The search for local native-like nucleation centers in the unfolded state of β-sheet proteins

Gregory V. Nikiforovich and Carl Frieden*

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, MO 63110

An approach involving the systematic computational conformational analysis of all overlapping hexapeptide segments in the protein sequence has found fragments with the higher than average propensity to adopt the native-like three-dimensional structure and other regular nonrandom structures in the unfolded states of four β-sheet proteins, namely IFABP (intestinal fatty acid-binding protein), ILBP (ileal fatty acid-binding protein), CRABP I (cellular retinoic acid-binding protein), and CRBP II (cellular retinal binding protein). The native three-dimensional structures of these four proteins are very similar even though they possess as little as ≈30% sequence similarity. The computational results were validated by comparison with the experimental data of the heteronuclear sequential quantum correlation NMR spectroscopy obtained earlier for IFABP at high urea concentrations. On this basis, a molecular model of the unfolded state of IFABP has been developed. The model presumes a dynamic equilibrium between various nonrandom structures (including the native-like structure) and random coil in the local segments of the protein sequence. The model explains experimental observations obtained earlier for folding of several mutants of IFABP, as well as the observed differences in molecular mechanisms of folding for the four β-sheet proteins. Because the computational approach itself does not employ any experimentally derived information in advance, it is not necessarily limited to the β-sheet proteins.

The protein folding problem has been one of the most discussed in molecular biology during the last several decades. It is quite clear that the process of folding must involve some mechanisms allowing a protein chain to avoid visiting all possible conformational states on its way to the native three-dimensional (3D) structure—otherwise the time of such search will be exceedingly long, the so-called Levinthal paradox [as discussed by Wetlaufer (1)]. One solution that has been suggested to resolve this problem is the hierarchical mechanism of folding. It postulates forming several local nucleation centers corresponding to small segments of the native 3D structure at the very first stages of folding (e.g., refs. 1 and 2). The local nucleation centers are thought to be α-helical fragments, β-strands, or β-turns (3, 4). β-turns may lead to formation of β-hairpins that, in turn, may propagate to form β-sheets, whereas α-helical fragments may subsequently form α-helical bundles. This mechanism predicts possible accumulation of intermediates of the "molten globule" type as shown experimentally for many proteins (see, e.g., ref. 5). Another possible solution to the folding problem is a nucleation-condensation mechanism in which transition from an unfolded state to the folded one occurs directly in one step without intermediates (the "two-state" model; see, e.g., ref. 6). According to this mechanism, the rate of folding depends on general protein topology rather than local interactions. The long-range interactions may be quantified roughly by the "contact order" parameter; the values of this parameter correlate with the rates of folding for 22 two-state proteins (7).

Despite differences in their views on the exact role of the local native-like centers in protein folding, both models assume that the unfolded state may be regarded as a dynamic ensemble of many local 3D structures (8). In fact, such an assumption has been successfully used in theoretical studies aimed at calculating the kinetics of protein folding from the known 3D structures (e.g., refs. 9 and 10). However, the location of the native-like centers in a sequence is rather difficult to determine in the unfolded state by experimental means, such as NMR spectroscopy (e.g., refs. 11–17). Therefore, the present study develops a computational approach focused on finding peptide fragments in the protein sequence, which may serve as the local native-like nucleation centers preexisting in the unfolded/denatured state.

The approach is applied to four β-sheet proteins that have been extensively studied: IFABP (the intestinal fatty acid-binding protein), ILBP (the ileal fatty acid-binding protein), CRABP I (the cellular retinoic acid-binding protein), and CRBP II (the cellular retinal binding protein; refs. 18–23). The native 3D structures of these four proteins are very similar [the PDB entries are 1IFC (IFABP), 1EIO (ILBP), 1CBI (CRABP I), and 1OPA (CRPB II), respectively], although they possess as little as ≈30% sequence similarity (21, 24). The native 3D structure of IFABP is shown schematically in Fig. 1 with the β-strands labeled A–J.

The available experimental data on the four proteins lead to various general conclusions as to their folding mechanisms. Hodsdon and Frieden, using the heteronuclear sequential quantum correlation (HSQC) NMR experiment, have shown that the turns between the D and E β-strands, as well as between the I and J β-strands, transiently exist in the unfolded state of IFABP (23). Earlier data suggested that the initial step in IFABP folding is hydrophobic collapse bringing together residues belonging to the C, D, E, F, and G β-strands (18), a view developed further in a recent publication (22). A somewhat different model has been proposed for ILBP (21). The authors suggest that ILBP is less stable than IFABP and proceeds through formation of a molten globule-like structure rather than through a relatively limited collapse of hydrophobic clusters, as IFABP does. This early intermediate retains practically all of its secondary structure (21), which is not the case for IFABP.

Methods

Computational Approach to Finding Local Native-Like Centers in the Protein Sequence. The approach itself consists of systematic conformational analysis of all overlapping hexapeptide segments in the protein sequence. The hexapeptide fragments were selected as the basic units for conformational studies mainly for two reasons. First, the length of six residues is sufficient to distinguish between most of the expected regular 3D structures (β-strand, α-helix, or β-turn) at the chosen criterion of geometrical similarity of the rms value equal to 2 Å (C^α atoms only). Second, complete conformational sampling of hexapeptides is still readily attainable using available computational resources.

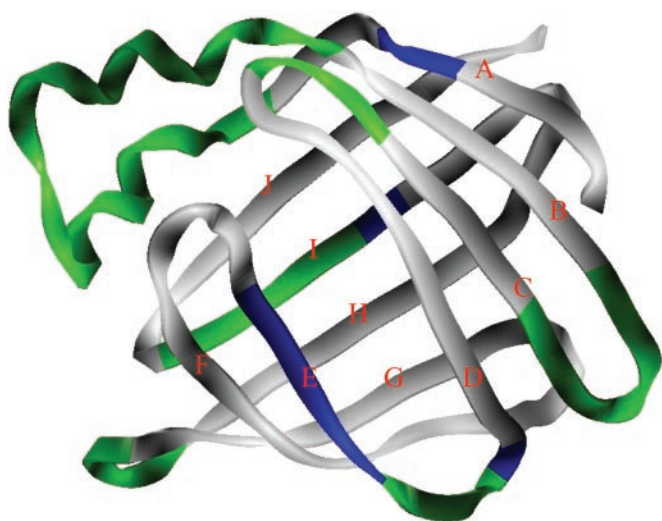For a given hexapeptide in the protein sequence, it is logical

---

Fig. 1. Schematic representation of the native 3D structure of IFABP (PDB entry 1IFC). β-strands are labeled. The green color marks the predicted native-like nucleation centers and the blue color marks the predicted near-native-like segments found for IFABP (see text).

to define an inherent propensity to adopt the native-like 3D structure as a ratio between the number of low-energy conformers geometrically similar to the native structure (which is known) and the total number of low-energy conformers. Generally, such an approach tacitly assumes that a computational procedure always finds the correct sets of low-energy conformers. In reality, this is not the case, mainly because of systematic errors of modeling very diverse interatomic interactions in peptides by the uniformly parameterized force fields (see, e.g., ref. 25). To alleviate this problem, some preconditions may be introduced. First, narrow sets of low-energy conformers (those with less than 20 conformers) are not considered because they may represent possible artifacts of the computational procedure. Second, the ratios *normalized* by the average ratio for all hexapeptides comprising the sequence in question are regarded as measures of *relative* propensities to adopt the native-like structure for each hexapeptide rather than the absolute ratios outlined above. Accordingly, if for any hexapeptide the *relative* propensity value is larger than 1, it indicates that a given hexapeptide possesses an above average propensity to adopt the native-like structure. Third, to evaluate the relative propensity of each residue to adopt low-energy conformation compatible to the native conformation of the protein, standard smoothing of the propensity values within the six-residue window can be performed.

Obviously, the above approach can find not only segments with the higher than average propensity to adopt the native-like 3D structures, but segments with high propensity to adopt any other type of nonrandom structure (β-strands, α-helices, or β-turns) as well. These segments may overlap in the protein sequence, which will point out the existence of dynamic equilibrium between various nonrandom structures in the overlapping fragments. Therefore, the approach may, to some extent, account for dynamic nature of the possible local nucleation centers present in the unfolded state.

Conformational Energy Calculations. Energy calculations for all linear hexapeptides were performed by employing the ECEPP/2 potential field (26, 27) assuming rigid valence geometry with planar trans-peptide bonds (including those in proline residues). In all cases, hexapeptides were considered acetylated at the N-terminal and N-methylamidated at the C-terminal. Aliphatic and aromatic hydrogens were generally included in the unified

atomic centers of CH_n type; H^α-atoms and amide hydrogens, as well as H^δ-atoms in prolines were described explicitly. All calculations were performed with the value of the dielectric constant $\varepsilon = 80$ (the macroscopic ε value for water) to mimic, to some extent, the effect of water solution. The starting points for energy calculations were all possible combinations of local energetic minima in the Ramachandran map selected to cover all conformational possibilities of peptide backbone (also see ref. 25). Specifically, we have selected the local energy minima with the dihedral angle ϕ, ψ values of $-140°$, $140°$; $-75°$, $140°$; $-75°$, $80°$; $-60°$, $-60°$ and $60°$, $60°$ for all nonglycine and nonproline residues with an addition of the ϕ, ψ values of $140°$, $-140°$; $75°$, $-140°$ and $75°$, $-80°$ for glycines; for prolines, the ϕ, ψ values of $-75°$, $140°$; $-75°$, $80°$ and $-75°$, $-60°$ were selected. The side chain dihedral angle values were optimized before energy minimization to achieve their most favorable spatial arrangements, employing an algorithm described previously (28). In total, the number of different conformations of peptide backbone under consideration for each hexapeptide was between 15,000 and 40,000. Low-energy conformers of peptide backbone were selected after energy minimization for each hexapeptide (the conformers with relative energies $\Delta E = E - E_{min} \leq 6$ kcal/mol; see ref. 25 for justifying the criterion of 1 kcal/mol per residue).

Results and Discussion

Possible Local Nucleation Centers in the Unfolded/Denatured State of β-Sheet Proteins. The computational procedure found all low-energy conformers of peptide backbones for all of the 126 hexapeptide fragments that together comprise the entire IFABP sequence. The numbers of low-energy conformers per fragment varied from 4 (fragment 113–118) to 407 (fragment 19–24) with the average value of ≈87. For each hexapeptide, the number of low-energy conformations geometrically similar to the 3D structure of the corresponding hexapeptide fragment in the known x-ray structure of IFABP (PDB entry 1IFC) has been determined.

The resulting values of the relative propensities for residues 6–126 of IFABP found to be compatible with the x-ray structure are depicted in Fig. 2. There are several continuous regions of the IFABP sequence showing the relative propensities above average, namely the fragments 12–35, 42–47, 52–54, 65–67, 97–102, and 116–119. They roughly correspond to the N-terminal α–α turn, and the β-turns BC, CD, DE, GH, and IJ in the x-ray structure of IFABP. Fig. 3 shows these results along with those obtained for the three other proteins under consideration. For IFABP these segments are also shown in green in Figs. 1 and 4.

The computational procedure also found IFABP segments and separate residues with the higher than average propensities to adopt the β-strand structures (6–9, 53–54, 65–71, 114–121), α-helical structures (6, 20–31, 33–39, 43–48, 59–61, 72–77, 79–82, 87–94, 99–107, 112–114), and the possible centers for β-turns (12–13, 16–17, 32–33, 43–44, 64–65, 66–67, 99–100, 121–122). Some of these segments overlap with each other and with the segments of the native-like structure. Accordingly, Fig. 3 also displays the residues within the native-like fragments that may exist in dynamic equilibrium between the native-like structure and any of the other regular structures contradictive to the native-like structure. There are also several segments outside the native-like fragments, which display higher than average propensities to adopt only one single type of the regular structure. Some of these segments may not be native-like according to the chosen criterion for geometrical similarity (rms < 2.0 Å), but may be considered as near-native-like by comparison to definitions of the regular structures in the corresponding PDB entries. They are colored blue in Figs. 1 and 4. The segments that show higher than average propensities to adopt the single type of the regular structure different from those defined in the correspond-
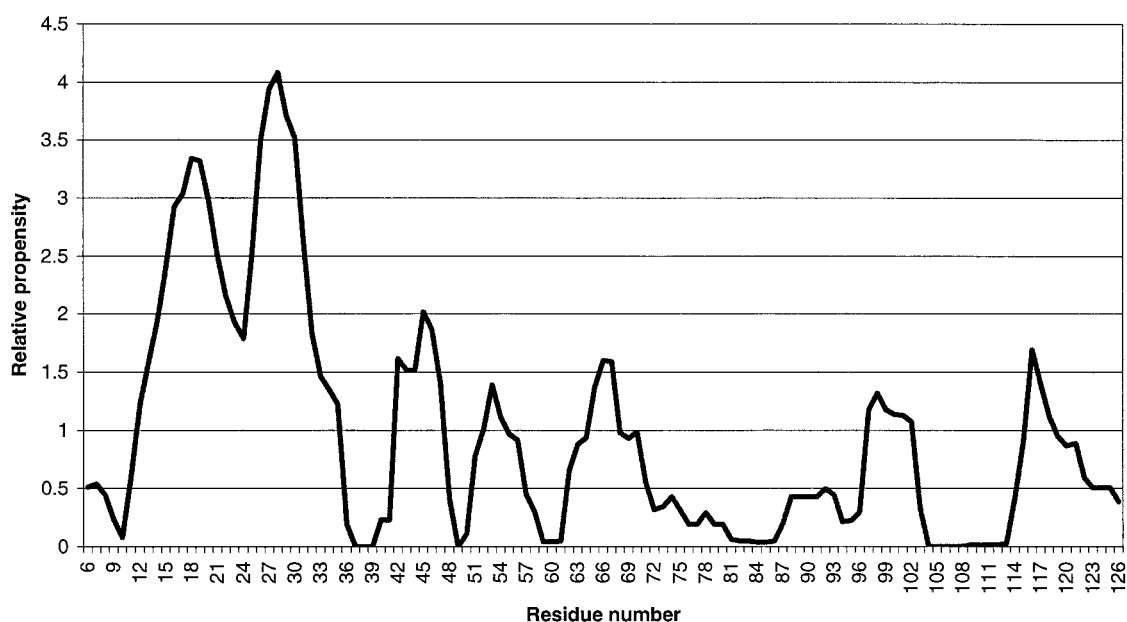
Fig. 2. Profile of relative propensities to adopt the native-like 3D structure for residues in IFABP.

ing PDB entries are colored red in Fig. 4; these segments represent the regular non-native structures in the unfolded state of IFABP. Interestingly, the AGADIR program, which estimates propensities of helix–coil transitions in peptides (Internet access is kindly provided by the Serrano group at www.embl-heidelberg.de/Services/serrano/agadir/agadir-start.html), also suggests at least two non-native helical regions at 100–108 and 120–129.

We suggest that IFABP segments with higher than average propensities to adopt the native-like 3D structure, as well as the other nonrandom structures, may be regarded as the local nucleation centers that transiently exist in the unfolded/denatured state of the protein in rapid equilibrium with random coil and with each other. The experimental data of the heteronuclear sequential quantum correlation (HSQC) NMR spectroscopy obtained for IFABP at various concentrations of urea (a denaturing agent) demonstrate that the ^{1}H-^{15}N cross-peaks reliably assigned to several residues retain at least 0.01 of the intensities of the native state in 6.0 M urea, when the unfolding process is virtually completed. Indeed, some peaks can be observed even at 6.5 M urea (23). This may indicate that local environment of these particular residues changes only slightly

during unfolding—i.e., that the local native 3D structure in the regions of these residues may transiently remain the same in the unfolded state. These residues in the fragment 6–126 are 7–9, 17, 22–23, 42, 50, 63–65, 67, 73, 95, 120–121, and 125. Fig. 3 shows that 14 of these 18 residues are located either exactly within the native-like or the near-native-like centers proposed by the computational procedure, or just next to them. The exceptions are residues 50, 73, 95, and 125. This agreement with experimental data validates both the computational procedure itself and the assumption that the predicted local native-like centers adopt the 3D structures compatible to the native structure of IFABP in the unfolded/denatured state. It agrees also with the molecular model of IFABP folding in which the nascent β-turns BC, CD, DE, GH, and IJ exist at the early steps. This model is similar to that proposed earlier as the result of experimental studies (23).

Local Native-Like Nucleation Centers at Further Stages of Folding. Several independent experimental studies suggest that folding of IFABP is a multistep process with distinct intermediates. Our hypothesis is that the local nucleation centers preexist in dynamic equilibrium in the unfolded state and/or at the very first phase

Fig. 3. Sequences of IFABP, ILBP, CRBP II, and CRABP I aligned according to their 3D native structures by the VAST procedure (www.ncbi.nlm.nih.gov/Structure/VAST/vast.search.html). Approximate positions of α-helical fragments and β-turns are shown below the sequences. Numbers flanking the sequences correspond to their actual residue numbering. The native-like nucleation centers are shown in green and the near-native-like nucleation centers (IFABP only) are shown in blue. Residues within the centers that may exist in dynamic equilibrium between the native-like structure and any of the other types of nonrandom structures (as β-strands, α-helices, and β-turns) are shown in magenta. Residues retaining NMR cross-peak intensities at 6.0 urea in IFABP are underlined.
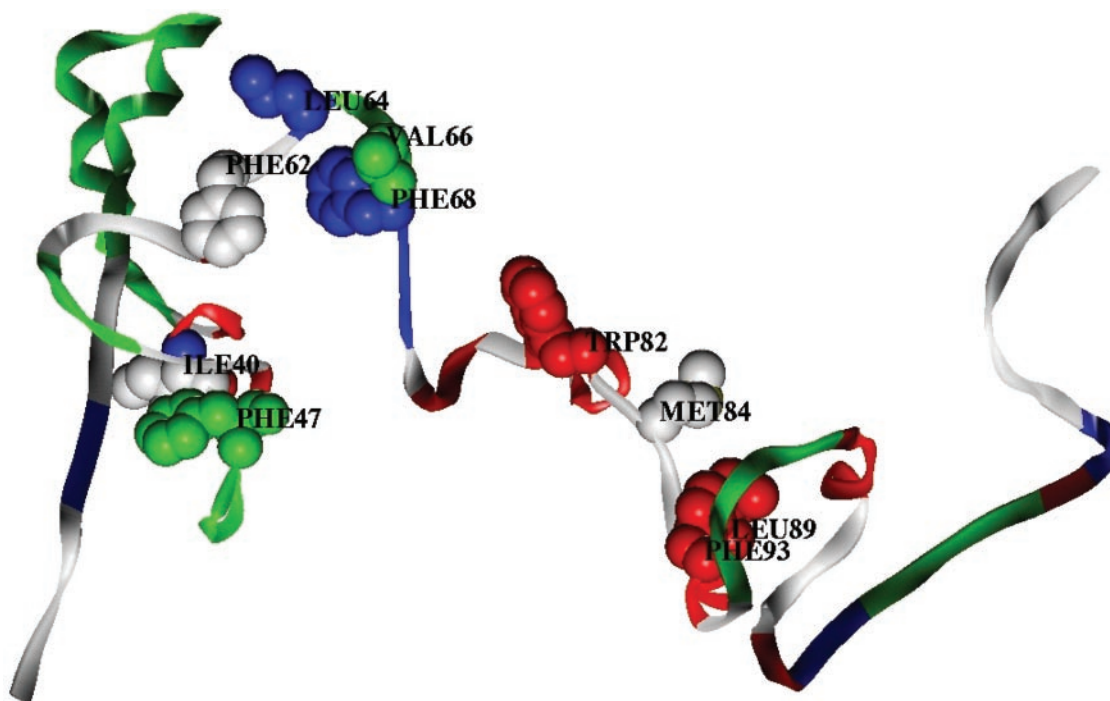
Fig. 4. Representation of transient structures in the unfolded state of IFABP. Residues mentioned in the text are shown as space-filled models. Regions that are defined as the native-like and the near-native-like are shown in green and in blue, respectively. The non-native regular structures are shown in red (see text for details).

of folding characterized by very fast kinetics ($\approx$10,000 s^{-1}; refs. 22 and 29). We also believe that our computational results may explain some experimental data obtained for the slower phases of folding ($\approx$1 s^{-1}). These data include results of measurements performed on the series of IFABP mutants where glycine residues in β-turns have been replaced by more conformationally rigid valines (20). Fig. 5 shows a comparison of the profiles of the relative propensities to adopt the native structure for IFABP and the "composite" profile for mutants G44V, G65V, G75V, G86V, G99V, G110V, and G121V [assuming that the native structure for mutants is the same as for IFABP (20)]. It is obvious that there are rather small differences between IFABP and mutants G44V, G65V, G75V, and G86V. At the same time, the local nucleation centers corresponding to the β-turns GH and IJ are different in the mutants G99V and G121V, respectively, compared with IFABP, and there is a new nucleation center in the region of the HI β-turn in the G110V mutant.

Fluorescence studies that measured rate constants on the order of 2,000 s^{-1} for refolding revealed negligible differences between these mutants (as well as between a series of the G/P mutants; ref. 29), but stopped flow measurements (the slower phases of folding) showed that whereas mutants G44V, G65V, G75V, and G86V refold relatively fast (at 10–100 s^{-1} scale), mutants G99V, G110V, and G121V refold 10 to 100 times slowly (20). This may mean that the "proper" folding pathway of IFABP in mutants G99V, G110V, and G121V is disturbed, because the starting unfolded state in these mutants is different. Thus, either location of local nucleation centers represented in IFABP is shifted (mutants G99V and G121V), or new nucleation centers (mutant G110V) emerge. That may be the reason for slow restructuring toward the native structure at the further stages of folding.

Molecular Models of the Unfolded State of IFABP and of the Early Stages of IFABP Folding. In earlier studies, Kim *et al.* (19) carried out random mutagenesis in the DE turn (residues 64–66)

yielding several different mutants. They concluded that one residue, Leu-64, was critical for stabilizing the final structure. Our computational procedure found that the profiles of the relative propensities to adopt the native structure for the wild-type IFABP and for four mutants in the DE turn, namely L64/A65/F66, L64/D65/H66, S64/A65/N66, and G64/A65/F66, are essentially the same (data not shown). However, the former two mutants refold at the same rate as the wild-type IFABP, whereas the two latter ones are significantly slower (19). Therefore, although the same nucleation center in the region of the DE β-turn may preexist in all four mutants in the unfolded state, the very presence of the highly hydrophobic Leu-64 side chain may be required for stabilizing of the native structure at the later stages of folding. This conclusion is essentially the same as previously proposed (19), where nonlocal hydrophobic interactions of Leu-64 with other hydrophobic residues in IFABP have been suggested for the final stages of folding. This finding emphasizes that the entire process of IFABP folding depends on contributions from both the local native-like centers and the nonlocal hydrophobic interactions.

Experimental data suggest that simultaneous hydrophobic collapse of Leu-64 with Trp-82, Met-84, and Leu-89 (the region of the FG turn), as well as with Phe-62, Phe-68 (the DE turn), and Phe-47 (the C leg of the BC β-turn) stabilizes the hydrophobic core of IFABP at the first (22) or final (19, 23) stages of folding. According to the results of this study, it seems more likely that the native-like nucleation centers around the BC, CD, and DE turns, which preexist in the unfolded state, may first initiate formation of missing parts of the first β-sheet (BCDE). This process would be promoted by hydrophobic interactions in the two clusters, Leu-64, Phe-62, Phe-68, and Val-66, and Ile-40 and Phe-47, which then would combine to form a larger hydrophobic cluster; in turn, the random and small non-native segments in region 30–70 would reorganize to the native-like structure of the BCDE β-sheet (see Fig. 4). Simultaneously, but perhaps more slowly (because relatively large non-native seg-
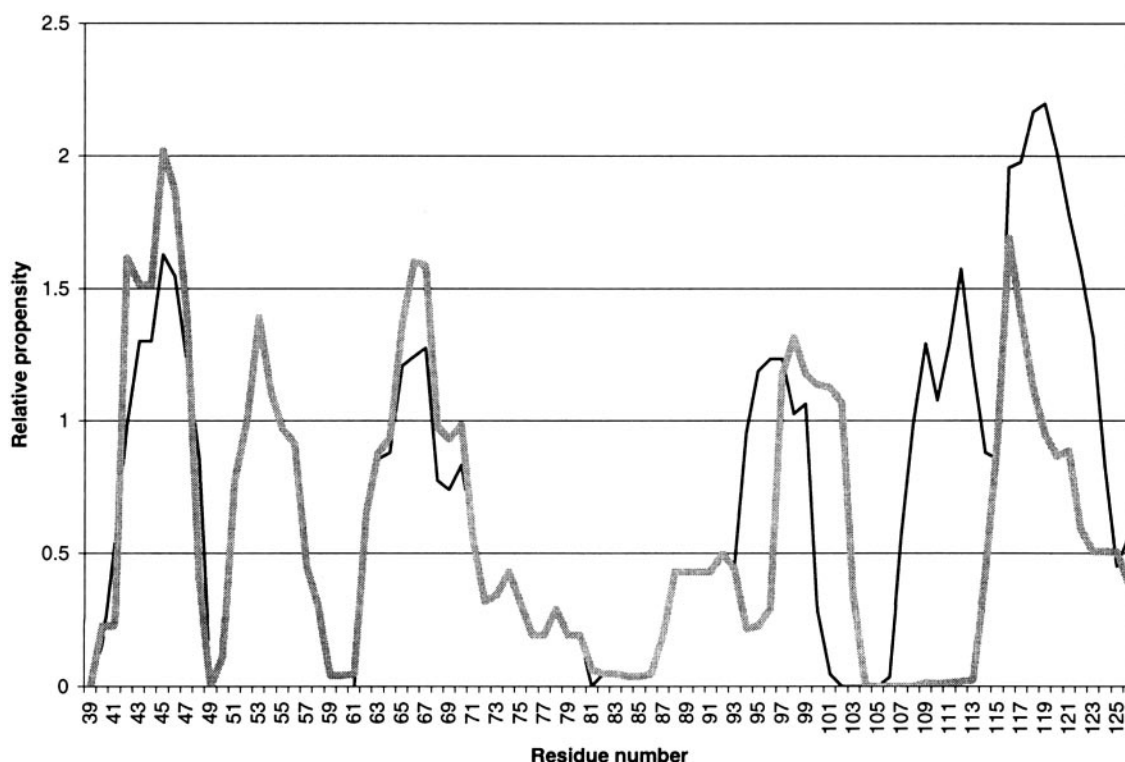
Fig. 5. Comparison of relative propensities to adopt the native-like structure of IFABP in the unfolded/denatured state for IFABP itself (thick gray line) and its G/V mutants (thin black line).

ments should be restructured in the process), hydrophobic residues located around the emerging FG β-turn (Trp-82, Met-84, Leu-89, and Phe-93) would come together to form the another hydrophobic cluster (see Fig. 4), which may collapse with the BCDE β-sheet at the next state of folding.

It is noteworthy that the peak maximum in Trp-82 fluorescence corresponding to the fully folded state is achieved gradually and not at the very early steps of folding (22, 29). This experimental observation agrees with our model in that a local nucleation center in the FG β-turn region may not form initially. At the same time, the model does not exclude stabilization of the entire hydrophobic core in IFABP (involving Trp-82) after the BCDE β-sheet is formed. The earlier work on ^{19}F NMR spectroscopy of selectively labeled IFABP showed an intermediate persisting at high concentrations of denaturant (i.e., at the later stages of unfolding/early stages of folding) and involving Trp-82 (18). It was hypothesized that Trp-82 would participate in the initial hydrophobic collapse with most of the above-mentioned residues, but no direct experimental evidence has been provided. From our point of view, a local native-like nucleation center involving the FG turn may not exist in the unfolded state, but rather may be formed at a later stage of folding.

Differences in Folding Mechanisms for the Four β-Sheet Proteins. As one can see from Fig. 3 (second row), seven local native-like nucleation centers, shown in green, may be identified for ILBP roughly corresponding to the α–α turn, and the β-turns BC, CD, EF, FG, GH, and IJ. The nascent β-turns may rapidly develop into six full-fledged β-hairpins, forming an intermediate that may show CD spectra very similar to those of the native structure

(eight β-hairpins total) in agreement with experimental observations (21). However, the local nucleation center in the region of the DE β-turn present in IFABP is notably absent in ILBP, ruling out early formation of a hydrophobic interaction that may be important for stabilization of the native structure (see discussion above and in ref. 24). This finding may explain why ILBP has been found to be less stable than IFABP (21).

For CRBP II, the results in Fig. 3 (third row) also indicate six distinct native-like nucleation centers in the β-turn regions (shown in green), but they are slightly different from those found for ILBP, namely BC, EF, FG, GH, HI, and IJ. Again, the important local nucleation center in the DE β-turn is absent, suggesting different folding/unfolding pathways between CRBP II and IFABP. In the case of CRABP I (last row of Fig. 3), only four distinct native-like nucleation centers in the β-turn regions were found, namely those for the BC, CD, DE, and EF β-turns. This finding may explain why CRBP II refolds 100 times faster than CRABP I (24). Interestingly, the values of the contact order parameter (7) for the all four β-sheet proteins should be close, because the topologies of these proteins are virtually identical; therefore, the contact order parameter cannot account for the observed two-order difference in the folding rates. It implies that the computational approach developed in this study may suggest sites for new mutations to produce protein mutants with predictable changes in folding/unfolding molecular mechanisms, including changes in the folding/unfolding rates. Because the approach does not employ any experimentally derived information in advance, it is not necessarily limited to β-sheet proteins.

1. Wetlaufer, D. B. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 697–701.
2. Ptitsyn, O. B. (1973) *Dokl. Acad. Nauk SSSR* **210,** 1213–1215.
3. Baldwin, R. L. & Rose, G. D. (1999) *Trends Biochem. Sci.* **24,** 26–33.
4. Baldwin, R. L. & Rose, G. D. (1999) *Trends Biochem. Sci.* **24,** 77–83.
5. Ptitsyn, O. B. (1995) *Curr. Opin. Struct. Biol.* **5,** 74–78.
6. Fersht, A. R. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 1525–1529.

7. Plaxco, K. W., Simons, K. T. & Baker, D. (1998) *J. Mol. Biol.* **277,** 985–994.
8. Plaxco, K. W. & Gross, M. (2001) *Nat. Struct. Biol.* **8,** 659–660.
9. Munoz, V. & Eaton, W. A. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11311–11316.
10. Galzitskaya, O. V. & Finkelstein, A. V. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11299–11304.
11. Arcus, V. L., Vuilleumier, S., Freund, S. M. V., Bycroft, M. & Fersht, A. R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 9412–9416.
12. Freund, S. M. V., Wong, K.-B. & Fersht, A. R. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 10600–10603.
13. Bai, Y., Chung, J., Dyson, H. J. & Wright, P. E. (2001) *Protein Sci.* **10,** 1056–1066.
14. Garcia, P., Serrano, L., Durand, D., Rico, M. & Bruix, M. (2001) *Protein Sci.* **10,** 1100–1112.
15. Kortemme, T., Kelly, M. J. S., Kay, L. E., Forman-Kay, J. & Serrano, L. (2000) *J. Mol. Biol.* **297,** 1217–1229.
16. Yao, J., Chung, J., Eliezer, D., Wright, P. E. & Dyson, H. J. (2001) *Biochemistry* **40,** 3561–3571.
17. Yi, Q., Scalley-Kim, M. L., Alm, E. J. & Baker, D. (2000) *J. Mol. Biol.* **299,** 1341–1351.
18. Ropson, I. J. & Frieden, C. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 7222–7226.
19. Kim, K., Ramanathan, R. & Frieden, C. (1997) *Protein Sci.* **6,** 364–372.
20. Kim, K. & Frieden, C. (1998) *Protein Sci.* **7,** 1821–1828.
21. Dalessio, P. M. & Ropson, I. J. (2000) *Biochemistry* **39,** 860–871.
22. Yeh, S.-R., Ropson, I. J. & Rousseau, D. L. (2001) *Biochemistry* **40,** 4205–4210.
23. Hodsdon, M. E. & Frieden, C. (2001) *Biochemistry* **40,** 732–742.
24. Burns, L. L., Dalessio, P. M. & Ropson, I. J. (1998) *Proteins Struct. Funct. Genet.* **33,** 107–118.
25. Nikiforovich, G. V. (1994) *Int. J. Peptide Protein Res.* **44,** 513–531.
26. Dunfield, L. G., Burgess, A. W. & Scheraga, H. A. (1978) *J. Phys. Chem.* **82,** 2609–2616.
27. Nemethy, G., Pottle, M. S. & Scheraga, H. A. (1983) *J. Phys. Chem.* **87,** 1883–1887.
28. Nikiforovich, G. V., Hruby, V. J., Prakash, O. & Gehrig, C. A. (1991) *Biopolymers* **31,** 941–955.
29. Chattopadhyay, K., Zhong, S., Yeh, S.-R., Rousseau, D. L. & Frieden, C. (2002) *Biochemistry* **41,** 4040–4047.