain of Arg 111 are difficult to predict.

Availability of a crystal structure for CRABPI also enables us to look in detail at specific interactions between the tryptophans and neighboring residues in an effort to understand their spectral signatures. Most intriguing among these is the interaction proposed earlier between Trp 109 and Arg 111, based on an R111Q mutation (Zhang et al., 1992). In this paper, we have confirmed the presence of this interaction with the construction of the reciprocal mutation, W109F. The charge-ring interaction between Trp 109 and Arg 111 can be observed spectroscopically

by every method described here: Removing either Arg 111 or Trp 109 alters an inflection in the far-UV CD spectrum at 228 nm, replacement of Arg 111 with glutamine abolishes the positive CD signal of wild-type CRABPI at 295 nm in the near-UV spectrum, and the same replacement also increases the fluorescence intensity of the native protein relative to wild-type CRABPI, suggesting that the fluorescence of Trp 109 is quenched by Arg 111 in the native state. The crystal structure of CRABPI (Kleywegt et al., 1994) provides additional evidence for the origin of these spectral signatures.

Mutation of Trp 87 to phenylalanine also has an effect on the far-UV inflection at 228 nm. Evaluation of the environment surrounding Trp 87 in the crystal structure shows that the aromatic ring of Trp 87 points toward the inside of the β-structure and is close to the aromatic rings of Phe 3 and Phe 50. Far-UV contributions from charge–ring and ring–ring interactions have been observed for many proteins; e.g., bovine pancreatic trypsin inhibitor (Manning & Woody, 1989) and barnase (Vuilleumier et al., 1993). Conversely, Trp 7, which has no solvent exposure, is surrounded by nonpolar, nonaromatic side chains and does not contribute to the far-UV CD inflection. Substitution of Trp 7 with phenylalanine generates the largest decrease in fluorescence intensity (ca. 45% relative to the wild-type protein), whereas mutation of either Trp 87 or Trp 109 generates only a small decrease in fluorescence intensity. Thus, Trp 7 is the major contributor to the fluorescence emission in the wild-type protein. This result is expected because Trp 7 is a completely buried tryptophan and is in the most hydrophobic environment. The common theme underlying all of these observations is the ability to clearly define specific molecular interactions in a dynamic molecule by gathering information from several experimental techniques; providing, of course, that the protein structure has been shown to remain the same under all conditions and throughout the mutagenesis.

The presence of three tryptophans insures that fluorescence is a valuable tool for monitoring CRABPI during folding. Proteins containing multiple tryptophan residues often show a phenomenon called the red-edge excitation effect, in which the tryptophan fluorescence emission spectrum shifts to longer wavelengths as the excitation wavelength is increased toward the red side of the tryptophan absorption (Demchenko, 1986; Wasylewski et al., 1992). This phenomenon can be used to decompose the overlapping fluorescence spectra of individual tryptophan residues located in different parts of the protein molecule (Wasylewski et al., 1992). As shown in Figure 5, the effect is seen for folded (but not unfolded) CRABPI, reinforcing the conclusion that the tryptophan residues have different environments in the native state. In addition, there is a small anomaly seen between the fluorescence spectra of the denatured single Trp-mutant proteins. The measurement of the emission fluorescence of denatured proteins provides a means to check the correctness of the protein concentrations because one could expect that the contribution of the three Trp residues to the fluorescence of CRABPI should be equal assuming the protein is completely denatured. In the denatured state, the fluorescence intensities of wild-type CRABPI and mutant R111Q and R131Q are similar within experimental error ($\pm 10\%$). As expected, mutation of either Trp 7 or Trp 87 decreases the fluorescence intensity by 30%, whereas mutation of two Trp residues (mutant W7F/W87F) yields a 67% decrease. However, the fluorescence intensity decreases ca. 50% in the W109F mutant. This could either be due to an artifact in the determination of the concentration of W109F (possibly due to the tendency of this mutant protein to aggregate) or it could indicate that there is residual structure around Trp 109 in denatured CRABPI. We favor the former possibility because the fluorescence emission spectrum of denatured CRABPI did not show any excitation wavelength dependency (Fig. 5B), and because fluorescence quenching with acrylamide showed the expected linear increase as a function of acrylamide concentration (data not shown), suggesting that the tryptophans in CRABPI are equivalent and exposed to rapid relaxing aqueous solvent. In addition, the fluorescence intensity of the W7F/W87F double mutant was one-third that of the wild-type, which is expected because only Trp 109 remains in the protein.

The stopped-flow fluorescence results described here represent a first glimpse of the dynamics of CRABPI during folding. In the reconstructed spectra, the evolution of a spectrum that corresponds to fluorescence components other than fully denatured and completely folded CRABPI is a strong indication of the presence of an intermediate species. The intermediate is formed very early, with changes in the spectrum occurring within the dead time of the experiment (<10 ms). The most salient characteristic of the intermediate is its enhanced fluorescence intensity in the 322–333-nm region of the emission spectrum, indicating a hydrophobic environment is formed around one or more of the tryptophan residues prior to the evolution of the specific native-state interactions. Although assignment of this effect to any particular tryptophan is, at this point, purely speculative, this part of the native emission spectrum is dominated by fluorescence from Trp 7 (see Fig. 6A), and presents the possibility that, early in the folding process, Trp 7 experiences an environment more hydrophobic than in the native state. Alternatively, it is possible that the increased intensity at 322–333 nm is due to the absence of the quenching effect of Arg 111 on Trp 109 in the early stages of the folding process. The similar time constants and amplitudes observed for each emission wavelength indicate that the three tryptophan residues experience folding-induced environment changes on the same time scales, implying that folding requires entire-chain involvement. The early fluorescence increase and the global nature of the folding events suggest a hydrophobic collapse event may initiate the folding of CRABPI, and argue against folding initiated via the formation of autonomous hairpins. At the moment, however, our interpretation of the dynamic folding changes is entirely dependent on extrapolations from the contributions of each tryptophan to the completely folded fluorescence spectrum. Unambiguous assignment of spectral changes and full understanding of the tryptophan environments during folding requires stable CRABPI variants containing only one wild-type tryptophan; these are currently under construction, incorporating the Trp → Tyr mutation at position 109 and the Arg → Gln mutation at position 131 in attempts to make variants with near-wild-type stability.

Materials and methods

Materials

The R111Q and R131Q mutants were obtained from previous studies (Zhang et al., 1992). Urea (ultrapure) was obtained from ICN Biomedicals, Inc. (Aurora, Ohio). Media used to grow bacteria were purchased from DIFCO laboratories (Detroit, Mich-

igan). *E. coli* BL21(DE3) cells were from Novagen (Madison, Wisconsin). Restriction enzymes were from New England BioLabs (Beverly, Massachusetts). All other reagents, unless specified, were purchased from Sigma (St. Louis, Missouri).

Oligonucleotide-directed mutagenesis

Single Trp-mutant forms of CRABPI (W7F, W87F, and W109F), where one tryptophan residue was replaced with phenylalanine, were generated by the procedure described for the R111Q and R131Q mutants in a previous study (Zhang et al., 1992). The following oligonucleotides were used to direct the mutagenesis (underlines indicate the non-wild-type nucleotides incorporated):

Trp 7 → Phe (W7F): 5′-CGCATCTTGAAGGTACCG-3′;

Trp 87 → Phe (W87F): 5′-TCATTCTCGAACGTGGT-3′;

Trp 109 → Phe (W109F): 5′-TCTCGGGTGAAGTAAGTT-3′.

The W87F/W109F double mutant was generated from the W87F mutant by a second round of mutagenesis with the oligonucleotide used for generating the W109F mutant. The *E. coli* expression vectors of single Trp-mutant forms of CRABPI and double Trp-mutant W87F/W109F (pT7-7/CRABPI/W7F, pT7-7/CRABPI/W87F, pT7-7/CRABPI/W109F, and pT7-7/CRABPI/W87F/W109F) were made by replacing the *Kpn* I/*Eco*R I cDNA fragments of the pT7-7/CRABPI vector with the *Kpn* I/*Eco*R I cDNA fragment of the corresponding mutants (Zhang et al., 1992). Digestion of pT7-7/CRABPI with *Bgl* I generated two fragments, with the larger fragment (2,400 bp) containing the mutation site Trp 7, and the smaller fragment (900 bp) containing the mutation sites Trp 87 and Trp 109. The expression vector pT7-7/CRABPI/W7F/W87F was constructed from the larger fragment of *Bgl* I digestion of pT7-7/CRABPI/W7F and the small fragment of *Bgl* I digestion of pT7-7/CRABPI/W87F. The expression vector pT7-7/CRABPI/W7F/W109F was constructed in the same way from pT7-7/CRABPI/W7F and pT7-7/CRABPI/W109F. The sequences of the mutants were verified by DNA sequencing.

Expression and purification

All proteins were overexpressed in *E. coli* strain BL21(DE3) according to Liu et al. (1994). Wild-type and Trp-mutant forms of CRABPI were purified from cell inclusion bodies according to Liu et al. (1994). The purity of the proteins was assayed by SDS-PAGE as more than 95% pure. Although expression vectors for the other two double Trp-mutants W7F/W109F and W87F/W109F were constructed, biophysical studies with these two mutants were not done in this work. Each protein was dialyzed against 10 mM Tris-HCl, pH 8.0, and stored at 4 °C.

Concentration determination

The extinction coefficients of the wild-type and mutant proteins were determined by the method of Gill and von Hippel (1989). For denatured proteins, the extinction coefficients at 280 nm were calculated from the number of tryptophan, tyrosine, and cysteine residues in the proteins using tabulated molar extinction coefficients (5,690, 1,280, and 128 M^{-1} cm^{-1} for Trp, Tyr, and Cys, respectively). The absorbances of native and denatured CRABPI (and mutants) were found to be the same within experimental error (data not shown). Therefore, the calculated extinction coefficients for denatured proteins were used for native proteins. The calculated extinction coefficients are 21,294 M^{-1} cm^{-1} for wild-type CRABPI, 15,604 M^{-1} cm^{-1} for the single Trp-mutants W7F, W87F, and W109F, and 9,914 M^{-1} cm^{-1} for the double Trp-mutant W7F/W87F. The protein concentrations were determined from the UV absorbance at 280 nm and the appropriate extinction coefficient.

Fluorimetric titration

The binding properties of mutant proteins were measured using a fluorimetric titration method (Cogan et al., 1976). Wild-type and mutant proteins (3 μM in 10 mM Tris-HCl, pH 8) were excited at 280 nm (2-nm slit width) and the emission spectrum was recorded from 300 to 380 nm at 2-nm slit width. RA, prepared in absolute ethanol, was added in 2-μL aliquots with the final added volume not exceeding 1% of the protein solution. The concentration of the stock RA solution was determined spectrophotometrically using a molar absorption coefficient of 45,000 M^{-1} cm^{-1} at 336 nm (Cogan et al., 1976).

Equilibrium fluorescence and CD measurements

Steady-state fluorescence measurements were performed with a PTI M-3 spectrofluorimeter (New Brunswick, New Jersey). A band width of 2 nm was used for both the excitation and emission slits. The fluorescence spectrum was not corrected to compensate for the differences in the lamp intensity across the wavelength range. The CD spectra were recorded on an Aviv 62DS spectropolarimeter (Lakewood, New Jersey). Cells had 1-mm and 1-cm path lengths for recording far-UV (190–250 nm) and near-UV CD (250–320 nm) spectra, respectively. All fluorescence and CD spectra were corrected by subtracting the solvent spectrum obtained under identical conditions. Solutions of wild-type and mutant proteins were prepared in 10 mM Tris-HCl, pH 8.0. Denatured proteins were prepared in 8 M urea and 10 mM Tris-HCl, pH 8.0. Protein concentrations are described in the corresponding figure legends. All measurements were made at 5 °C.

Stopped-flow fluorescence measurements

Kinetic fluorescence traces were acquired using a Bio-Logic SFM-4 stopped-flow apparatus (Claix, France) linked to the excitation and emission monochrometers of a PTI QM-1 fluorimeter via two randomly wound fiber optic cables. Data were collected using the fluorimeter's PMT and processing software, initiated by an external trigger pulse from the motor driver of the stopped-flow apparatus. Band widths were 1.5 nm for excitation slits and 5 nm for emission slits and the data collection rate was 500 pts/s. Under our experimental conditions, the dead time of the stopped-flow apparatus was <10 ms. Data shown are the average of six separate folding shots for each emission wavelength (λ_{ex} = 280 nm). Kinetic traces were fit using the program SigmaPlot (Jandel Scientific). Fits for all data were best to triple, rather than second or quadruple, exponential equations (based on randomness of residuals) and characterized by time constants of approximately 200 ms, 1 s, and 15 s. Corresponding amplitudes were approximately 70%, 15%, and 15%. Points

for the stopped-flow spectra were obtained by either plotting the earliest kinetic data point (~10-ms spectrum) or by averaging the data points obtained on 10 ms of either side of the given spectrum (100 ms and 1 s spectra). Wild-type CRABPI (lyophilized) was unfolded at a concentration of 1 mg/mL in 7 M urea with 10 mM Tris-HCl, pH 8.0, and allowed to equilibrate for >2 h at room temperature. Folding was initiated via a 13-fold dilution of the unfolded protein into 10 mM Tris-HCl, pH 8.0, at 25 °C; final conditions were 0.08 mg/mL CRABPI in 0.54 M urea with 10 mM Tris-HCl, pH 8.0. Equilibrium fluorescence spectra corresponding to fully unfolded and native CRABPI were obtained by diluting the unfolded protein stock into 7 M urea with 10 mM Tris-HCl, pH 8.0, or 10 mM Tris-HCl, pH 8.0, respectively, within the stopped-flow apparatus, and measuring the resulting fluorescence spectrum after approximately 1 min of equilibration. During the time scale of our experiments, there was no significant deterioration of the fluorescence signal from either unfolded or native protein due to lamp exposure.

Acknowledgments

We are grateful to Phil Thomas for the use of the PTI M-3 fluorimeter and for designing the fiber optic adapter for the stopped-flow fluorescence experiments. We thank Josep Rizo for many helpful discussions. This research was supported by NIH grant GM27616. P.L.C. is supported by the NIH Biophysics Predoctoral Training Program (GM08237-07).

References

Banaszak L, Winter N, Xu Z, Bernlohr DA, Cowan S, Jones TA. 1994. Lipid-binding proteins: A family of fatty acid and retinoid transport proteins. *Adv Protein Chem 45*:89–151.

Cogan U, Kopelman M, Mokady S, Shinitzky M. 1976. Binding affinities of retinol and related compounds to retinol binding proteins. *Eur J Biochem 65*:71–78.

Demchenko AP. 1986. *Ultraviolet spectroscopy of proteins*. New York: Springer-Verlag. pp 159–166.

Gill SC, von Hippel PH. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Analyt Biochem 182*:319–326.

Goldenberg DP. 1992. Mutational analysis of protein folding and stability. In: Creighton TE, ed. *Protein folding*. New York: W.H. Freeman and Company. pp 353–355.

Johnson WC Jr. 1988. Secondary structure of proteins through circular dichroism spectroscopy. *Annu Rev Biophys Biophys Chem 17*:145–166.

Johnson WC Jr. 1990. Protein secondary structure and circular dichroism: A practical guide. *Proteins Struct Funct Genet 7*:205–214.

Kleywegt GJ, Bergfors T, Senn H, Motte PL, Gsell B, Shudo K, Jones TA. 1994. Crystal structure of cellular retinoic-acid-binding proteins I and II in complex with all-*trans*-retinoic acid and a synthetic retinoid. *Structure 2*:1241–1258.

Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946–950.

Kuwajima K, Yamaya H, Miwa S, Sugai S, Nagamura T. 1987. Rapid formation of secondary structure framework in protein folding studied by stopped-flow circular dichroism. *FEBS Letters 221*:115–118.

Kwon OH, Churchich JE. 1994. Interaction of 70-kDa heat shock cognate protein with peptide and *myo*-inositol monophosphatase. *J Biol Chem 269*:266–271.

Liu ZP, Rizo J, Gierasch LM. 1994. Equilibrium folding studies of cellular retinoic acid binding protein, a predominantly β-sheet protein. *Biochemistry 33*:134–142.

Locke BC, MacInnis JM, Qian SJ, Gordon JI, Li E, Fleming GR, Yang NC. 1992. Fluorescence studies of rat cellular retinol binding protein-II produced in *E. coli* – An analysis of four tryptophan substitution mutants. *Biochemistry 31*:2376–2383.

Loewenthal R, Sancho J, Fersht AR. 1991. Fluorescence spectrum of barnase – Contributions of three tryptophan residues and a histidine-related pH dependence. *Biochemistry 30*:6775–6779.

Rizo J, Liu ZP, Gierasch LM. 1994. ^{1}H and ^{15}N resonance assignments and secondary structure of cellular retinoic acid-binding protein with and without bound ligand. *J Biomol NMR 4*:741–760.

Ropson IJ, Gordon JI, Frieden C. 1990. Folding of a predominantly β-structure protein: Rat intestinal fatty acid binding protein. *Biochemistry 29*:9591–9599.

Ropson IJ, Frieden C. 1992. Dynamic NMR spectral analysis and protein folding: Identification of a highly populated folding intermediate of rat intestinal fatty acid-binding protein by ^{19}F NMR. *Proc Natl Acad Sci USA 89*:7222–7226.

Rule GS, Pratt EA, Simplaceanu V, Ho C. 1987. Nuclear magnetic resonance and molecular genetic studies of the membrane-bound D-lactate dehydrogenase of *Escherichia coli*. *Biochemistry 26*:549–556.

Strickland EH. 1974. Aromatic contributions to circular dichroism spectra of proteins. *CRC Crit Rev Biochem 2*:113–175.

Varley P, Gronenborn AM, Christensen H, Wingfield PT, Pain RH, Clore GM. 1993. Kinetics of folding of the all β-sheet protein interleukin-1-β. *Science 260*:1110–1113.

Vuilleumier S, Sancho J, Loewenthal R, Fersht AR. 1993. Circular dichroism studies of barnase and its mutants: Characterization of the contribution of aromatic side chains. *Biochemistry 32*:10303–10313.

Wasylewski Z, Koloczek H, Wasniowska A, Slizowska K. 1992. Red-edge excitation fluorescence measurements of several two-tryptophan-containing proteins. *Eur J Biochem 206*:235–242.

Zhang J, Liu ZP, Jones TA, Gierasch LM, Sambrook JF. 1992. Mutating the charged residues in the binding pocket of cellular retinoic acid-binding protein simultaneously reduces its binding affinity to retinoic acid and increases its thermostability. *Proteins Struct Funct Genet 13*:87–99.