ELSEVIER

Biophysical Chemistry 100 (2003) 421–436

Biophysical Chemistry

www.elsevier.com/locate/bpc

Native structural propensity in cellular retinoic acid-binding protein I 64–88: the role of locally encoded structure in the folding of a β-barrel protein☆

Kenneth S. Rotondi[a], Linda F. Rotondi[b], Lila M. Gierasch[a,b,*]

[a]*Department of Chemistry, University of Massachusetts–Amherst, Amherst, MA 01003, USA*
[b]*Department of and Biochemistry & Molecular Biology, University of Massachusetts–Amherst, Amherst, MA 01003, USA*

Received 11 April 2002; accepted 6 June 2002

Abstract

A central question in protein folding is the relative importance of locally encoded structure and cooperative interactions among residues distant in sequence. We have been exploring this question in a predominantly β-sheet protein, since β-structure formation clearly relies on both local and global sequence information. We present evidence that a 24-residue peptide corresponding to two linked hairpins of cellular retinoic acid-binding protein I (CRABP I) adopts significant native structure in aqueous solution. Prior work from our laboratory showed that the two turns contained in this fragment (turns III and IV) had the highest tendency of any of the eight turns in this anti-parallel β-barrel to fold into native turns. In addition, the primary sequence of these two turns is well conserved throughout the structural family to which CRABP I belongs, and residues in the turns and their associated hairpins participate in a network of conserved long-range interactions. We propose that the strong local-sequence biases within the chain segment comprising turns III and IV favor longer-range interactions that are crucial to the folding and native-state stability of CRABP I, and may play a similar role in related intracellular lipid-binding proteins (iLBPs).

Keywords: Protein folding; Peptide fragment; β-Sheet; β-Turn; Diffusion–collision; Conserved interaction; Intracellular lipid-binding proteins

☆ We dedicate this manuscript to John T. Edsall in memory of his long career as a scientific statesman and eminent biophysicist. We have been fortunate to pursue our scientific careers during the time of his influence on biophysics and biophysical chemistry. We have benefited from his wisdom, his rigor, his broad knowledge and his integrity. The present view of protein structure and folding has been molded by his fundamental contributions, his critical reviews and his thoughtful editorial guidance. Beyond this he was a scientist with a social conscience; we aspire to follow his example. He will be greatly missed.

*Corresponding author. Tel.: +1-413-545-6094; fax: +1-413-5451289.
E-mail address: gierasch@biochem.umass.edu (L.M. Gierasch).

0301-4622/03/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0301-4622(02)00296-X

1. Introduction

The seminal work of Anfinsen in the 1960s demonstrated that the information necessary to guide the acquisition of the native state of a protein resides in its primary sequence [1]. The last four decades of research on protein folding has focused on understanding how this information is expressed. A central question in these investigations has been the relative role of local and global forces in directing the folding process. Two broad models have been developed as potential mechanisms for the folding reaction. The hydrophobic collapse model posits that global information in the pattern of hydrophilic and hydrophobic residues directs acquisition of the native state through a collapse that sequesters the hydrophobic residues from solvent [2–4]. In contrast, the framework model suggests that locally defined, native-like secondary and super-secondary structure forms early and assembles to form the native state [5,6]. Two variants of the framework model of folding are the nucleation–condensation [7,8] and the diffusion–collision [9,10] models. In the former, a region of locally defined structure provides a nucleus upon which the rest of the chain can organize, while in the latter, two or more such regions of locally defined structure are separately encoded, and the productive collision of these regions favors the folding reaction. In framework models, the acquisition and retention of native regions of structure lead to a build-up of native structure throughout the folding reaction, and thus these models are hierarchical. Central to framework models is the ability of primary sequence to encode native conformation independently of the other regions of the folding chain. By contrast, the hydrophobic collapse model does not require linear sequence to favor native structure when isolated from the entire protein chain.

The framework model has been tested in many studies of protein fragments. Considerable evidence for the behavior of short peptides shows that the α-helix can be locally defined [11–16] and that the stability of helices in protein fragments correlates with their roles in the folding of helical proteins [17–20]. In addition, these locally defined helices have been shown to function in both nucleation–condensation [8] and diffusion–collision folding processes [10]. By contrast, there have been fewer studies of peptide fragments from primarily β-sheet proteins. The topology of β-structure requires that residues far removed in primary sequence make specific contacts. Nonetheless, anti-parallel up–down β-structure relies upon the presence of turns between adjacent strands. Turns constitute a region where interactions in local sequence can have significant effects on longer-range interactions.

In recent years there has been an increasing number of studies showing native conformational bias in hairpin fragments of β-sheet proteins [21–24]. The distal loop of the spectrin SH3 domain showed a tendency to fold into a native-like conformation in a fragment [25], and this loop also appears to be well formed in the folding transition state [26]. Similarly, a 16-residue fragment corresponding to the N-terminal hairpin of ubiquitin populates a native-like conformation [23], and hydrogen exchange results show that this same region shows early protection [27]. Recent work suggests that the ability of this region of ubiquitin to fold relies on the presence of a five-residue turn with a glycine bulge [28]. Removing the glycine in the fragment disrupted the turn conformation, and mutating it in the protein slowed folding by a factor of two, arguing that it plays a critical role. In the protein G B1 domain, the second of two turns persists in a fragment [29] and is formed in the transition state [30]. All of these examples underline the potential roles played by turns and hairpins in folding and demonstrate that their behavior in excised fragments can be used to examine their roles in folding. However, all of these examples come from small (<100 residues) proteins, none of which contain regular β-structure.

The question remains to what extent local sequence biases in turns are important in the folding of larger β-proteins. In β-barrel proteins, is this bias towards native conformations a feature of all of the turns, and if not, do the propensities of turns to adopt native conformations determine a hierarchical process of folding? A better understanding of these local sequence-encoded steps in β-sheet folding may help to provide general prin-

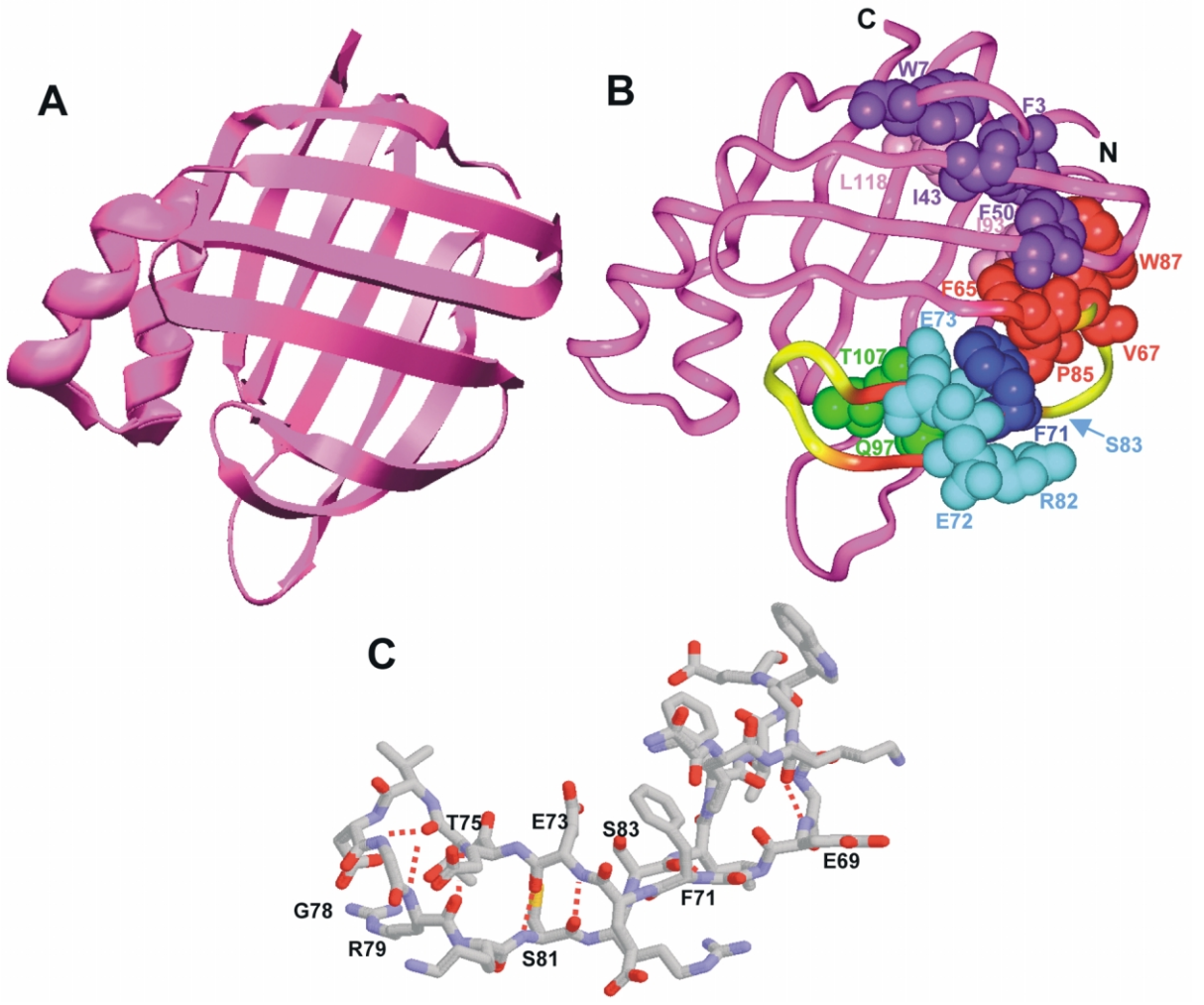

Fig. 1. (A) Cartoon depiction of CRABP I [33] showing the overall architecture. (B) Ribbon diagram of CRABP I with residues involved in the conserved hydrophobic and polar clusters (Gunasekaran et al., unpublished) shown in spacefill. Shown in red are residues in the network of hydrophobic residues that are on the 64–88 segment. Other members of the hydrophobic network are in dark purple if they are on the N-terminal sheet, or in light purple if they are on the C-terminal sheet. In light blue are residues in the polar cluster that reside on the 64–88 segment; other residues in the polar cluster are shown in green. Phe71 (depicted in dark blue) forms conserved hydrophobic interactions with F65 and P85 in the hydrophobic cluster, as well as a conserved hydrophobic–polar interaction with S83 in the polar cluster. As such, it bridges the two clusters. (C) Stick representation of CRABP I 64–88 fragment. Intra-fragment hydrogen bonds available in the native conformation are shown, and the hydrogen bond donors are labeled.

ciples applicable to the folding mechanisms of this relatively under-studied yet highly important structural class [31]. Aberrant folding of β-sheets may have deleterious consequences, and indeed may be causative in many diseases [32], and thus enhanced understanding may help in the design of therapeutic interventions.

To examine these issues, we have focused our attention on the folding of a β-barrel protein with a simple anti-parallel, up–down topology—cellular retinoic-acid-binding protein I (CRABP I, Fig. 1). CRABP I is a member of the large family of intracellular lipid-binding proteins (iLBPs), allowing us to use comparative sequence and structure

data to extend our findings [33]. The physiological role of CRABP I is the transport and sequestration of the sparingly soluble ligand retinoic acid. As occurs for all iLBPs, the hydrophobic ligand binds in a central cavity formed by a 10-stranded β-barrel. The folding of CRABP I has been well characterized [34–36]. The earliest kinetic event (~300 μs) involves hydrophobic collapse accompanied by acquisition of considerable secondary structure ([36]; Habink and Gierasch, unpublished). In a second kinetic phase (100–200 ms), the population of folding species largely attains its native topology, as indicated by docking of the N- and C-termini and formation of the binding pocket [36]. Intriguingly, the 100–200-ms intermediate population is devoid of stable hydrogen bonds in the β-barrel. The bulk of the folding population reaches the native state, with all tertiary interactions and a stable hydrogen-bonding network in the sheets occurring an order of magnitude later in a 1-s time phase. A small population (~15%) of the folding protein is rate-limited by *cis–trans* isomerization of the Leu84–Pro85 bond, leading to an observable 15–20-s time phase [37].

The hypothesis that the local sequence biases early folding events was tested through a fragment study focusing on the β-turns of CRABP I. Seven short (six hexa- and one hepta-) peptides corresponding to the β-turns in CRABP I were subject to biophysical characterization (Rotondi and Gierasch, in preparation). Results from nuclear magnetic resonance (NMR) and circular dichroism (CD) indicate that, while the balance of the turns are primarily random coil in nature, turns III and IV possess considerable bias toward the native conformation. Indicators include non-random CD spectra, native-like deviations from random-coil chemical shift, native hydrogen bonding monitored by amide proton temperature shifts, and native-like cross-strand nuclear Overhauser enhancements (nOes). These two turns occur at intriguing positions in the CRABP I structure: turn III joins a hairpin devoid of cross-strand hydrogen bonding, while the turn IV hairpin forms the transition between the N- and C-terminal sheets of the β-barrel. The fact that these turns form linked hairpins in CRABP I suggested that the region encompassing them and the intervening strand might play a significant role in the folding of the intact protein. In fact, the turn III/IV-linked hairpin and the helix–turn–helix sub-domain, which we earlier showed to have a strong tendency to adopt native structure as an isolated fragment [38], constitute the most likely regions to behave in a framework-like way. Thus, we seek to explore the extent to which local sequence-encoded structure formation by the turn III/IV-linked hairpin plays a role in productive folding of CRABP I.

An additional motivation for the present studies of the turn III/IV-linked hairpin region of CRABP I derives from observations based on sequence comparisons in the iLBP family. We have identified interacting residue pairs in the CRABP I structure that have been conserved in the iLBP family as hydrophobic–hydrophobic or polar–polar, based on the premise that these pairwise interactions are likely to contribute favorably to stability and folding (Gunasekaran et al., in preparation). Interactions between amino acids within five residues in primary sequence were considered local, and those beyond this window, global. The analysis of local conserved interactions found only three in the β-domain of CRABP I; remarkably, two were associated with turn III and one with turn IV. In addition, the global analysis identified a hydrophobic network primarily involving residues on the N-terminal β-sheet, as well as a smaller, polar cluster in the β-domain of CRABP I (Fig. 1b). Five of the seven residues involved in conserved polar interactions and six of 19 residues involved in conserved hydrophobic interactions reside in the turn III and IV hairpins, suggesting a critical role for this portion of the molecule in organization of this interaction network.

We hypothesize that early steps in the folding of CRABP I are critically dependent on the tendency of turns III and IV to take up native conformation, which in turn favors the organization of the subset of residues into the conserved network of long-range interactions. The resultant early organization of this region of the molecule provides an interaction surface upon which additional longer-range conserved interactions can form (Fig. 1). To refine and test this hypothesis, we examined the behavior of the turn III/IV-linked hairpins in a 25-residue fragment cor-

responding to residues 64–88 in CRABP I (^{64}N*FKVGEG*FEE*ETVDGRK*SRSLPTWE^{88}, turn III and IV regions previously studied as peptides are in italics) (Fig. 1b,c). Biophysical characterization of CRABP I 64–88 supports the notion that this region of CRABP I populates a native-like conformation in the unfolded state and influences the early folding events of the intact protein. The resultant fragment is monomeric in nature. While there is no indication of stable tertiary structure, there are numerous data showing that the two turns and their connecting strand populate native conformation. Supporting this native bias are the chemical shift deviations from random coil, nonsequential nOes, and amide shift with temperature NMR data in the turn III and IV regions of the large fragment. Results of the conformational analysis of this fragment and the implications for the folding of the intact protein are presented in this paper.

2. Materials and methods

2.1. Peptide synthesis and characterization

The CRABP I 64–88 fragment was synthesized by solid-phase methods [39] utilizing 9-fluorenylmethyloxycarbonyl (Fmoc) protection and *N*-[(dimethylamino)-1*H*-1,2,3,-triazolo[4,5-b]-pyridin-1-ylmethylene]-*N*-methylmethanaminium hexyfluorophosphate *N*-oxide (HATU) as a coupling agent. The synthesis was carried out on a PE Biosystems (Framingham, MA) Pioneer peptide synthesis system with PE Biosystems protocols. The N- and C-termini were left unprotected to enhance solubility. The crude peptide was purified using reverse-phase high-performance liquid chromatography (RP-HPLC). Support resin and HATU were purchased from PE Biosystems (Framingham, MA). Fmoc-amino acids were purchased from Peptides International (Louisville, KY). Vydac preparatory and analytical HPLC columns were purchased through the NEST group (Southboro, MA). HPLC-grade dimethylformamide, dichloromethane, acetonitrile and methanol were purchased from EM-Scientific (St. Louis, MO). Phenol, ethyl ether and triisopropyl silane were purchased from VWR (West Chester, PA) and Sigma Scientific (St. Louis, MO). Spectrophotometric-grade trifluoroacetic acid was purchased from Aldrich chemicals (Milwaukee, WI).

The identity of the pure peptide was confirmed by NMR spectroscopy, amino acid analysis (Cornell Biotechnology Analytical Chemistry Facility, Ithaca, NY) and mass spectrometry (Mass Spectrometry Facility, University of Massachusetts, Amherst). The purity of the peptide was assessed by analytical RP-HPLC. The extinction coefficient at 280 nm was calculated by the method of Gill and von Hippel [40]; subsequently, peptide concentration was determined by A_{280} or quantitative NMR spectroscopy [41].

2.2. CD spectroscopy

CD spectra were collected at 5 °C on a JASCO J-715 spectropolarimeter and processed with Jasco J-700 software version 1.50. A JASCO PTC-348WI Peltier piezo electric controller maintained temperature control. CD samples were in 10 mM phosphate, pH 5.3, at peptide concentrations of 24 μM and 4.4 mM. The pathlength of the CD cuvettes (Hellma, Germany) ranged from 0.01 to 1 cm. A 2-cm jacketed cell from Uvonic Instruments Inc was used with a circulating water bath for temperature control in some of the low concentration analyses. CD spectra reported are the average of five scans, using a 1-nm bandwidth at a scan rate of 10 nm/min, and were corrected by subtracting a solvent spectrum acquired under identical conditions. Spectra were recorded above 200 nm to maintain a photomultiplier voltage of <600 V.

2.3. NMR spectroscopy

NMR analysis was performed on a Bruker 600-MHz Avance NMR spectrometer using a TXI SB 5-mm probe with z gradients. Temperature was maintained with a Bruker BCU 05 temperature control unit (calibrated against 4% methanol in methanol-d_4 with a trace of hydrochloric acid). Samples were initially analyzed at concentrations of 100 μM and 1 mM to check for aggregation. The samples were in 10 mM potassium phosphate, pH 5.3, 10% D_2O, with 0.11 mM 2,2-dimethyl-2-

silapentane-5-sulfonic acid as a chemical shift reference, with a final volume of 650 μl. The pH was adjusted with 1 M potassium hydroxide and 1 M hydrochloric acid and read directly from a Fisher pH meter with no correction for the deuterium isotope effect. 1D spectra were collected with 14 000 points and a spectral width of 7183.9 Hz, or approximately 12 ppm. Spectra were zero-filled to 32768 points prior to Fourier transformation, giving an effective spectral resolution of 0.2 Hz/point or 0.0004 ppm/point. Water suppression in all experiments was achieved using the excitation sculpting modification of the WATERGATE gradient selective excitation pulse scheme [42].

Total correlation 2D spectroscopy (TOCSY) [43] was used for identification of peptide spin systems, measurement of amide-proton chemical shift change with temperature (temperature coefficient, $\Delta\delta/\Delta T$) and determination of ^{1}Hα deviation from random coil. A low temperature coefficient indicates that the amide proton is shielded from bulk solvent, whether through participation in a hydrogen bond or through steric occlusion [44]. As a result, the temperature coefficient is a powerful tool for identifying the presence of intramolecular hydrogen bonds in fragments. The shift from random coil for the α-proton of a residue provides information on the secondary structure populated in a site-specific manner. An upfield or negative shift from the random coil value is indicative of participation in a helical or turn conformation, while a downfield or positive shift indicates extended conformation [45,46]. This site-specific secondary structure information is extremely useful in determining the overall conformational tendencies of a protein fragment.

TOCSY mixing was effected using an MLEV-17 sequence [47] with a mixing time of 78 ms, while phase sensitivity was accomplished using States-TPPI phase cycling [47]. The chemical shift data reported were recorded at 3 °C. Amide temperature-shift data were collected at 3, 5, 10, 15 and 25 °C. TOCSY spectra were acquired with a spectral width of 6127.4 Hz or approximately 10 ppm in both dimensions, with 2048 complex data points in $t2$ and 1024 experiments of eight scans each in $t1$. A Gaussian window function was used with a line broadening of −2 Hz and a Gaussian broadening of 0.1 and 0.05 in t2 and t1, respectively. The processed matrices were 1024×1024 points with an effective resolution of 6 Hz/point or 0.01 ppm/point.

Proton nuclear Overhauser effect spectroscopy (^{1}H NOESY) provides a powerful tool for determining inter-proton distances in biomolecules [48]. The intensity of the nuclear Overhauser effect (nOe) is related to interproton distance by an inverse sixth power, and the maximum distance observable is approximately 5 Å. NOESY spectra were collected at 3 °C for mixing times of 150 and 300 ms, with a spectral width of 6127.4 Hz or approximately 10 ppm in both dimensions, 2048 complex points in $t2$ and 1024 experiments of 32 scans each in $t1$. Phase sensitivity was achieved using States-TPPI phase cycling [47]. Gaussian window functions were applied with line broadening of −2 Hz and a Gaussian broadening of 0.09 and 0.04 in $t2$ and $t1$, respectively. The data in $t1$ were zero-filled to 2048 points prior to Fourier transformation, resulting in matrices that were 2048×2048 points with an effective resolution of 3 Hz/point or 0.005 ppm/point. Sequential assignment of the fragments was accomplished using the NOESY sequential α-proton to amide proton [$d_{\alpha N}(i,i+1)$] nOes; in combination with TOCSY data, nearly complete assignment of the peptide was accomplished (Table 1).

2.4. Conservation of secondary structure

Secondary structure conservation was determined between CRABP I and 53 sequences of other iLBP family members with sequences greater than 30% homology (to increase reliability of sequence alignment [49,50]). Sequence alignment was accomplished using the CLUSTALW program [51]. The conservation analysis is based upon the primary sequence and secondary structure of a reference protein, in this case CRABP I. The residue of the reference protein is used to calculate the conservation of each position in the sequence. The positional conservation score is a simple percentage of the occurrence of the reference residue in the available sequences at that position. The crystal structure of CRABP I was used to

Table 1
Chemical shifts of the protons in CRABP I 64–88

Residue	Chemical shift (ppm)					
	NH	αH	βH	Other		
F65	8.73	4.66	3.00, 3.09	2,6: 7.33	3,5: 7.30	4: 7.28
K66	8.4	4.28	1.71	γ1.32	δ1.61, ε3.00	
V67	8.32	3.98	2.02	γ0.92		
G68	8.71	3.87, 4.04				
E69	8.3	4.32	1.9, 2.06	γ2.23		
G70	8.55	3.78, 3.86				
F71	8.14	4.62	2.97,3.10	2,6: 7.14	3,5: 7.48	4: 7.22
E72	8.52	4.26	1.88, 1.99	γ2.23		
E73	8.51	4.24	1.94, 2.04	γ2.30		
E74	8.59	4.37	1.96, 2.04	γ2.31		
T75	8.36	4.43	4.27	γ1.19		
V76	8.45	4.04	2.08	γ0.94		
D77	8.52	4.55	2.70			
G78	8.41	3.90, 3.95				
R79	8.14	4.28	1.80, 1.86	γ1.58, 1.63	δ3.17	NH 7.40
K80	8.35	4.33	1.85, 1.77	γ1.41, 1.46	δ1.66, ε2.97	NH7.60
S81	8.4	4.41	3.86			
R82	8.42	4.38	1.77, 1.88	γ1.59, 1.63	δ3.15	NH 7.28
S83	8.4	4.42	3.86			
L84	8.26	4.60	1.56	γ1.56	δ0.91	
P85		4.32	1.92,1.99	γ1.51	δ3.59, 3.80	
T86	8.07	4.23	4.19	γ1.14		
T86c	8.41		4.16	γ1.32		
W87	7.99	4.75	3.28, 3.32	2: 7.21	5: 7.17	7: 7.44
				4: 7.63	6: 7.39	NH 10.19
W87c	8.18	4.72	3.26, 3.32	2: 7.24	7: 7.44	NH 10.15
E88	7.80	4.09	1.79, 1.98	2.08		
E88c	7.84	4.08	1.80, 1.98	2.09		

determine the secondary structure of each residue [33]. Secondary structure conservation scores for all turns, strands, loops and helices are the average of the positional scores within those elements of secondary structure.

3. Results and discussion

3.1. CRABP I 64–88 is monomeric and lacks stable tertiary structure

The CD spectrum of CRABP I 64–88 (not shown) shows a strong minimum below 200 nm and a very weak negative band centered on 232 nm, consistent with a largely random distribution of conformations [52]. Seemingly in contradiction, the amide proton NMR resonances of many residues in this peptide have temperature coefficients below 4 ppb/°C (see below), suggesting intramolecular hydrogen bonding. An alternative possibility that could account for the low temperature coefficients was intermolecular association. However, the CD spectrum does not change with concentration between 24 μM and 4.4 mM. In addition, the 1D NMR spectrum of this fragment exhibits narrow lines (the width at half height of the Trp87 indole proton is 6 Hz) at the concentration of 4.4 mM.

3.2. CRABP I 64–88 samples native-like structure

^{1}Hα chemical shift deviations from random coil provide a residue-specific measure of the conformational state populated by a polypeptide [45,46]. We compared the deviations from random coil for residues in the CRABP I 64–88 fragment with the

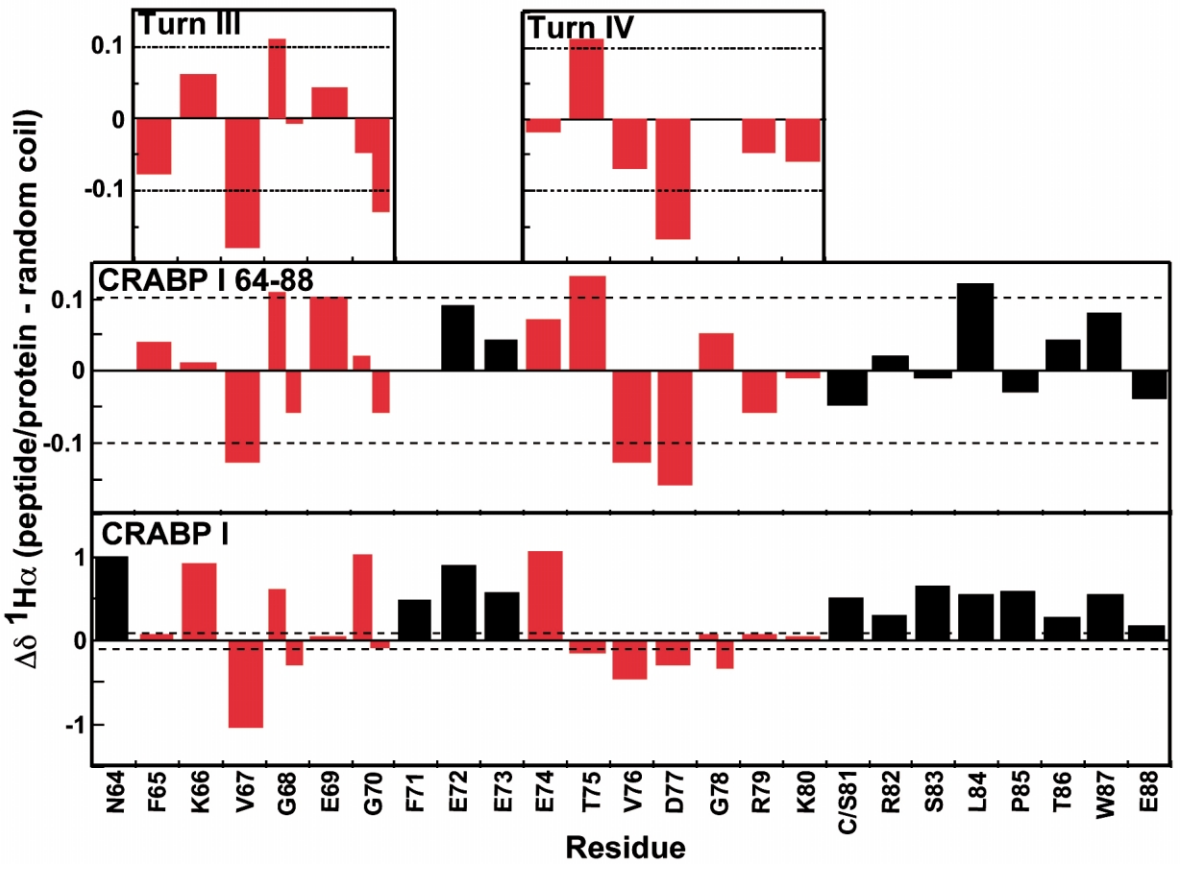

Fig. 2. Hα chemical-shift deviation from random coil for residues in the CRABP I 64–88 fragment. For comparison, the deviations from random coil shifts are shown for the short turn fragments studied (Rotondi and Gierasch, in preparation) and intact CRABP I [57]. Residues in the turn regions are highlighted in red. Note native-like deviations in the turn regions, as well as the strand bridging turns III and IV. Weak and alternating positive and negative deviations in the region beyond turn IV suggest a more random, coil-like behavior. The dotted lines indicate the generally accepted threshold deviations for confident assignment of secondary structure propensity [45,46].

shifts exhibited by short fragments of turns III and IV and with those of intact native CRABP I residues 64–88 (Fig. 2). The magnitudes of the shifts from random coil are considerably smaller both in the short turn fragments and in CRABP I 64–88 (0–0.16 ppm) than in the native protein (0.05–1.05 ppm), as expected for short sequences that populate an ensemble of conformations. Nonetheless, the behavior of residues in turns III and IV is strikingly similar to that exhibited in the intact protein. The pattern of deviation from random coil in the turn sequences is more native-like in the longer fragment. Therefore, this analysis reveals a native bias in the turn III and IV portions of the CRABP I 64–88 fragment. Moreover, the native-like bias of the long fragment continues in the strand that links turns III and IV. With the exception of Phe71, all of the residues in this strand show a downfield shift from random coil, i.e. toward extended conformation. The shift of Phe71 in the intact protein is toward extended conformation, while its chemical shift in the large fragment is random coil. Phe71 forms many tertiary contacts in the native protein, including stacking interactions with Phe50 and Phe65. We suggest that the conformation and environment of this

residue are significantly affected by these interactions, which are absent in the fragment. The behavior of Thr75 is curious. From the crystal structure ϕ,ψ values, Thr75 in the intact protein is near to extended, with a kink that enables it to take part in the conserved polar cluster, and its ^{1}Hα resonance exhibits a modest upfield shift towards turn or helical conformation. Interestingly, the chemical shift of this residue in both the short and long fragments is typical of an undistorted extended conformation. Thus, the difference likely reflects the absence in the fragments of the longer-range network of native contacts. Most of the sequence beyond turn IV in the CRABP I 64–88 fragment varies between weakly extended and weakly helical biases, which indicates a more random coil conformation in this region of the fragment. Only Leu84 appears to be in an extended conformation. Overall, based on deviations in the ^{1}Hα chemical shifts from random coil, the CRABP I 64–88 fragment appears to sample native-like structure in turns III and IV, as well as the strand connecting them.

The ability of the NOESY experiment to report on inter-proton distance can be exploited to determine the short-range conformational biases in a protein fragment. The distance between sequential amide and α-protons is dependent upon the local conformation. In extended regions, the sequential amide–amide distance is longer (~4.5 Å), so the NN(i,i+1) nOe is weaker, while the sequential α-amide distance is rather short (~2.5 Å) and the corresponding αN(i,i+1) nOe is strong [48]. In contrast, the sequential distances reverse in a helix or in the i+3–i+4 positions of a β-turn, with the sequential amide–amide distance falling to ~2.5 Å and the α-amide distance reaching ~3.2 Å. These sequential nOes provide information on the local backbone conformation in short fragments [53].

The sequential nOe pattern observed in the turn regions of CRABP I 64–88 is that expected for the native turns (Figs. 3 and 4). Both the turn regions show a strong NN nOe with a weak αN nOe between the i+3 and i+4 residues—between Gly68 and Glu69 in turn III, and between Asp77 and Gly78 in turn IV. As a result of the glycine bulge at Gly78 in turn IV, the amide protons of

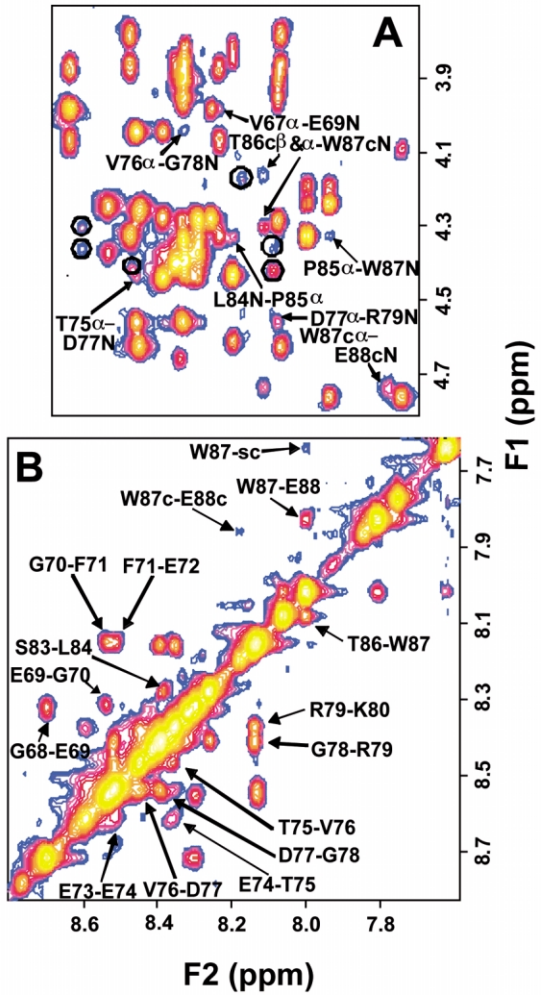

Fig. 3. (a) Fingerprint and (b) amide regions of the CRABP I 64–88 NOESY spectrum. Peaks labeled with a 'c' result from a small population (~15%) of peptides containing a *cis* Leu84–Pro85 bond. Conditions: 3 °C, 4.4 mM CRABP I 64–88, pH 5.3, 10 mM PO_4^{2-}, 10% D_2O in H_2O, t_m=150 ms. Cross-peaks arising from an impurity are circled.

residues Gly78 and Arg79 are considerably closer in the crystal structure (3.0 Å) than expected in extended conformation. The presence of a medium NN and weak αN nOe between this pair is in accord with a native-like conformation at the C-terminus of turn IV, including the glycine bulge. In addition, NN and αN nOes observed in the strand that bridges turns III and IV are consistent with the native extended conformation. The region

Fig. 4. Summary of results of the NOESY analysis, as well as the amide temperature coefficients for the CRABP I 64–88 fragment. The native conformations of residues in the fragment are indicated above the residue numbers (E for extended, T for turn). Strong NN($i,i+1$) and αN(i,i) nOes in the turn regions, coupled with weaker αN($i,i+1$) nOes, indicate that the turns are populated in the 64–88 fragment. By comparison, the presence of stronger sequential αN and weaker NN nOes in the strand linking turns III and IV indicates a more extended conformation. Note that observable non-sequential nOes are associated with turns III and IV, with only one in the more 'random coil' C-terminal region of the fragment. Amide temperature coefficients ($\Delta\delta/\Delta T$ ppb/°C) were measured between 3 and 25 °C; ‡ indicates a residue that would be an intra-fragment hydrogen-bond donor in the native conformation. An asterisk indicates that overlap precluded measurement of a given nOe; no entry for temperature dependence indicates that overlap did not allow that amide resonance to be followed.

beyond turn IV in the CRABP I 64–88 fragment is harder to interpret due to the large amount of spectral overlap; on the other hand, the spectral overlap itself is telling, pointing to a random distribution of conformations in this region. Despite this, the presence of a strong αN nOe coupled with a medium NN nOe provides some indication of extended conformation around Ser83 and Leu84, consistent with the chemical shift of the Leu α-proton. Progressively weaker αN and NN nOes observed approaching the C-terminus of the CRABP I 64–88 fragment suggest a more random coil conformation.

A series of sequential αα nOes is present in the CRABP I 64–88 fragment (data not shown). The residue pairs, with their crystal structure distances are: Lys66–Val67, 4.3 Å; Gly68–Glu69, 4.4 Å; Thr75–Val76, 4.4 Å; Val76–Asp77, 4.7 Å; Asp77–Gly78, 4.3 Å; and Thr86–Trp 87, 4.3 Å. There is also a sequential βα nOe between Thr75 and Val76 (5.1 Å). With the exception of that corresponding to Thr86–Trp87, these sequential nOes are associated with the turn III and IV residues. The sequential αα interproton distance is relatively invariant from strand to turn, remaining at approximately 4.5 Å, which is near the distance limit for an observable Overhauser effect [54,55]. The fact that these nOes are observable suggests that the turns are conformationally stable, with little segmental motion. The complete sequential and non-sequential nOe analyses, as well as the amide temperature coefficients measured for the sufficiently resolved amide protons, are summarized in Fig. 4.

3.3. The native-like structure in CRABP I 64–88 is persistent

The ^{1}Hα chemical shift deviation from random coil and the sequential nOe data argue that indi-

vidual residues and pairs of residues visit local conformations that correspond to the native structure of CRABP I a significant portion of the time. On the other hand, these very local data do not indicate whether the native structure is persistent throughout the peptide. The nOes observed for interproton pairs from non-sequential residues should be more useful indicators of the sampling of native-like structure throughout fragment. In addition, the presence of a native-like hydrogen bonding network, as evidenced by reduced amide temperature coefficients for those amides that are intramolecularly hydrogen bonded, would be strong evidence that native conformation is persistent throughout the fragment. Indeed, these data support the presence of native-like turns in the CRABP I 64–88 fragment.

As previously mentioned, the amide temperature coefficients for CRABP I 64–88 are rather low (average 3.3 ppb/°C). However, inspection of the data (Fig. 4) reveals that, in general, the most reduced temperature coefficients are associated with amides that would be involved in intra-fragment hydrogen bonds in the native conformation (average 2.6 ppb/°C for those amides that are hydrogen-bonded in the protein vs. 3.6 ppb/°C for those that are not; Fig. 1C). The turn hydrogen bonds of Glu69 (2.6 ppb/°C) in turn III and the bifurcated hydrogen bond in turn IV involving the amides of Gly78 (2.7 ppb/°C) and Arg79 (1.6 ppb/°C) show decreased shift with temperature, indicative of the persistence of these turns in the larger fragment. In addition, the temperature coefficients of the expected hydrogen-bond donor amides in the strand between turns III and IV are uniformly lower than their non-hydrogen-bonding neighbors (Fig. 4). We therefore conclude that the native bias exhibited in the ^{1}Hα chemical shift analysis is sufficient to cause partial protection of the amide protons expected to be involved in native hydrogen bonds. More interestingly, this requires at least the partial concomitant presence of native conformation in both turns and the intervening strand.

The non-sequential nOe data likewise support persistent structure over a longer range. Specifically, an $\alpha N(i,i+2)$ nOe is evident between Val67 and Glu69 (Figs. 3 and 4). This nOe is consistent with a native turn, since these residues would occupy the $i+1$ and $i+3$ positions of the turn. The analogous $\alpha N(i,i+2)$ nOe is observed between Val76 and Gly78 in the turn IV region (Fig. 3); the interproton distance in the X-ray crystal structure [33] is consistent with the presence of this nOe. A glycine bulge at Gly78 in turn IV results in close approach of residues Asp77 and Arg79 in this turn in native CRABP I, and a non-sequential $\alpha N(i,i+2)$ nOe observed between these residues in CRABP I 64–88 suggests that the bulge feature is retained for at least some fraction of the time. In addition, an $\alpha N(i,i+2)$ nOe is evident between Thr75 and Asp77. One additional, non-sequential nOe, an $\alpha N(i,i+2)$ between Pro85 and Trp87, was observed in this fragment. This nOe would not be expected in the native conformation of the protein (interproton distance 6.3 Å [33]), suggesting that the fragment loses its native structural bias towards its C-terminus. The presence of the expected non-sequential nOes in the turn III and IV regions of CRABP I 64–88 indicates that the native-like turns continue to be populated in this fragment. Equally important is the absence of such non-sequential nOes in the strand joining turns III and IV, as these would not be expected in a primarily extended region. The sequence beyond turn IV, which appeared more random coil in the ^{1}Hα analysis, is the only area in the fragment where an unexpected non-sequential nOe was observed. In total, the results of the non-sequential nOe analysis are in complete support of the results of the ^{1}Hα chemical shift analysis, indicating native bias in turns III and IV and suggesting an extended conformation in the strand linking them.

It is noteworthy that a minor conformer corresponding to a *cis* isomer of Pro85 can be observed in the TOCSY spectrum of CRABP I 64–88 (Fig. 5). Integration of the Trp87 indole-proton *cis* and *trans* isomers indicates that the *cis* isomer represents approximately 15% of the population. This is intriguing, as approximately 15% of the population of CRABP I folds through a slow phase associated with the *cis–trans* isomerization of Pro85 [34,37]. Thus, the populations of *cis* and *trans* isomers are the same in the fragment and in

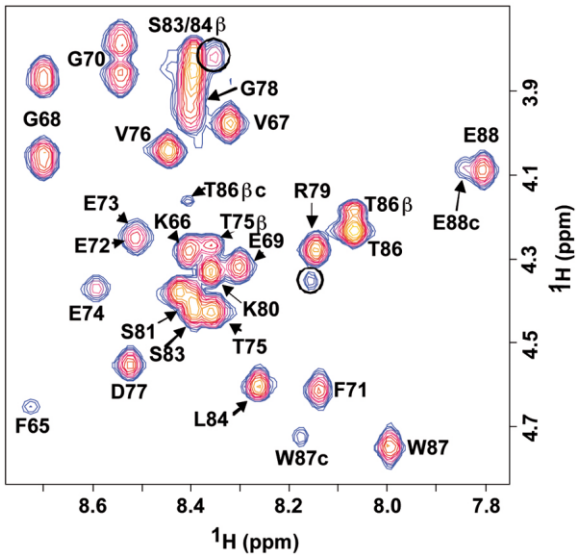

Fig. 5. Fingerprint region of the TOCSY spectrum of the CRABP I 64–88 fragment. Peaks labeled with a 'c' arise from the small population of molecules (~15%) containing a *cis* Leu84–Pro85 bond. *Conditions*: 3 °C, 4.4 mM CRABP I 64–88, pH 5.3, 10 mM PO_4^{2-}, 10% D_2O in H_2O, $t_m = 78$ ms. The circled peaks arise from an impurity.

the denatured protein, arguing that the fragment is a good model of the unfolded state of CRABP I.

Taken together, the results of the biophysical characterization of CRABP I 64–88 suggest continued native bias in turn III and IV, as well as in their intervening strand. Both the sequential nOe and $^1H\alpha$ results produce patterns characteristic of the native conformation in these regions. The non-sequential nOes are located in the turns and expected if the sequence is populating a native-like ensemble, with the exception of one in a portion of the fragment that appears as random coil. The amide temperature-shift data support the presence of native hydrogen bonding in the turns.

3.4. Sequence conservation

The results of the sequence conservation on the secondary structural level show turns III and IV to be the best-conserved turns in CRABP I. Turn IV is, in fact, the best-conserved secondary structural element in the protein (Fig. 6). The intervening strand, strand 5, which is biased toward its native conformation, also displays higher than average conservation. In contrast, strand 6 appears to be primarily random coil conformation in our structural analysis on CRABP I 64–88. This strand has considerably lower sequence conservation on the secondary structure level. These results suggest that regions of CRABP I that display a tendency to adopt native conformation when examined as fragments are also found to be well conserved in the iLBP family. This type of analysis may prove to be a powerful tool in identifying regions of sequence likely to influence early folding events.

3.5. Implications for folding

The presence of a native-like structural bias in the 64–88 region of CRABP I suggests a role in the early folding events of this protein. An intriguing question is whether the locally encoded native structure is manifest in the unfolded state or, alternatively, is characteristic of the earliest intermediate ensemble of CRABP I (formed with a 300-μs time constant). We earlier found that peptides corresponding to turns III and IV retained native-like conformation in 7 M urea, a strong argument that they populate their native-like conformations in the unfolded state of the intact protein. We raised the question whether CRABP I 64–88 shows the same structural tendencies in 7 M urea as in aqueous solution. The $\Delta\delta$ $^1H\alpha$ analysis was repeated under these conditions (data not shown). Despite a slight (0–0.05 ppm) shift towards extended conformation for nearly all the amide protons in the fragment, the pattern retains a native-like quality. The lone exception is Phe71, which shifts upfield. As has been discussed, Phe71 is involved in many global interactions, including one with Phe65, also on the large fragment. We conclude that urea disrupts some longer-range hydrophobic clustering involving Phe71, but that otherwise the native-like bias is retained in urea. Together, these findings suggest that pre-existing biases towards native local structure may exist in the unfolded state. Hydrophobic interactions in this region of the iLBP family that are critical for productive folding likely come into play when the protein is introduced into folding conditions upon dilution from urea. A similar role to that suggested

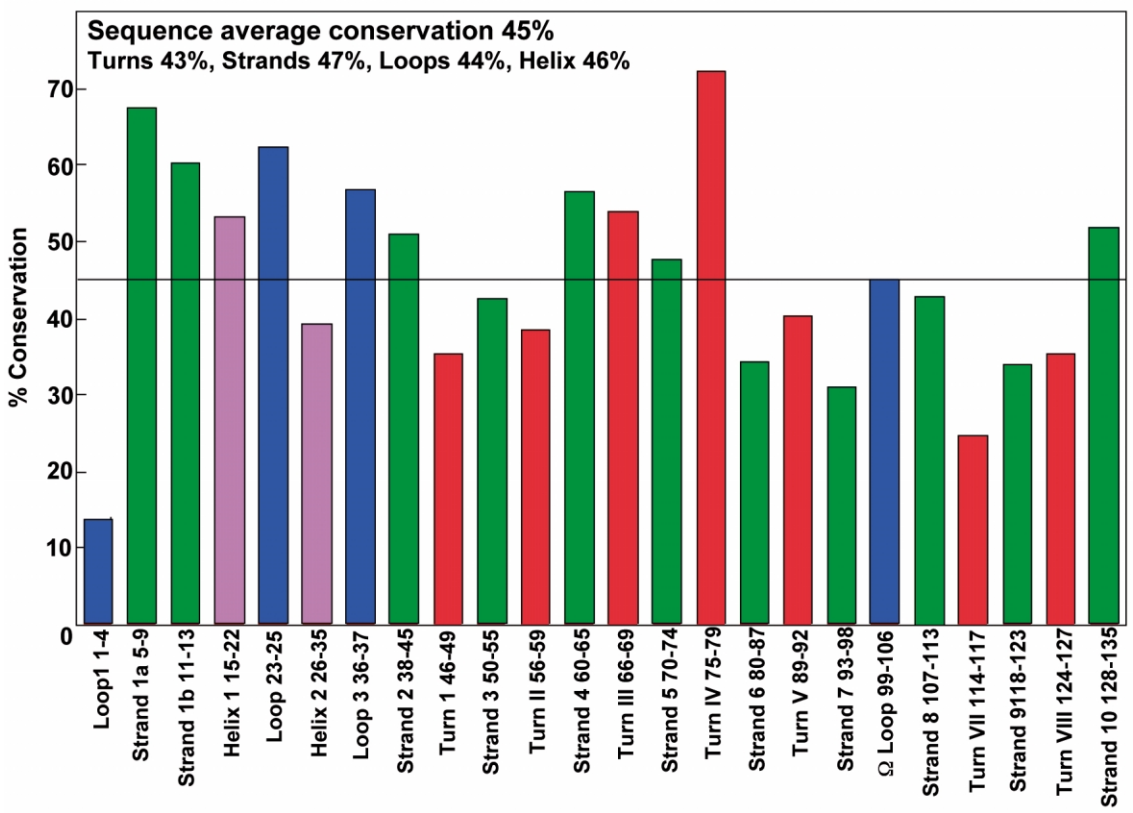

Fig. 6. Conservation of secondary structural elements in CRABP I. Turns are shown in red, strands in green, helix in purple and loops in black. The solid horizontal line indicates the average conservation of all secondary structural elements in this protein. Note that turn IV is the most conserved element in CRABP I, while turn III is the second-best conserved turn. In addition, the strand connecting turns III and IV, strand 5, shows better than average conservation.

for Phe71 has been postulated for a hydrophobic residue in turn III of intestinal fatty-acid-binding protein [56]. The retention of a native-like $\Delta\delta$ $^1H\alpha$ pattern in turns III and IV, as well as in the intervening strand 5, in 7 M urea indicates that the CRABP I 64–88 fragment represents a model of the unfolded state of CRABP I.

Adoption of native-like structure by turns III and IV and their intervening strand in the unfolded state of CRABP I suggests a role as a microdomain, as postulated by the diffusion–collision model of protein folding [9,10]. Prior work in our laboratory has shown that the helix–turn–helix sub-domain of CRABP I is also a locally organized piece of super-secondary structure that may function as a microdomain and may exist in the unfolded state of CRABP I [38]. The presence of two microdomains at either end of a β-sheet suggests that this portion of the molecule could form through the productive collision of these two microdomains. If this were the case, it might be hypothesized that the N-terminal sheet of CRABP I plays a central role in organizing the folding of the rest of the molecule. Such a role is supported by the distribution of conserved pairwise interactions in CRABP I (Gunasekaran et al., in preparation). A highly cross-linked network of conserved interactions was observed in the iLBPs, and seven of the nine residues that participate in four or more conserved interactions are located on the N-terminal sheet of CRABP I and in the turn III–turn IV linked hairpins. These residues form a total of 21 interactions with other N-terminal sheet residues, 11 interactions with residues on the

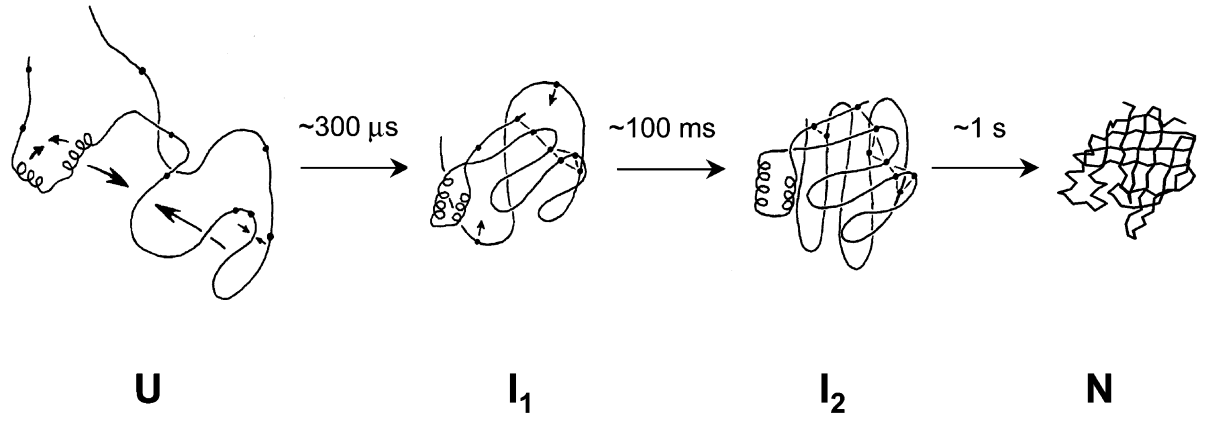

Fig. 7. Proposed model for the kinetic steps in the folding for CRABP I (see text).

linked hairpins and 13 interactions with residues on the C-terminal sheet. Only two of the highly cross-linked conserved interacting residues are on the C-terminal sheet, and these two interact more with residues on the N-terminal sheet (six) than with other residues on the C-terminal sheet (four). As a result, formation of the N-terminal sheet may proceed in the absence of an organized C-terminal sheet and would allow the highly interacting residues on the N-terminal sheet to enter into the bulk of their native interactions. In contrast, the highly interacting conserved residues on the C-terminal sheet would be without the majority of their native interactions in the absence of the organized N-terminal sheet.

Together, these results and observations suggest a model for the folding of CRABP I (Fig. 7). This work has demonstrated that the linked turn III and turn IV hairpins likely populate an ensemble with a native-like bias in the unfolded state. We hypothesize that in the earliest (300 μs) folding phase of CRABP I, the two microdomains identified, which are separated by only 26 residues, diffuse and collide productively. In this event, the highly interacting residues of the N-terminal sheet and the linked hairpins form a network of side-chain interactions that serves to stabilize the intermediate. In the subsequent (100 ms) folding phase, the loosely established N-terminal sheet, with its network of conserved interactions, may serve as a point of organization for the C-terminal sheet. The residues on the C-terminal sheet that are involved in conserved interactions are then positioned to engage in their contacts with N-terminal-sheet side chains, bringing the N- and C-termini together and forming the binding cavity—features that have been observed to form in this kinetic phase [36]. The last step in folding probably involves exclusion of solvent from the strands, packing of side chains into van der Waals contacts, and the cooperative formation of interstrand hydrogen bonds.

Acknowledgments

We appreciate the critical reading of the manuscript by Joanna Swain and Jennifer Habink. During much of this work, KSR was supported on a Chemistry–Biology Interface NIH training grant (GM08515). This work was supported by a grant to LMG from the National Institutes of Health (GM27616).

References

[1] C.B. Anfinsen, Principles that govern the folding of protein chains, Science 181 (1973) 223–230.
[2] K.A. Dill, K.M. Fiebig, H.S. Chan, Cooperativity in protein folding kinetics, Proc. Natl. Acad. Sci. USA 90 (1993) 1942–1946.
[3] K.A. Dill, S. Bromberg, K. Yue, et al., Principles of protein folding—a perspective from simplest exact models, Protein Sci. 4 (1995) 561–602.

[4] V.R. Agashe, M.C.R. Shastry, J.B. Udgaonkar, Initial hydrophobic collapse in the folding of barstar, Nature 377 (1995) 754–757.
[5] R.L. Baldwin, G.D. Rose, Is protein folding hierarchic? I. Local structure and peptide folding, Trends Biochem. Sci. 24 (1999) 26–33.
[6] R.L. Baldwin, G.D. Rose, Is protein folding hierarchic? II. Folding intermediates and transition states, Trends Biochem. Sci. 24 (1999) 77–83.
[7] D.B. Wetlaufer, Nucleation, rapid folding, and globular intrachain regions in proteins, Proc. Natl. Acad. Sci. USA 70 (1973) 697–701.
[8] L.S. Itzhaki, D.E. Otzen, A.R. Fersht, The structure of the transition state for folding of chymotrypsin inhibitor 2 analyzed by protein engineering methods: evidence for a nucleation–condensation mechanism for protein folding, J. Mol. Biol. 254 (1995) 260–288.
[9] M. Karplus, D.L. Weaver, Protein folding dynamics, Nature 260 (1976) 404–406.
[10] M. Karplus, D.L. Weaver, Protein folding dynamics: the diffusion–collision model and experimental data, Protein Sci. 3 (1994) 650–669.
[11] J.E. Brown, W.A. Klee, Helix–coil transition of the isolated amino terminus of ribonuclease, Biochemistry 10 (1971) 470–476.
[12] P.S. Kim, R.L. Baldwin, A helix stop signal in the isolated S-peptide of ribonuclease A, Nature 307 (1984) 329–334.
[13] J. Sancho, J.L. Neira, A.R. Fersht, An N-terminal fragment of barnase has residual helical structure similar to that in a refolding intermediate, J. Mol. Biol. 224 (1992) 749–758.
[14] M.J. McLeish, K.J. Nielsen, J.D. Wade, D.J. Craik, A peptide corresponding to the N-terminal 13 residues of T4-lysozyme forms an α-helix, FEBS Lett. 315 (1993) 323–328.
[15] S. Williams, T.P. Causgrove, R. Gilmanshin, et al., Fast events in protein folding: helix melting and formation in a small peptide, Biochemistry 35 (1996) 691–697.
[16] R. Aurora, T.P. Creamer, R. Srinivasan, G.D. Rose, Local interactions in protein folding: lessons from the α-helix, J. Biol. Chem. 272 (1997) 1413–1416.
[17] R.J. Pinker, L. Lin, G.D. Rose, N.R. Kallenbach, Effects of alanine substitutions in α-helices of sperm whale myoglobin on protein stability, Protein Sci. 2 (1993) 1099–1105.
[18] F. Chiti, N. Taddei, P. Webster, et al., Acceleration of the folding of acylphosphatase by stabilization of local secondary structure, Nat. Struct. Biol. 6 (1999) 380–387.
[19] S. Cavagnero, H.J. Dyson, P.E. Wright, Effect of H helix destabilizing mutations on the kinetic and equilibrium folding of apomyoglobin, J. Mol. Biol. 285 (1999) 269–282.
[20] N. Taddei, F. Chiti, T. Fiaschi, et al., Stabilization of α-helices by site-directed mutagenesis reveals the importance of secondary structure in the transition state for acylphosphatase folding, J. Mol. Biol. 300 (2000) 633–647.
[21] F.J. Blanco, M.A. Jimenez, J. Herranz, M. Rico, J. Santoro, J.L. Nieto, NMR evidence of a short linear peptide that folds into a β-hairpin in aqueous solution, J. Am. Chem. Soc. 115 (1993) 5887–5888.
[22] F. Blanco, L. Serrano, Folding of protein G B1 domain studied by the conformational characterization of fragments comprising its secondary structure elements, Eur. J. Biochem. 230 (1995) 634–649.
[23] M.S. Searle, H.W. Dudley, L.C. Packman, A short linear peptide derived from the N-terminal sequence of ubiquitin folds into a water-stable non-native β-hairpin, Nat. Struct. Biol. 2 (1995) 999–1006.
[24] M.S. Searle, R. Zerella, D.H. Dudley, L.C. Packman, Native-like β-hairpin structure in an isolated fragment from ferredoxin: NMR and CD studies of solvent effects on the N-terminal 20 residues, Protein Eng. 9 (1996) 559–565.
[25] A.R. Viguera, M.A. Jiménez, M. Rico, L. Serrano, Conformational analysis of peptides corresponding to β-hairpins and a β-sheet that represent the entire sequence of the α-spectrin SH3 domain, J. Mol. Biol. 255 (1996) 507–521.
[26] J.C. Martínez, L. Serrano, The folding transition state between SH3 domains is conformationally restricted and evolutionarily conserved, Nat. Struct. Biol. 6 (1999) 1010–1016.
[27] M.S. Briggs, H. Roder, Early hydrogen-bonding events in the folding reaction of ubiquitin, Proc. Natl. Acad. Sci. USA 89 (1992) 2017–2021.
[28] P.-Y. Chen, B.G. Gopalacushina, C.-C. Yang, S.I. Chan, P.A. Evans, The role of a β-bulge in the folding of the β-hairpin structure in ubiquitin, Protein Sci. 10 (2001) 2063–2074.
[29] F. Blanco, L. Serrano, Folding of protein G B1 domain studied by the conformational characterization of fragments comprising its secondary structure elements, Eur. J. Biochem. 230 (1995) 634–649.
[30] E.L. McCallister, E. Alm, D. Baker, Critical role of β-hairpin formation in protein G folding, Nat. Struct. Biol. 7 (2000) 669–673.
[31] A. Capaldi, S.E. Radford, Kinetic studies of β-sheet protein folding, Curr. Opin. Struct. Biol. 8 (1998) 86–92.
[32] M. Bucciantini, E. Giannoni, F. Chiti, et al., Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases, Nature 416 (2002) 507–511.
[33] G.J. Kleywegt, T. Bergfors, H. Senn, et al., Crystal structures of cellular retinoic acid binding proteins I and II in complex with all-*trans* retinoic acid and a synthetic retinoid, Structure 2 (1994) 1241–1258.
[34] P.L. Clark, Z.-P. Liu, J. Zhang, L.M. Gierasch, Intrinsic tryptophans of CRABP I as probes of structure and folding, Protein Sci. 5 (1996) 1108–1117.

[35] P.L. Clark, B.F. Weston, L.M. Gierasch, Probing the folding pathway of a β-clam protein with single-tryptophan constructs, Fold. Des. 3 (1998) 401–412.
[36] P.L. Clark, Z.-P. Liu, J. Rizo, L.M. Gierasch, Cavity formation before stable hydrogen bonding in the folding of a β-clam protein, Nat. Struct. Biol. 4 (1997) 883–886.
[37] S.J. Eyles, L.M. Gierasch, Multiple roles of prolyl residues in structure and folding, J. Mol. Biol. 301 (2000) 737–747.
[38] M. Sukumar, L.M. Gierasch, Local interactions in a Schellman motif dictate interhelical arrangement in a protein fragment, Fold. Des. 2 (1997) 211–222.
[39] R.B. Merrifield, Solid phase synthesis of linear peptides, J. Am. Chem. Soc. 85 (1963) 2149–2154.
[40] S.C. Gill, P.H. von Hippel, Calculation of protein extinction coefficients from amino acid sequence data, Anal. Biochem. 182 (1989) 319–326.
[41] F. Kasler, Quantitative Analysis by NMR Spectrometry, Academic Press, New York, 1973, pp. 78–88.
[42] M. Liu, C. Mao, H. He, J.K. Huang, J.K. Nicholson, J.C. Lindon, Improved WATERGATE pulse sequences for solvent suppression in NMR spectroscopy, J. Magn. Reson. 132 (1998) 125–129.
[43] L. Braunschweiler, R.R. Ernst, Coherence transfer by isotopic mixing: application to proton correlation spectroscopy, J. Magn. Reson. 53 (1983) 521–528.
[44] G.D. Rose, L.M. Gierasch, J.A. Smith, Turns in peptides and proteins, Adv. Protein Chem. 37 (1985) 1–109.
[45] D.S. Wishart, B.D. Sykes, F.M. Richards, Relationship between nuclear magnetic resonance chemical shift and protein secondary structure, J. Mol. Biol. 222 (1991) 311–333.
[46] D.S. Wishart, B.D. Sykes, F.M. Richards, The chemical shift index—a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy, Biochemistry 31 (1992) 1647–1651.
[47] A. Bax, D.G. Davis, MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy, J. Magn. Reson. 65 (1985) 355–360.
[48] K. Wüthrich, NMR of Proteins and Nucleic Acids, J. Wiley & Sons, New York, 1986.
[49] C. Sander, R. Schneider, Database of homology-derived protein structures and the structural meaning of sequence alignment, Proteins: Struct. Funct. Genet. 9 (1991).
[50] B. Rost, Twilight zone of protein sequence alignments, Protein Eng. 12 (1999) 85–94.
[51] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.
[52] R.W. Woody, in: S. Udenfriend, J. Meienhofer (Eds.), The Peptides, vol. 7, Academic Press, Orlando, 1985, p. 15.
[53] P.E. Wright, H.J. Dyson, R.A. Lerner, Conformation of peptide fragments of proteins in aqueous solution: implications for initiation of protein folding, Biochemistry 27 (1988) 7167–7175.
[54] H.J. Dyson, P.E. Wright, Defining solution conformations of small linear peptides, Annu. Rev. Biophys. Biophys. Chem. 20 (1991) 519–538.
[55] H.J. Dyson, P.E. Wright, Peptide conformation and protein folding, Curr. Opin. Struct. Biol. 3 (1993) 60–65.
[56] K. Kim, C. Frieden, Turn scanning by site-directed mutagenesis: application to the protein folding problem using the intestinal fatty-acid-binding protein, Protein Sci. 7 (1998) 1821–1828.
[57] J. Rizo, Z.P. Liu, L.M. Gierasch, ^{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 (1994) 741–760.