Role of Local Sequence in the Folding of Cellular Retinoic Acid Binding Protein I: Structural Propensities of Reverse Turns†

Kenneth S. Rotondi and Lila M. Gierasch*

Departments of Chemistry and of Biochemistry and Molecular Biology, University of Massachusetts, Amherst, Massachusetts 01003

Received February 24, 2003; Revised Manuscript Received May 9, 2003

ABSTRACT: The experiments described here explore the role of local sequence in the folding of cellular retinoic acid binding protein I (CRABP I). This is a 136-residue, 10-stranded, antiparallel β-barrel protein with seven β-hairpins and is a member of the intracellular lipid binding protein (iLBP) family. The relative roles of local and global sequence information in governing the folding of this class of proteins are not well-understood. In question is whether the β-turns are locally defined by short-range interactions within their sequences, and are thus able to play an active role in reducing the conformational space available to the folding chain, or whether the turns are passive, relying upon global forces to form. Short (six- and seven-residue) peptides corresponding to the seven CRABP I turns were analyzed by circular dichroism and NMR for their tendencies to take up the conformations they adopt in the context of the native protein. The results indicate that two of the peptides, encompassing turns III and IV in CRABP I, have a strong intrinsic bias to form native turns. Intriguingly, these turns are on linked hairpins in CRABP I and represent the best-conserved turns in the iLBP family. These results suggest that local sequence may play an important role in narrowing the conformational ensemble of CRABP I during folding.

Despite many years of research, the code that directs the acquisition of the native structure of a protein is not yet well-understood (*1*−*3*). On the basis of extensive work from many laboratories, it has become increasingly clear that folding is governed by a balance between the propensity of local sequence regions to adopt secondary structure and the tendency of the chain to collapse to minimize the hydrophobic surface area, based largely on the global pattern of hydrophobic and hydrophilic amino acids (*4*). Assessing the relative importance of local and global sequence information in folding of different protein structural classes is facilitated by characterizing the behavior of peptide fragments "excised" from proteins and thus removed from the influence of long-range interactions (*5*, *6*).

Numerous studies of fragments of α-helical proteins suggest that locally defined α-helices may actively limit the conformational space of the folding chain (*6*−*8*). These results with peptide fragments have been shown to correlate directly with the behavior of the same segments in the context of the intact folding polypeptide (*9*, *10*). Together, these findings have provided strong support for "diffusion−collision" models of protein folding (*11*), wherein the formation of local metastable secondary structure precedes specific collapse into supersecondary and tertiary structure (*12*, *13*). In turn, the specific collapse stabilizes native secondary structural components. Moreover, recent studies have demonstrated that there is considerable residual structure present in unfolded proteins, and much of this residual structure corresponds to regions of high local structural propensity (*14*, *15*).

An equivalent understanding of the relative roles played by local and global sequence and the nature of any residual structure in the unfolded ensemble has lagged for proteins rich in β-structure (*6*, *16*). Formation of β-structure necessarily requires long-range interactions, since regions distant in sequence come together to form partner strands. On the other hand, in β-hairpin models, it is well-established that strand association and cross-strand interactions are favored by formation of the intervening reverse turn (*6*, *17*, *18*). Yet there are relatively few examples where these propensities have been assessed in the context of the folding reaction for a β-rich protein. Turns are observed in folding transition state ensembles of protein G B1 and SH3 domains (*19*, *20*). In fact, mutations that are predicted to alter the stability of these turns influence the structure of the transition state of the SH3 domain (*21*). Thus, turn propensity clearly can play a determinative role in folding mechanisms. The proteins for which these results were obtained are small and fold by two-state mechanisms (i.e., they do not show observable intermediates in folding). Whether the relatively weak biases toward native turns observed in short peptides can play a significant role in the folding of larger proteins and how turn formation is balanced by longer-range interactions in different topological classes of β-protein remain important, unanswered questions.

Our laboratory has been examining these questions in the β-clam protein cellular retinoic acid binding protein I (CRABP I).[1] CRABP I is a 136-residue member of the extensive family of intracellular lipid binding proteins

† This work was supported by NIH Grant GM27616 (L.M.G.). K.S.R. was supported in part by a NIH Chemistry-Biology Interface predoctoral training grant (GM08515).

* To whom correspondence should be addressed. E-mail: gierasch@biochem.umass.edu.

10.1021/bi034304k CCC: $25.00
Published on Web 06/12/2003

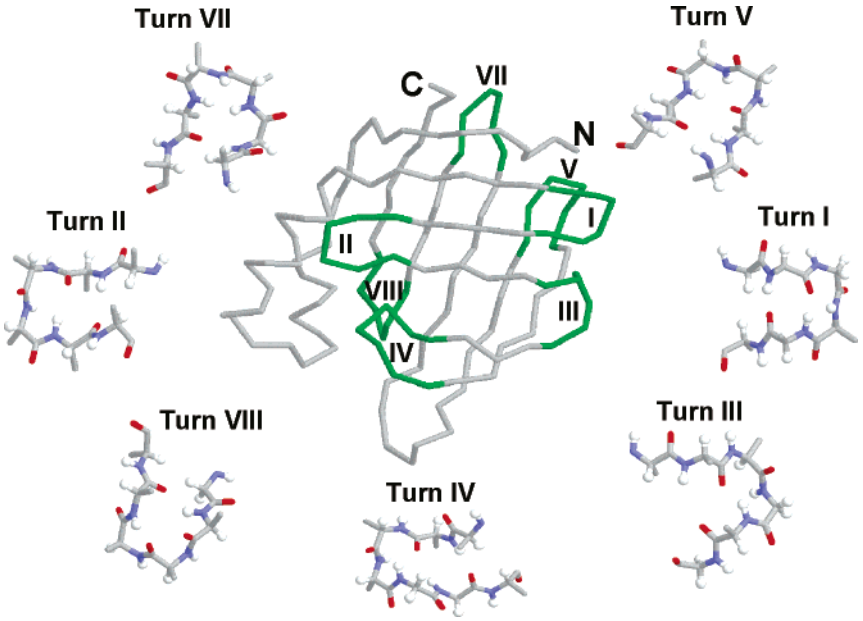

FIGURE 1: CRABP I and its component turns. Backbone representations of the fragments synthesized are shown as balls and sticks. Omitted from the turn structures are the N-acetyl and C-amide protecting groups [figure prepared from PDB entry 1cbi (*23*) using RasMol].

(iLBPs), all of which are made up of 10 β-strands in up−down antiparallel β-structure (*22*). The resulting β-barrel iLBPs contain seven β-hairpins [Figure 1 (*23*)]. A small helix−turn−helix subdomain is inserted between strands 1 and 2, while an Ω-loop links strands 7 and 8. We have analyzed in detail the kinetic phases in the folding of CRABP I using stopped-flow fluorescence, stopped-flow CD, quenched flow and competition hydrogen exchange, and ligand binding assays (*24*−*26*). These data indicate an early collapse with a time constant of 300 μs in which a large amount of secondary structure is formed. Topology formation and the ability to bind ligand occur on a subsequent time scale of 100 ms, while the bulk of the protein achieves a stable tertiary structure on a time scale of 1 s.

This study focuses on fragments corresponding to the turn regions from CRABP I. The conformational ensembles of short (six- and seven-residue) peptide fragments composed of the CRABP I β-turn sequences were characterized to test the potential for local sequence in CRABP I turns to influence its folding. The synthesized sequence and turn type (*27*, *28*) are shown in Table 1. Our results demonstrate that two and only two of the turns (III and IV) have a tendency to take up their native structures in these excised fragments. Interestingly, these two share a common β-strand and occur midway along the length of the chain. Moreover, these two turns exhibit the highest level of sequence conservation among all turns in a set of 52 homologous iLBPs. In related work, we have also shown that the native structural propensity of turns III and IV persists in a longer fragment (CRABP I 64−88) (*29*), reinforcing their potential influence on the folding pathway of this protein.

[1] Abbreviations: CD, circular dichroism; CRABP I, cellular retinoic acid binding protein I; Fmoc, 9-fluorenylmethoxycarbonyl; HPLC, high-pressure liquid chromatography; iFABP, intestinal fatty acid binding protein; iLBP, intracellular lipid binding protein; NMR, nuclear magnetic resonance; rOe, rotating-frame Overhauser enhancement; ROESY, rotating-frame Overhauser enhancement spectroscopy; TOCSY, total correlation spectroscopy.

Table 1: β-Turns of CRABP I

turn[a]	sequence synthesized[b]	native turn type[c]
I	^{45}QDGDQF^{50}	II′ (2:2)
II	^{55}STTVRT^{60}	I (2:2)
III	^{65}FKVGEG^{70}	II (irregular)
IV	^{74}ETVDGRK^{80}	I with Gly bulge (3:5)
V	^{88}ENENKI^{93}	I (2:2)
VII	^{113}LANDEL^{118}	irregular (2:4)
VIII	^{123}GADDVV^{128}	irregular (2:4)

[a] An alternative notation based upon the strands linked has been used in ref *24*: BC for turn I, CD for turn II, DE for turn III, EF for turn IV, FG for turn V, HI for turn VII, and IJ for turn VIII. [b] Peptides were synthesized with N- and C-termini blocked with acetyl and amide groups, respectively. [c] Turn types in the standard notation (*28*) and that of Sibanda et al. (*29*).

MATERIALS AND METHODS

Peptide Synthesis, Purification, and Characterization. Peptides comprised of the residues in the reverse turns of CRABP I plus two flanking residues were synthesized by the solid-phase technique of Merrifield (*30*) either manually or using a Perkin-Elmer (PE) Biosystems 9050 Plus automated synthesizer. The N- and C-termini of the fragments were protected by acetylation and amidation, respectively, to prevent interactions of charged termini. Support resin and the coupling reagent *N*-[(dimethylamino)-1*H*-1,2,3-triazolo[4,5-*b*]pyridin-1-ylmethylene]-*N*-methylmethanaminiumhexyfluorophosphate *N*-oxide were from PE Biosystems (Framingham, MA). Fmoc amino acids were from Peptides Int. (Louisville, KY).

Peptides were purified by reverse-phase HPLC (RP-HPLC) using C-18 preparative columns (Vydac). The purity was checked by analytical HPLC. Identities of the purified peptides were confirmed by NMR spectroscopy, amino acid analysis (Cornell Biotechnology Analytical Chemistry Facility, Ithaca, NY), and mass spectrometry (the Mass Spectrometry Center at the University of Massachusetts).

CD Spectroscopy. CD spectra were collected on a JASCO J-715 spectropolarimeter in 10 mM phosphate buffer at pH 5.3 and 5 °C as previously reported (*29*).

NMR Spectroscopy. NMR analysis was performed on a Bruker 600 MHz Avance NMR spectrometer with a 5 mm TXI SB probe with Z gradients as previously reported (*29*). Concentrations of NMR samples were between 1 and 3 mM; no concentration dependence was observed by CD or NMR for any of the fragments studied from micro- to millimolar concentrations. Peptide spin systems were identified, and amide temperature coefficients ($\Delta\delta/\Delta T$) and ^{1}Hα deviation from random coil values were measured using two-dimensional (2D) total correlation (TOCSY) spectroscopy as previously reported (*29*). Interproton distances in the peptides were monitored using the rotating-frame nuclear Overhauser effect spectroscopy (ROESY) experiment, using States−TPPI phase cycling (*31*−*33*). ROESY spectra were collected at 3 °C with mixing times of 75, 150, and 300 ms with a spectral width of 6127.4 Hz or approximately 10 ppm in both dimensions, 2048 complex points in t_2, and 1024 experiments of 32 or 64 scans each in t_1. The data in t_1 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.

Folded Population Analysis. The tendencies of peptides to take up nonrandom structure were assessed by either of two methods: calculation of their ^{1}Hα chemical shift deviations from random coil values and comparison to literature values for 100% turn (*34*) or comparison of peptide ^{1}Hα chemical shifts to those of native CRABP I as a reference for a 100% folded population (*35*). The chemical shifts from random coil values of the turn residues ($i + 2$ and $i + 3$ for turns I−III, V, VII, and VIII and $i + 2$, $i + 3$, and $i + 4$ for turn IV) were used in this analysis. The shift at each position was divided by the reference shift to generate a percent folded for each residue. These were averaged over the turn residues to yield the percent folded for the fragment.

Calculation of Secondary Structure Conservation. The sequences of 52 members of the iLBP family were aligned, and the number of occurrences of the CRABP I residue at a given position was divided by the number of available sequences at that position to generate the positional conservation score. These positional conservation scores were then averaged over the secondary structural elements to generate secondary structural conservation, as previously reported (*29*).

RESULTS

Overview. Because peptides that are the length of the turn fragments of CRABP I studied here sample a broad ensemble of states, with biases toward native structure anticipated to be small, analysis of their conformational distributions relied on multiple parameters. Interpretation of observed parameters was challenged by the lack of reliable values for standard conformational states. The most compelling parameters were interproton Overhauser effects, particularly those between non-neighboring amino acids, as they argued for persistence of folded states. Inter-residue (between neighboring residues) and intra-residue Overhauser effects were used to test the consistency of any conformational interpretations for the peptides. The careful analyses of Smith and co-workers were consulted to assess the implications of observed NMR data more rigorously (*36*, *37*). Deviations of α-proton chemical shifts from random coil values ($\Delta\delta_{rc}$ Hα) also indicate nonrandom sampling of conformational space (*38*, *39*). It is generally agreed that these deviations should be considered with caution in deriving conformational conclusions unless they are greater than 0.1 ppm. Correlation of observed deviations of chemical shifts from random coil values with those for the same segment in the folded protein may prove to be useful, but caution should be exercised because of perturbations due to tertiary packing or ring-current effects in the folded protein. The temperature dependence of the amide proton chemical shift, the so-called amide temperature coefficient ($-\Delta\delta/\Delta T$, parts per billion per degree Celsius), identifies a significant population with a given amide proton either intramolecularly hydrogen-bonded or otherwise sterically shielded from solvent in folded proteins. Recent work suggests that amide temperature coefficients of ≤4.6 ppb/°C can be safely assumed to arise from solvent shielding via either hydrogen bonding or steric effects, again in folded proteins (*40*, *41*). For peptides undergoing conformational averaging, the interpretation of amide proton chemical coefficients is less straightforward (*42*). In these instances, combining the amide $\Delta\delta/\Delta T$ data with amide proton shifts from random coil ($\Delta\delta_{rc}$ NH) has been recommended by Andersen and co-workers (*42*), and we have examined our data in this manner as well. Last, $^3J_{\alpha NH}$ coupling constants point to residues that have nonrandom conformational biases. We examined all of these NMR parameters, together with CD spectra, to draw conclusions about the behavior of the CRABP I turn peptides. The proton chemical shifts and $^3J_{\alpha NH}$ coupling constants for all turn peptides are provided as Supporting Information.

Only Peptides Corresponding to Turns III and IV Display Significant Bias toward Native Structure. In Figure 2 are shown the sequential and nonsequential rotating-frame interproton Overhauser effects (rOes) for all seven turn peptides, along with the amide proton $\Delta\delta/\Delta T$ and $^3J_{\alpha NH}$ values, while in Figure 3, their CD spectra in aqueous buffer are shown. Although several of the peptides display nonrandom rOe patterns, only the data for peptides corresponding to turns III and IV are consistent with their structures in folded CRABP I. We discuss briefly here the salient features of the turn peptides that are not native-like, and then expand in greater detail on the data supporting strikingly strong biases toward native structure in turns III and IV.

Turn I, ^{45}QDGDQF^{50}, is a type II′ turn in CRABP I. As shown in Figure 2, however, the pattern of rOes for the turn I peptide is neither native nor a combination of native with random coil. The uniform and intense NN(i,i+1) rOes suggest that this fragment adopts a series of overlapping turns, as has been correlated in other peptides with so-called "nascent helix" (*43*). Observed weak i,i+2 rOes are not unexpected in a fragment populating a nascent helix conformation. The intensity of the CD spectrum for this fragment (Figure 3) is weak with a minimum around 200 nm, characteristic of random coil peptides, and suggests the presence of a large fraction of random conformers (*44*). The $^3J_{\alpha NH}$ coupling constants for this fragment likewise support a conclusion of random coil as all are near the calculated random coil value of 6.6 Hz (*37*). The amide temperature coefficients are not consistent with a bias toward the native conformation since the resonance of the expected hydrogen bond donor, Gln49, does not have a reduced temperature dependence.

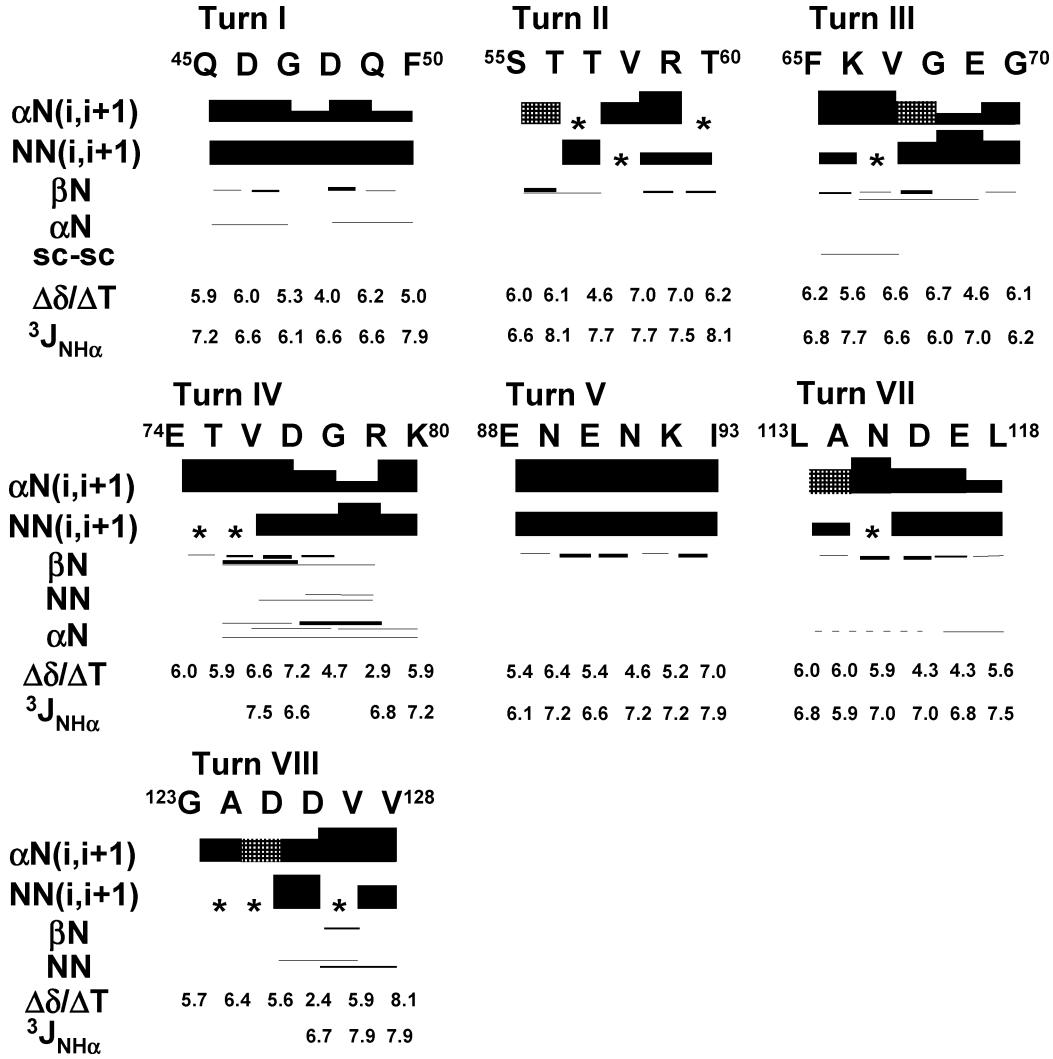

FIGURE 2: ROESY sequential and nonsequential rOes, amide temperature coefficients ($-\Delta\delta/\Delta T$ in parts per billion per degree Celsius), and $^3J_{\alpha NH}$ (Hz) data. The hatched boxes are used to signify an αN sequential rOe that is overlapped with an autocorrelation peak. In each case, these peaks were very high in intensity; however, in the absence of the ability to quantify them, they are shown with medium intensity. Asterisks are used to denote sequential rOes that were not measurable due to overlap either with the diagonal (NN) or with other sequential rOes (αN).

Similarly, the peptide corresponding to turn II, ^{55}STTVRT^{60}, which adopts a type I β-turn in CRABP I, shows no indication of a bias toward a native-like conformation. Here, data on the peptide are consistent with a primarily extended conformation. All observable αN rOes are medium or strong, while the NN rOes are medium or weak. The $^3J_{\alpha NH}$ coupling constants are large (7.5−8.1 Hz), with the exception of that of Ser55, which has a characteristic random coil value of 6.6 Hz. With the exception of Ser55, all of the ^{13}Cα chemical shifts indicate an extended bias (data not shown). The CD spectrum for this fragment (Figure 3) is indicative of a primarily random coil conformation.

In CRABP I, turn V, ^{88}ENENKI^{93}, is a type I β-turn centered on E90 and N91. The corresponding peptide shows no indications of a native-like turn populating its conformational ensemble. The CD spectrum of this fragment (Figure 3) is weak and suggests a primarily random coil conformation. The $^3J_{\alpha NH}$ coupling constants are, with the exception of that of I93 (7.9 Hz), all near the random coil value (6.1−7.2 Hz). The sequential rOe pattern for the fragment derived from turn V in CRABP I is not indicative of any structural bias. The fragment displays uniformly strong sequential αN rOes and uniformly medium sequential NN rOes at all positions (Figure 2).

In an analogous fashion, careful inspection of the data for the peptides corresponding to the two irregular turns from CRABP I, turns VII and VIII, shows no indications of a native bias in their conformational ensembles. For the turn VII peptide, A114 demonstrates a reduced $^3J_{\alpha NH}$ coupling constant of 5.9 Hz for A114. However, the sequential NN rOes are of medium intensity with the exception of the L113−A114 NN rOe, which is weak; there are no significant nonsequential rOes, and the CD spectrum is random-like. There are two consecutive reduced amide temperature coefficients in the turn VII peptide, those of residues Asp116 and Glu117, but the absence of rOe data indicating any structure that could account for the reduced temperature coefficients, and the lack of a correlation between the $\Delta\delta/\Delta T$ and $\Delta\delta_{rc}$ NH (*42*) suggests that they are due to conformational averaging among non-native structures. For the turn VIII peptide, ^{123}GADDVV^{128}, again the CD spectrum is indicative of a primarily random coil conformation. There are several chemical shift overlaps in the NMR spectrum of this fragment, resulting in an incomplete set of data; the lack

of spectral dispersion itself is telling, suggesting a random distribution of conformers. Moreover, the available sequential pattern does not reflect the pattern predicted by the CRABP I crystal structure. The first cross-strand hydrogen bond donor, if the turn VIII peptide adopted a native-like structure, would be Ala124. While this residue does not exhibit a reduced temperature coefficient, Asp126 does (2.4 ppb/°C). This is not supportive of a native-like turn, but raises the possibility of a non-native turn centered on residues Ala124 and Asp125. However, the absence of support for any turn conformation by ROESY makes this unlikely. Here also, the amide temperature coefficient versus amide random coil shift analysis of Andersen and co-workers (*42*) does not support a well-defined conformational equilibrium.

In contrast with all of the peptides discussed thus far, the fragments corresponding to turns III and IV display a number of spectral indicators suggesting that they significantly populate native turns in their conformational ensemble. The CD spectrum of the turn III peptide, ^{65}FKVGEG^{70}, which is a type II turn in native CRABP I, is remarkable (Figure 3), resembling the reference spectrum for a type II β-turn (*44*), and immediately suggests the presence of native conformers. The potential for this CD spectrum to be due to the presence of an aromatic group was tested in a turn III variant incorporating a cyclohexylalanine residue for Phe65. This fragment continues to display a CD spectrum indicative of a type II β-turn, as well as the NMR indicators of native structure seen in the parent fragment (data not shown). The sequential rOe pattern for the turn III fragment strongly indicates a native bias. Its native conformation should produce a strong αN rOe between the $i + 1$ and $i + 2$ residues (Val67 and Gly68, respectively) and a strong NN rOe between the $i + 2$ and $i + 3$ residues (Gly68 and Glu69, respectively) (*45*). An intense NN rOe is present between Gly68 and Glu69, indicative of native conformation; however, the αN rOe between Val67 and Gly68 cannot be observed due to resonance overlap. The αN sequential rOe pattern shows strong rOes in the N-terminal region of this fragment coupled with a weak observable NN rOe (Figure 2). In addition to the evidence in the sequential rOe pattern, the nonsequential rOes of this fragment are equally supportive of a conformational bias toward a native turn. The side chain−side chain (i,i+2) rOe seen between the Phe65 2,6 and 3,5 protons and the Val67 γ proton would be expected in a fragment populating a native turn (Figure 2), and suggests a hydrophobic interaction in this short fragment. Most remarkable is the βN(i,i+3) rOe observed between Lys66 and Glu69 (Figure 2); these protons are 3.6 Å apart in the CRABP I crystal structure. This distance is highly indicative of a turn centered about residues Val67 and Gly68.

Additional support for a native-like conformational bias in the turn III peptide comes from $^3J_{\alpha NH}$ coupling constant analysis, showing that Gly68 (expected to be α_L in the native conformation) has the second lowest coupling constant (6.0 Hz) seen in all the fragments; amide temperature coefficients, where the amide shift with temperature for residue Glu69 is reduced (4.6 ppb/°C); and comparison of the ^{1}Hα chemical shifts from random coil for the turn residues in the fragment with those of intact CRABP I (data not shown), which shows similar patterns for the central four residues. The analysis of Andersen et al. (*42*) also argues that this peptide is in equilibrium between a well-defined conformation and a less ordered state.

The folded population for the turn III fragment was calculated in two ways, either using the ^{1}Hα chemical shifts from random coil in CRABP I (*35*) as a reference for 100% folded (CRABP I Hα shifts are −1.05 ppm for Val67 and 0.61 ppm for Gly68) or using the average chemical shifts from random coil for residues in turns as a reference for 100% folded (*34*, *46*). The ^{1}Hα chemical shifts from random coil measured in the turn III fragment are −0.18 ppm for Val67 and 0.11 ppm for Gly68. These two methods yield estimates of the population of native conformer of 18 and 63%, respectively. Extending this analysis to the i and i + 4 residues in the turn and using the CRABP I and average extended shifts result in populations of 12 and 37%, respectively. The divergence of these results is not surprising given the limitations of the respective methods of quantifying conformational ensembles, but both numbers are high enough to demonstrate that this short peptide has a strong conformational bias. Comparisons with other peptides in the literature show this to be one of the best-formed short peptide turns and/or hairpins studied (*47*). This is especially significant given the general relationship between hairpin population and hairpin length (*47*).

Like the turn III peptide, the fragment corresponding to turn IV, ^{74}ETVDGRK^{80}, exhibits a large number of indicators of a native-like turn populating its conformational ensemble. Turn IV in CRABP I is a type I turn with a glycine bulge into α_L conformational space. While the CD spectrum suggests the presence of random coil conformers, the large number of nonsequential rOes that would be expected in the native turn presents strong evidence for a native bias in this fragment (Figures 2 and 4). Unambiguous evidence for a population with a native-like turn is the presence of the βN-(i,i+4) rOe between residues Thr75 and Arg79, which are 4.1 Å apart in the CRABP I crystal structure (*23*). The expected αN(i,i+2) rOe between residues Val76 and Gly78 is observed (Figure 4A) (*45*). The NN(i,i+2) and βN(i,i+2) rOes seen between residues Asp77 and Arg79 are also consistent with a native-like turn that includes the α_L bulge at Gly78, as this conformation brings these protons into proximity. Serving to confirm the proximity of the Thr75 and Asp77 residues, expected in a turn conformation, are αN(i,i+2) and βN(i,i+2) rOes. The distances between these protons in the crystal structure are 4.6 and 3.4 Å, respectively (*23*). The observation of a αN(i,i+2) rOe between residues Gly78 and Lys80 and an αN(i,i+5) rOe between Thr75 and Lys80 [7.4 and 7.2 Å in the crystal structure, respectively (*23*)] suggest a non-native ψ angle for residue Arg79 at the C-terminus of this fragment. Despite the deviation from the native conformation implied at the C-terminus of the fragment, the sequential and nonsequential rOe data support a population of a remarkably native conformation in the structural ensemble of the turn IV fragment.

The expected NN sequential rOe pattern for a type I turn includes medium and strong rOes between residues $i + 1$ and $i + 2$ (Val76 and Asp77, respectively) and $i + 3$ and i + 4 (Asp77 and Gly78, respectively) in the turn (*45*). The αN rOe pattern is expected to be weak in the turn residues (*45*). In the turn IV fragment, medium NN rOes between Val76 and Asp77 and between Asp77 and Gly78 and a strong NN rOe between Gly78 and Arg79 are seen (Figure 2). These

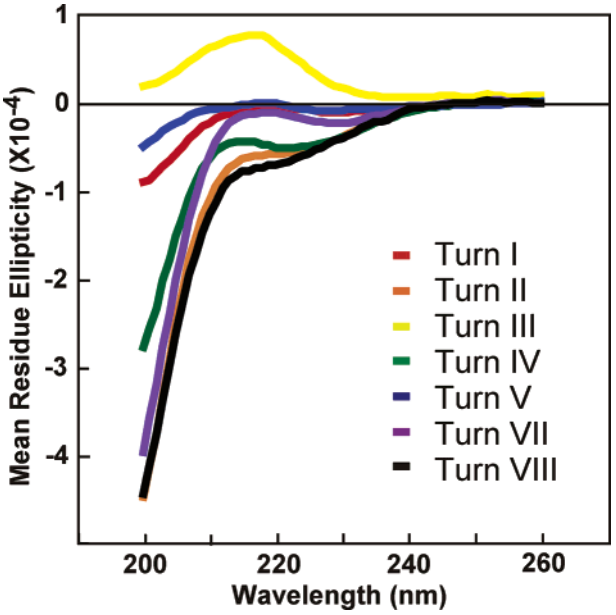

FIGURE 3: CD spectra of the CRABP I turn fragments.

rOes suggest that a native conformation is populated, but that it is in conformational exchange with more extended conformations. The strong NN rOe seen between residues Gly78 and Arg79, while not expected in a normal type I turn, is the result of the entry of Gly78 into α_L space in the β-bulge. This bulge results in a separation of 3.0 Å between the Gly78 and Arg79 amide protons in the CRABP I crystal structure. A series of strong αN rOes in the N-terminus and between R79 and K80 at the C-terminus of the fragment suggest extended regions, while their diminution through the center of the fragment implies a turn conformation (Figure 2).

A significant population of native turn is also supported by the amide temperature coefficient analysis of the CRABP I turn IV fragment. As a result of the glycine β-bulge, the native conformation of turn IV in CRABP I includes both Gly78 and Arg79 as hydrogen bond donors in a bifurcated hydrogen bond to the Thr75 carbonyl oxygen. A strongly decreased amide shift with temperature is seen for residue Arg79 (Figure 2, 2.9 ppb/°C), while a slight decrease is seen for Gly78 (4.7 ppb/°C). These parameters argue that the native bifurcated hydrogen bond with both Gly78 and Arg79 as donors is populated in this fragment. As was seen in the turn III fragment, the analysis of Andersen et al. (*42*) shows a clear correlation between $\Delta\delta/\Delta T$ and $\Delta\delta$rc NH indicating that the turn IV fragment is biased toward a single conformation.

The sequential rOe pattern and numerous cross-strand rOes (Figure 4) and the temperature coefficient results are clear indicators of a strong native bias in the conformational ensemble of the CRABP I turn IV fragment. Therefore, the turn population of this fragment was calculated as in the case of turn III. The deviations from random coil values for the turn IV fragment turn residues, ^{76}VDG^{78}, were −0.07, −0.17, and 0 ppm, respectively. Using the Hα chemical shifts from CRABP I as a reference for the 100% folded state (^{76}VDG^{78}, −0.46, −0.29, and 0.07 ppm, respectively) resulted in an estimate of 25% folded. Using the average ^{1}Hα chemical shifts derived from proteins of known structure (*34*) as a reference, the folded population is estimated to be 49%.

Table 2: CRABP I Turn Conservation

turn	sequence	conservation[a] (%)	turn	sequence	conservation[a] (%)
I	^{46}DGDQ^{49}	35	V	^{89}NENK^{92}	40
II	^{56}TTVR^{59}	38	VII	^{114}ANDE^{117}	25
III	^{66}KVGE^{69}	54	VIII	^{124}ADDV^{127}	36
IV	^{75}TVDGR^{79}	72			

[a] Among the members of the iLBP family, the overall average is 45% (see Materials and Methods). The following sequences were used: RET3_BOVIN, RET3_CHICK, RET3_HUMAN, RET4_HUMAN, RET4_RAT, FABE_HUMAN, FABE_MOUSE, RET2_MOUSE, RET2_RAT, FABP_ECHGR, FABP_SCHMA, RET2_HUMAN, RET2_PIG, FABH_MOUSE, RET3_XENLA, RET4_MOUSE, FABH_HUMAN, FABH_PIG, FABH_RAT, FABL_GINCI, FABH_ONCMY, MYP2_BOVIN, MYP2_HUMAN, MYP2_RABBIT, RET1_MOUSE, RET1_RAT, TLBP_MOUSE, TLBP_RAT, FABB_MOUSE, FABH_BOVIN, MYP2_MOUSE, RET1_HUMAN, FABA_HUMAN, FABB_CHICK, FABB_RAT, FABM_SCHGR, FABA_BOVIN, FABA_RAT, FABB_HUMAN, FABA_MOUSE, FABE_BOVIN, FABM_LOCMI, FABP_ANOGA, FABL_CHICK, FABL_HUMAN, FABL_PIG, FABL_RHASA, FAB1_AMBME, FABE_RAT, FABI_HUMAN, FABI_XENLA, and FABI_MOUSE.

Extending this analysis to the *i* and *i* + 5 residues results in calculated populations of 15 and 25% using the CRABP I and average shifts, respectively. As in the turn III fragment, these values present reasonable upper and lower bounds to the turn population. The large number of nonsequential rOes consistent with native conformation, the large reduction of the Arg79 amide temperature coefficient, and the amide temperature coefficient versus amide chemical shift analysis suggest that the population of this fragment is likely to be nearer the higher end of these estimates (25−49%). In total, the data demonstrate that both the turn III and turn IV hairpins represent excellent examples of locally defined turns (Figure 4).

Turns III and IV Are the Most Conserved Turns in the iLBP Family. The presence of any folding information encoded in the primary sequence of turns in CRABP I raises the question of whether this folding information might be evolutionarily conserved in the iLBP family. We searched for a correlation between sequence conservation in the turns of the iLBP family and any native conformational bias identified in the CRABP I turn fragments. The framework folding models posit the presence of native bias in secondary and supersecondary structural elements. Hence, the analysis of sequence conservation was undertaken not on an individual residue basis but on units of secondary structure, specifically the turns. There are sequences available with levels of sequence homology as low as 18% with respect to CRABP I; however, 52 sequences with a level of homology of greater than 30% were used in the analysis (Table 2; aligned sequences are available in the Supporting Information). The large size of the iLBP family makes the results statistically significant.

This analysis demonstrates that turns III and IV are the most conserved turns in the iLBP family (Table 2). The average secondary structural sequence conservation for the iLBP family was 45%. Turn III is 54% conserved, while turn IV is 72% conserved. These two turns were the only ones to show greater than average levels of sequence conservation. Additionally, turn IV is the most conserved secondary structural element in CRABP I (*29*). The results of this analysis of the iLBP family suggest that secondary

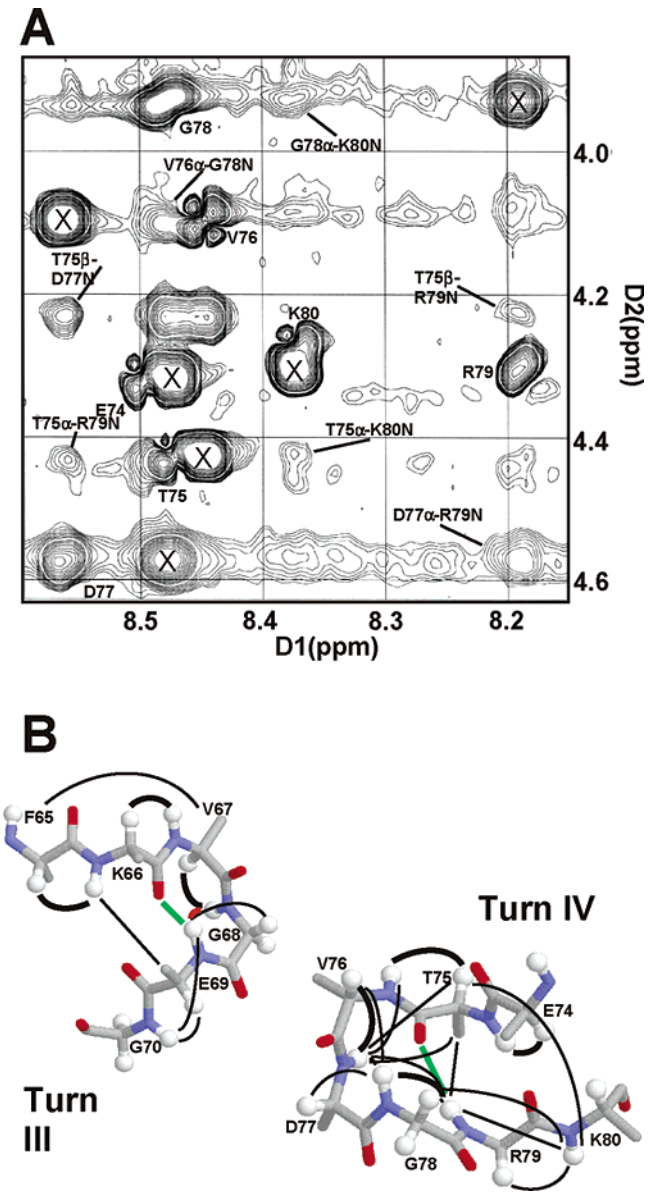

FIGURE 4: (A) Fingerprint region of the ROESY spectrum of turn IV. Nonsequential rOes are labeled, while sequential rOes are marked with an X. (B) Indicators of structure in peptides corresponding to turns III and IV. rOes are shown in black, and the thickness of the line correlates with rOe intensity. Hydrogen bonds are shown as green lines. The αN(*i*,*i*+2) rOe between Asp77 and Arg79 is omitted for clarity.

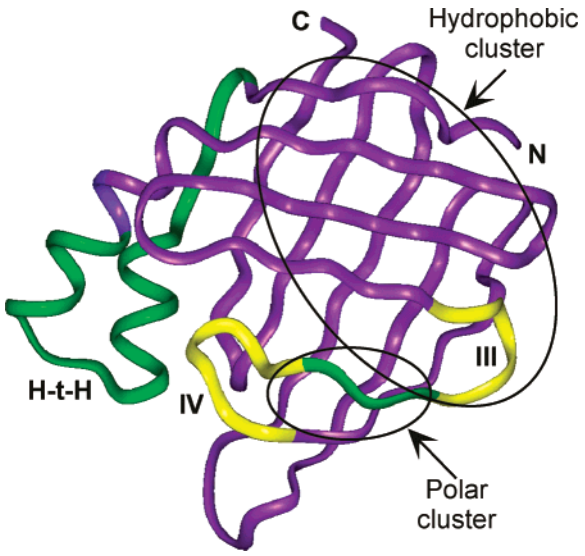

FIGURE 5: CRABP I structure with those regions possessing local sequence-encoded conformational bias color-coded. The turns identified in this work are shown in yellow. Regions previously identified are shown in green and include the helix−turn−helix region (H-t-H) containing the Schellman motif (*41*) and the strand linking turns III and IV (*29*) [figure prepared from PDB entry 1cbi (*23*) using InsightII].

structural conservation is a good predictor of native bias in that sequence.

DISCUSSION

A central puzzle in the protein folding field is the extent to which the conformational distribution in the unfolded ensemble is biased toward the native state. Recent studies have demonstrated not only local structural motifs that are present in unfolded ensembles but also residual structure that involves some longer-range interactions (*48*). Clearly, for protein sequences, the "random coil" is a specific state for a given sequence (*36*), and its character will directly influence the mechanism of folding to the native state, particularly the early events. Yet obtaining descriptions of the unfolded state presents many technical challenges. In this work, we have sought to understand the impact of local sequence on the conformational behavior of the predominantly β-sheet protein, CRABP I. Our finding that isolated peptides corresponding to two of the turns in CRABP I have a strong tendency to fold into their native structures (this study) as does the strand joining them (*29*) adds to our previous work which showed that the helix−turn−helix and Schellman motif of this protein is strongly locally encoded (Figure 5) (*49*). Together, these results paint a picture of a significantly narrowed unfolded ensemble, even without the influence of potential longer-range interactions that may persist in the full-length unfolded chain.

Each of the regions of strong local sequence bias in CRABP I plays a key role in defining its structure. The turn III hairpin is unique to the iLBP family of proteins and is characterized as a diverging hairpin without a cross-strand hydrogen bonding network, with the exception of the turn hydrogen bond (*50*). The presence of a locally defined turn in this hairpin suggests that the turn could be serving to set the register of the hairpin in the absence of more intimate cross-hairpin side chain interactions. In a recent study, we examined the 52 iLBP family members to identify interacting pairs of residues in which the nature of the interaction was conserved to be hydrophobic−hydrophobic, polar−polar, or hydrophobic−polar (with a favorable specific interaction) (*50*). This analysis identified a network of conserved hydrophobic interactions in CRABP I. A subset of the interacting residues (10 of 25) participated in at least three conserved pairwise interactions to form a highly cross-linked hydrophobic cluster. Both Phe65 and Val67 in turn III are part of this cluster (*50*). The mutagenesis work from Frieden and co-workers on intestinal fatty acid binding protein (iFABP) provides insight into the role turn III plays in the folding of iLBP family members. These authors found that mutations in turn III (turn DE in their terminology) substantially destabilized the protein and led to slow folding (*51*). It was concluded that interactions of the highly conserved hydrophobic residues in the turn III region are responsible

for initiating folding and stabilizing the folded state, and specifically that Leu64, analogous to Val67 in CRABP I, was a key residue in the native state and in the folding pathway of iFABP (*52*). These results indicate that turn III is important in organizing residues involved in long-range interactions in the iLBP family of proteins and suggest that the folding information encoded in this turn is conserved in the iLBP family. The implied conservation of folding information throughout the iLBP family is enforced by the results of our secondary structure conservation analysis.

Turn III shares strand 5 with turn IV. The turn IV hairpin represents the transition from the front to the back sheet of the β-clamshell, and is thus a key feature of barrel closure. Local bias for native conformation at this turn provides a mechanism for fast and faithful formation of this critical transition. We have shown that turns III and IV, along with the strand 5, continue to display a native conformational bias in a large fragment encompassing both turns (*29*). Formation of turns III and IV with their shared strand biases the interaction of several of the residues in the conserved hydrophobic network (Phe65, Val67, Pro85, and Trp87). This is intriguing, as studies on iFABP implicate turn III as interacting with elements of the turn IV hairpin in early folding events (*52*). Trp82 on the turn IV hairpin in iFABP, analogous to Trp87 in CRABP I, has been shown to be involved in residual structure at high denaturant concentrations (*53*). Additionally, the formation of turn IV and strand 5 may stabilize contacts in a network of conserved polar contacts also found through sequence analysis of the iLBP family [Figure 5 (*50*)]. As with turn III, these results suggest that the native structural bias observed in turn IV is conserved throughout the iLBP family. Our findings that the turn IV secondary structure is the best-conserved element in the iLBP family support this conclusion.

The apparent conservation of secondary structural elements important in folding the iLBP family is in contrast with results from similar studies on smaller proteins. The proteins G B1 and protein L B1 have a similar fold, two hairpins overlaid with an α-helix. Conformational analysis on fragments of protein G B1 indicates that the second hairpin is locally defined (*54*), while similar experiments on protein L B1 suggest that neither hairpin is locally defined (*55*). The results of the study presented here in concert with the those of other iLBP family members suggest that in these larger, primarily β-structured proteins, topological conservation of folding information may be critical to proper folding.

While the data obtained do not speak directly to the roles of the locally encoded structural regions in the folding mechanism of CRABP I, it is plausible that their influence on the unfolded ensemble will favor formation of specific structural domains at early stages in folding. The nature of these structural domains in turn will guide the cooperative formation of structure as folding proceeds. In the context of CRABP I, the other locally encoded structural region, the helix−turn−helix region, favors the proximity of strands 1 and 10, and thus may play a role in ensuring barrel closure. It is remarkable that these two regions of locally encoded structure both appear to prevent intramolecular association of β-strands, which would result in aggregation. Equally intriguing is the fact that these structures are placed on either end of the front sheet of CRABP I (Figure 5). As such, they are poised to function as the metastable elements of structure in a diffusion−collision mechanism to form the front sheet. We have advanced a model of CRABP I folding based upon such an event (*29*). This model suggests that the front sheet is formed early in folding and is supported by our analysis of conserved interactions, which shows that both the highly interacting residues and their interaction partners are biased toward residues on the front sheet. We hypothesize that the 300 μs phase of folding is dominated by the productive interaction of the two identified regions of structure to form the front sheet with its conserved network of hydrophobic interactions. Subsequently, the organization of the back sheet in the 100 ms time phase results in the solution of the topology problem. Finally, the squeezing out of solvent, allowing the intimate backbone interactions necessary for stable hydrogen bond formation, occurs on the 1 s time scale.

The finding of native bias in the CRABP I turn IV fragment serves to underscore an emerging similarity in turns found to be locally encoded. Turn IV in CRABP I is a type I turn with a glycine bulge. There are few native sequence fragments that form stable turns and/or hairpins in aqueous solution (*54*, *56*−*58*). Three of them, hairpins from platelet factor 4, ubiquitin, and protein G B1, possess type I turns with glycine bulges (*54*, *57*, *58*). Turn IV of CRABP I presents another example of this motif in independently encoded β-turns. Turns with this motif have been shown to be less dependent upon cross-strand interactions (*59*), which suggests that incorporating a turn of this type at a site critical in folding results in more robust turn formation. Removal of this feature has been shown to disrupt the ubiquitin turn fragment and have a negative impact on the folding of ubiquitin (*60*). Furthermore, this motif is suggested as an element of "negative design" (*61*), utilized at the edges of β-sheets to prevent intermolecular aggregation. In each case, these independently encoded turns are at the exposed edge of a β-sheet. This suggests a successful strategy of controlling aggregation at exposed β-sheet edges by employing early formation or preformation of turns with this element. In total, these data show that larger regular β-structured proteins can utilize locally encoded structure in folding, and that they do so with strategies similar to those of smaller proteins that fold with two-state kinetics.

We are following up these studies in several fashions. We are investigating the forces available in the local sequences of turns III and IV that are responsible for their native bias. We are examining the effects on the folding kinetics and equilibrium stability of point mutations in the turn III and IV regions of intact CRABP I. Finally, we are initiating exploration of the unfolded ensemble of CRABP I through both denatured state studies and simulations. Any persistence of structure in the denatured ensemble will result in a significant impact on the early folding events. Each of these avenues of research will provide additional insight into the role of turns in the folding of larger primarily β-structured proteins.

ACKNOWLEDGMENT

We thank Linda Rotondi and Ariana Ornelas for assistance with peptide synthesis and purification. We are grateful to Stephen Eyles, Zoya Ignatova, and Joanna F. Swain for critical reading of the manuscript.

SUPPORTING INFORMATION AVAILABLE

$^3J_{\alpha NH}$ and ^{1}H chemical shift data for the CRABP I turn fragments and aligned sequences of the turns in the iLBP family of proteins. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Baldwin, R. L. (1989) How does protein folding get started? *Trends Biochem. Sci. 14*, 291−294.
2. Dill, K. A., Fiebig, K. M., and Chan, H. S. (1993) Cooperativity in protein folding kinetics, *Proc. Natl. Acad. Sci. U.S.A. 90*, 1942−1946.
3. Brockwell, D. J., Smith, D. A., and Radford, S. E. (2000) Protein folding mechanisms: New methods and emerging ideas, *Curr. Opin. Struct. Biol. 10*, 16−25.
4. Daggett, V., and Fersht, A. R. (2003) Is there a unifying mechanism for protein folding? *Trends Biochem. Sci. 28*, 18−25.
5. Dyson, H. J., and Wright, P. E. (1993) Peptide conformation and protein folding, *Curr. Opin. Struct. Biol. 3*, 60−65.
6. Serrano, L. (2000) The relationship between sequence and structure in elementary folding units, *Adv. Protein Chem. 53*, 49−85.
7. Brown, J. E., and Klee, W. A. (1971) Helix-coil transition of the isolated amino terminus of ribonuclease, *Biochemistry 10*, 470−476.
8. Kim, P. S., and Baldwin, R. L. (1984) A helix stop signal in the isolated S-peptide of ribonuclease A, *Nature 307*, 329−334.
9. Dyson, H. J., and Wright, P. E. (2002) Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance, *Adv. Protein Chem. 62*, 311−340.
10. Chakrabartty, A., and Baldwin, R. L. (1995) Stability of α-helices, *Adv. Protein Chem. 46*, 141−176.
11. Karplus, M., and Weaver, D. L. (1994) Protein folding dynamics: The diffusion-collision model and experimental data, *Protein Sci. 3*, 650−669.
12. Taddei, N., Chiti, F., Fiaschi, T., Bucciantini, M., Capanni, C., Stefani, M., Serrano, L., Dobson, C. M., and Ramponi, G. (2000) Stabilisation of α-helices by site-directed mutagenesis reveals the importance of secondary structure in the transition state for acylphosphatase folding, *J. Mol. Biol. 300*, 633−647.
13. Myers, J. K., and Oas, T. G. (2001) Preorganized secondary structure as an important determinant of fast protein folding, *Nat. Struct. Biol. 8*, 552−558.
14. Schwalbe, H., Fiebig, K. M., Buck, M., Jones, J. A., Grimshaw, S. B., Spencer, A., Glaser, S. J., Smith, L. J., and Dobson, C. M. (1997) Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea, *Biochemistry 36*, 8977−8991.
15. Shortle, D., and Ackerman, M. S. (2001) Persistence of native-like topology in a denatured protein in 8M urea, *Science 293*, 487−489.
16. Capaldi, A., and Radford, S. E. (1998) Kinetic studies of β-sheet protein folding, *Curr. Opin. Struct. Biol. 8*, 86−92.
17. Haque, T. S., and Gellman, S. H. (1997) Insights on β-hairpin stability in aqueous solution from peptides with enforced type I′ and type II′ β-turns, *J. Am. Chem. Soc. 119*, 2303−2304.
18. Blanco, F., Ramírez-Alvarado, M., and Serrano, L. (1998) Formation and stability of β-hairpin structures in polypeptides, *Curr. Opin. Struct. Biol. 8*, 107−111.
19. McCallister, E. L., Alm, E., and Baker, D. (2000) Critical role of β-hairpin formation in protein G folding, *Nat. Struct. Biol. 7*, 669−673.
20. Kortemme, T., Kelly, M. J., Kay, L. E., Forman-Kay, J., and Serrano, L. (2000) Similarities between the spectrin SH3 domain denatured state and its folding transition state, *J. Mol. Biol. 297*, 1217−1229.
21. Nauli, S., Kuhlman, B., and Baker, D. (2001) Computer-based redesign of a protein folding pathway, *Nat. Struct. Biol. 8*, 602−605.
22. Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S., and Jones, T. A. (1994) Lipid-binding proteins: a family of fatty acid and retinoid transport proteins, *Adv. Protein Chem. 45*, 89−151.
23. Kleywegt, G. J., Bergfors, T., Senn, H., Le Motte, P., Gsell, B., Shudo, K., and Jones, T. A. (1994) Crystal structures of cellular retinoic acid binding proteins I and II in complex with all-*trans*-retinoic acid and a synthetic retinoid, *Structure 2*, 1241−1258.
24. Clark, P. L., Liu, Z.-P., Zhang, J., and Gierasch, L. M. (1996) Intrinsic tryptophans of CRABPI as probes of structure and folding, *Protein Sci. 5*, 1108−1117.
25. Clark, P. L., Liu, Z.-P., Rizo, J., and Gierasch, L. M. (1997) Cavity formation before stable hydrogen bonding in the folding of a β-clam protein, *Nat. Struct. Biol. 4*, 883−886.
26. Gierasch, L. M., Rotondi, K. S., Gunasekaran, K., Habink, J. A., and Hagler, A. T. (2001) Local and long-range sequence contributions to the folding of a predominantly β-sheet protein, in *Proceedings of the Seventeenth American Peptide Symposium* (Lebl, M., and Houghten, R. A., Eds.) pp 391−393, The American Peptide Society, San Diego.
27. Rose, G. D., Gierasch, L. M., and Smith, J. A. (1985) Turns in peptides and proteins, *Adv. Protein Chem. 37*, 1−109.
28. Sibanda, B. L., Blundell, T. L., and Thornton, J. M. (1989) Conformation of β-hairpins in protein structures: A systematic classification with applications to modeling by homology, electron density fitting and protein engineering, *J. Mol. Biol. 206*, 759−777.
29. Rotondi, K. S., Rotondi, L. F., and Gierasch, L. M. (2003) 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, *Biophys. Chem. 100*, 421−436.
30. Merrifield, R. B. (1963) Solid phase synthesis of linear peptides, *J. Am. Chem. Soc. 85*, 2149−2154.
31. Bothner-By, A. A., Stephens, R. L., Lee, J.-M., Warren, C. D., and Jeanloz, R. W. (1984) The CAMELSPIN-experiment (cross-relaxation appropriate of minimolecules emulated by locked spins), *J. Am. Chem. Soc. 106*, 811−813.
32. Bax, A., and Davis, D. G. (1985) MLEV-17 based two-dimensional homonuclear magnetization transfer spectroscopy, *J. Magn. Reson. 65*, 355−360.
33. Bax, A., and Davis, D. G. (1985) Practical aspects of two-dimensional transverse NOE spectroscopy, *J. Magn. Reson. 63*, 207−213.
34. Williamson, M. P. (1990) Secondary-structure dependent chemical shifts in proteins, *Biopolymers 29*, 1423−1431.
35. Rizo, J., Liu, Z.-P., and Gierasch, L. M. (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.
36. Smith, L. J., Feibig, K. M., Schwalbe, H., and Dobson, C. M. (1996) The concept of a random coil. Residual structures in peptides and denatured proteins, *Folding Des. 1*, R95−R106.
37. Smith, L. J., Bolin, K. A., Schwalbe, H., MacArthur, M. W., Thornton, J. M., and Dobson, C. M. (1996) Analysis of main chain torsion angles in proteins: Prediction of NMR coupling constants for native and random coil conformations, *J. Mol. Biol. 255*, 494−506.
38. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure, *J. Mol. Biol. 222*, 311−333.
39. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy, *Biochemistry 31*, 1647−1651.
40. Cierpicki, T., and Otlewski, J. (2001) Amide proton temperature coefficients as hydrogen bond indicators in proteins, *J. Biomol. NMR 21*, 249−261.
41. Cierpicki, T., Zhukov, I., Byrd, R. A., and Otlewski, J. (2002) Hydrogen bonds in human ubiquitin reflected in temperature coefficients of amide protons, *J. Magn. Reson. 157*, 178−180.
42. Andersen, N. H., Neidigh, J. W., Harris, S. M., Lee, G. M., Liu, Z., and Tong, H. (1997) Extracting information from the temperature gradients of polypeptide NH chemical shifts. 1. The importance of conformational averaging, *J. Am. Chem. Soc. 119*, 8547−8561.
43. Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., and Lerner, R. A. (1988) Folding of immunogenic peptide fragments of proteins in water solution: II. The nascent helix, *J. Mol. Biol. 201*, 201−217.
44. Woody, R. W. (1985) Circular dichroism of peptides, in *The Peptides: Analysis, Synthesis, Biology* (Udenfriend, S., and Meienhofer, J., Eds.) pp 15−114, Harcourt Brace Jovanovich, New York.
45. Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, John Wiley and Sons, New York.

46. Ramírez-Alvarado, M., Blanco, F. J., and Serrano, L. (1996) De novo design and structural analysis of a model β-hairpin peptide system, *Nat. Struct. Biol. 3*, 604−612.
47. Santiveri, C. M., Santoro, J., Rico, M., and Jimenez, M. A. (2002) Thermodynamic analysis of β-hairpin-forming peptides from thermal dependence of ^{1}H NMR chemical shifts, *J. Am. Chem. Soc. 124*, 14903−14909.
48. Shortle, D. (2002) The expanded denatured state: An ensemble of conformations trapped in a locally encoded topological space, *Adv. Protein Chem. 62*, 1−23.
49. Sukumar, M., and Gierasch, L. M. (1997) Local interactions in a Schellman motif dictate interhelical arrangement in a protein fragment, *Folding Des. 2*, 211−222.
50. Gunasekaran, K., Hagler, A. T., and Gierasch, L. M. (2003) Sequence and structural analysis of cellular retinoic acid binding proteins reveals a network of conserved hydrophobic interactions, *Proteins: Struct., Funct., Genet.* (in press).
51. Kim, K., Ramanathan, R., and Frieden, C. (1997) Intestinal fatty acid binding protein: A specific residue in one turn appears to stabilize the native structure and be responsible for slow refolding, *Protein Sci. 6*, 364−372.
52. Kim, K., and Frieden, C. (1998) Turn scanning by site-directed mutagenesis: Application to the protein folding problem using the intestinal fatty acid binding protein, *Protein Sci. 7*, 1821−1828.
53. Ropson, I. J., and 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. U.S.A. 89*, 7222−7226.
54. Blanco, F., and Serrano, L. (1995) Folding of protein G B1 domain studied by the conformational characterization of fragments comprising its secondary structure elements, *Eur. J. Biochem. 230*, 634−649.
55. Ramírez-Alvarado, M., Serrano, L., and Blanco, F. (1997) Conformational analysis of peptides corresponding to all the secondary structure elements of protein L B1 domain: Secondary structure propensities are not conserved in proteins with the same fold, *Protein Sci. 6*, 162−174.
56. Blanco, F. J., Jimenez, M. A., Herranz, J., Rico, M., Santoro, J., and Nieto, J. L. (1993) NMR evidence of a short linear peptide that folds into a β-hairpin in aqueous solution, *J. Am. Chem. Soc. 115*, 5887−5888.
57. Ilyina, E., Milius, R., and Mayo, K. H. (1994) Synthetic peptides probe folding initiation sites in platelet factor-4: Stable chain reversal found within the hydrophobic sequence LIATLKNGRKISL, *Biochemistry 33*, 13436−13444.
58. Searle, M. S., Dudley, H. W., and Packman, L. C. (1995) 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*, 999−1006.
59. de Alba, E., Jiménez, M. A., and Rico, M. (1997) Turn residue sequence determines β-hairpin conformation in designed peptides, *J. Am. Chem. Soc. 119*, 175−183.
60. Chen, P. Y., Gopalacushina, B. G., Yang, C. C., Chan, S. I., and Evans, P. A. (2001) The role of a β-bulge in the folding of the β-hairpin structure in ubiquitin, *Protein Sci. 10*, 2063−2074.
61. Richardson, J. S., and Richardson, D. C. (2002) Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation, *Proc. Natl. Acad. Sci. U.S.A. 99*, 2754−2759.

BI034304K