Biochemistry. In the article “Characterization of residual structure in the thermally denatured state of barnase by simulation and experiment: Description of the folding pathway” by Chris J. Bond, Kam-Bo Wong, Jane Clarke, Alan R. Fersht, and Valerie Daggett, which appeared in number 25, December 9, 1997, of Proc. Natl. Acad. Sci. USA (94, 13409–13413), one of the authors regrets that she inadvertently omitted references to the computer program and protein potential function that the authors used for their simulations of barnase cited above. The following sentence should have been the first sentence of the Methods section: Molecular dynamics simulations were performed with the program ENCAD (44) and the potential energy function of Levitt et al. (45).


Biochemistry. In the article “Negative control of bacterial DNA replication by a cell cycle regulatory protein that binds at the chromosome origin” by Kim C. Quon, Bing Yang, Ibrahim J. Domian, Lucy Shapiro, and Gregory T. Marczynski, which appeared in number 1, January 6, 1998, of Proc. Natl. Acad. Sci. USA (95, 120–125), the authors wish to note that the institutional affiliations in the author line were incorrectly attributed. The correct affiliations are as follows. Bing Yang and Gregory T. Marczynski are at McGill University in Montreal, and Kim C. Quon is now at the Netherlands Cancer Institute.

Biochemistry. In the article “The 30-kDa C-terminal domain of the RecB protein is critical for the nuclease activity, but not the helicase activity, of the RecB/RecD enzyme from Escherichia coli” by Misook Yu, Jehanne Souaya, and Douglas A. Julin, which appeared in number 3, February 3, 1998, of Proc. Natl. Acad. Sci. USA (95, 981–986), the following correction should be noted. The symbols in the graph (Fig. 3C) were identified incorrectly in the manuscript. The corrected legend and graph with accompanying symbols are printed below.

![Graph](image)

**Fig. 3.** Nuclease assays. All reaction mixtures contained buffer A and 10 mM MgCl₂. (A) ssDNA exonuclease assay with 100 nM single-stranded linear 5'-32P-labeled 25-mer as the substrate. (B) dsDNA exonuclease assay using 6 nM single-stranded circular M13 phage DNA as the substrate. (C) dsDNA exonuclease assay using 6 nM single-stranded linear 5'-32P-labeled 25-mer as the substrate. Each reaction contained 2.3 μM (nucleotides) double-stranded [3H]pTZ19R, 250 μM ATP, and 9 nM RecB1-92CD, 0.2 nM RecB/ChD, or 9 nM RecB/ChD.

Biochemistry. In the article “Escherichia coli RNA polymerase terminates transcription efficiently at rho-independent terminators on single-stranded DNA templates” by Susan M. Uptain and Michael J. Chamberlin, which appeared in number 25, December 9, 1997, of Proc. Natl. Acad. Sci. USA (94, 13548–13553), the authors request that the following correction be noted. It is critical that the bands in lanes 6 and 8 of Fig. 3 indicated by the T7Te arrow be visible. The existence of these terminated bands is a major point on which the conclusions of the paper depend. Therefore, to enhance their visibility, Fig. 3 and its accompanying legend are reprinted below with greater contrast.

![Image](image)

**Fig. 3.** Assaying for intrinsic transcript termination at T7Te on ssDNA. C46 complexes bound to Ni²⁺-NTA agarose in TGK-B90M4 were chased with 500 μM ATP, 500 μM GTP, 500 μM CTP, and 500 μM UTP for 10 min at 37°C in the presence of rifampicin at 20 μg/ml and yeast Torula RNA at 0.8 mg/ml. The lanes are assigned as for Fig. 2A. The T7Te terminator is at positions +95 and +96, whereas transcription to the end of the DNA template generates a run-off RNA of +146 nucleotides. C46 complexes in lanes 5–8 were digested with Exo III at 5,000 units/ml for 5 min at 37°C. Unlike C47, some of the C46 complexes failed to resume elongation after treatment with Exo III (see lanes 6 and 7). The mechanism of this inactivation is unknown but similar observations have been made by others (29, 30).

Cell Biology. In the article “Scavenger receptor class B, type I (SR-BI) is the major route for the delivery of high density lipoprotein cholesterol to the steriodogenic pathway in cultured mouse adrenocortical cells” by Ryan E. Temel, Bernardo Trigatti, Ronald B. DeMattos, Salman Azhar, Monty Krieger, and David L. Williams, which appeared in number 25, December 9, 1997, of Proc. Natl. Acad. Sci. USA (94, 13600–13605), the following correction should be noted. The equation on page 13601 should be as follows:

\[ P_{total} = \frac{[P_{max}][HDL]}{K_{HA} + [HDL]} + C[HD] \]
Cell Biology. In the article “Identification of a family of low-affinity insulin-like growth factor binding proteins (IGFBPs): Characterization of connective tissue growth factor as a member of the IGFBP superfamily” by Ho-Seong Kim, Srinivasa R. Nagalla, Youngman Oh, Elizabeth Wilson, Charles T. Roberts, Jr., and Ron G. Rosenfeld, which appeared in number 24, November 25, 1997, of Proc. Natl. Acad. Sci. USA (94, 12981–12986), the authors request that the following corrections be noted. In Fig. 8, the units on the scale should indicate the number of substitution events, rather than “million years.” The lengths of the branches represent the relative distance between the sequences of mammalian IGFBPs compared in this figure. In the Discussion, the statement “The dendrogram depicted in Fig. 8 indicates that, based upon structural similarities, all ten members of the superfamily can be traced back to an ancestor gene 60 million years ago” should read: “The dendrogram depicted in Fig. 8 indicates that all ten members of the superfamily share a common ancestral gene based upon their sequence similarities.”

Neurobiology. In the article “Hair cell-specific splicing of mRNA for the α1D subunit of voltage-gated Ca²⁺ channels in the chicken’s cochlea” by Richard Kollmar, John Fak, Lisa G. Montgomery, and A. J. Hudspeth, which appeared in number 26, December 23, 1997, of Proc. Natl. Acad. Sci. USA (94, 14889–14893), the authors wish to note that the quality of reproduction of Fig. 1 was below standard. In all three panels, the middle parts were affected. Specifically, the reverse (white-on-black) type denoting exons 9a, 22a, and 30a was illegible; parts of the arrows that represented primers such as F9 were missing; and the outlines of several of the boxes that depicted exons such as 9, 10, and 20 were defective. The figure and its legend are reproduced below.

**FIG. 1.** Alternative splicing of the α1D mRNA in the basilar papilla and the brain. (A) Southern blot of PCR products amplified with primers flanking the insert in the I-II loop (exon 9a). Marker sizes in base pairs are indicated on the left. The diagram below of the putative genomic structure (not drawn to scale) depicts exons as rectangles, introns as horizontal lines, and PCR primers as arrows. To amplify all isoforms together, we used primers F9 and R14. To amplify rare isoforms without interference from more abundant ones, we used exon-specific primers: primer F9a binds across the splice junction of exons 9 and 9a, and primer F10 binds across that of exons 9 and 10. The table at the bottom lists product size and occurrence for each splice variant and primer pair. (+), abundant; (+), detectable; (−), barely so; (−), not detectable. (B) Same as A, but for the alternative IIIS2 segment (exon 22a). (C) Same as A, but for the insert in the IVS2–3 loop (exon 30a). Primer F30a binds across the splice junction of exons 30 and 30a, primer F31 binds across that of exons 30 and 31, and primer F31a binds across that of exons 30 and 31a. For the basilar papilla, the lengths of even the minor products were consistent only with splice isoforms containing exon 30a; for the brain, they were consistent only with isoforms lacking exon 30a. Note the abundance in the brain of mRNAs with exons for both IVS3 segments.
Characterization of residual structure in the thermally denatured state of barnase by simulation and experiment: Description of the folding pathway

(protein folding/molecular dynamics/NMR)

CHRIS J. BOND*, KAM-BO WONG*, JANE CLARKE†, ALAN R. FERSHT‡†, AND VALERIE DAGGETT*¶

†Centre for Protein Engineering, Medical Research Council Centre, Hills Road, Cambridge, CB2 2QH, United Kingdom; and *Department of Medicinal Chemistry, University of Washington, Seattle, WA 98195-7610

ABSTRACT Residual structure in the denatured state of a protein may contain clues about the early events in folding. We have simulated by molecular dynamics the denatured state of barnase, which has been studied by NMR spectroscopy. An ensemble of 10⁴ structures was generated after 2 ns of unfolding and following for a further 2 ns. The ensemble was heterogeneous, but there was nonrandom, residual structure with persistent interactions. Helical structure in the C-terminal portion of helix α₁ (residues 13–17) and in helix α₂ as well as a turn and nonnative hydrophobic clustering between β₃ and β₄ were observed, consistent with NMR data. In addition, there were tertiary contacts between residues in α₁ and the C-terminal portion of the β-sheet. The simulated structures allow the rudimentary NMR data to be fleshed out. The consistency between simulation and experiment inspires confidence in the methods. A description of the folding pathway of barnase from the denatured to the native state can be constructed by combining the simulation with experimental data from φ value analysis and NMR.

The denatured state of a protein is an ensemble of rapidly interconverting structures (1–3). It is only recently that structural information for the denatured state has become available at atomic resolution because of developments in high field NMR spectroscopy (4–18). These studies have focused on areas of persistent structure, with the aim of obtaining insight into the relationship between sequence, structure, and in particular, the mechanism of folding. The energy landscape available to a protein is complex, leading to the question of how proteins navigate it and fold on a reasonable, and often very fast, timescale. One hypothesis is that persistent structure under denaturing conditions may guide the folding of the polypeptide chain by limiting the conformational space available or even actively nucleating and promoting structure formation. To test this hypothesis, the structural attributes of the denatured state must be characterized for proteins for which detailed folding information is available.

Barnase is particularly well suited for such endeavors. It is a 110-residue ribonuclease from Bacillus amyloliquefaciens, whose folding/unfolding properties have been well characterized (19). Barnase is a single-domain protein with three helices (α₁–α₃) in the first half of the sequence followed by a five-stranded antiparallel β-sheet (20) (Fig. 1). The major core is formed by the packing of hydrophobic residues in α₁ and the antiparallel β-sheet.

The pH, temperature, and urea-denatured states of barnase have been studied by NMR spectroscopy (6, 8, 14). Structural parameters, such as chemical shifts, coupling constants, and nuclear Overhauser effect crosspeaks, suggest that there is residual structure in the regions corresponding to α₁, α₂, and β(3–4) of native barnase (6, 8). Unfortunately, transformation of these data into specific structural models is not possible. α₁ and a portion of β(3–4) make significant interactions in the intermediate and transition states, as determined by the protein engineering method (21–23). The second helix, α₂, remains at least partially helical under most denaturing conditions, yet it does not appear to be structured in either the intermediate or transition state (21–23). What then is the relationship between stable structures observed under denaturing conditions at equilibrium and the kinetic folding pathway and its associated structures?

We hope both to obtain a more detailed description of the denatured state and to gain insight into the potential importance of residual structure in folding by complementing the experimental picture with a molecular dynamics-generated ensemble of denatured conformations of barnase and the unfolding pathway. To this end, we simulated the thermal denaturation of barnase, whose early unfolding behavior has been the subject of previous unfolding simulations (24, 25). Here, the denatured ensemble is taken as the collection of structures late (2–4 ns) in an all atom, high temperature simulation of barnase in water. The structural properties of the denatured ensemble are presented and compared with experimental NMR data (6, 8, 14). The extent to which residual structure in the unfolded, denatured ensemble contributes to the kinetic pathway is discussed.

METHODS

An all atom representation was used for both the protein and solvent box. The initial starting structure was the average NMR solution structure (20). Protonation states of side chains were chosen to mimic low neutral pH (Lys, Arg, and His residues were positively charged, and Asp and Glu were negatively charged). The first 2 ns of this denaturation simulation at 498 K, as well as the control at 298 K, have been described (26). The solvent density was set to the experimental value for liquid water at 498 K (27), which corresponds approximately to a pressure of 26 atm (1 atm = 101.3 kPa). Extensive tests of the water model (F3C) as a function of temperature and pressure have been described (28). The denaturation simulation was continued to 4 ns to generate an ensemble of unfolded structures. Another simulation was also performed, and its denatured state is in broad agreement with

¶To whom reprint requests should be addressed at: University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K. e-mail: adggett@cam.ac.uk.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1997 by The National Academy of Sciences 0027-8424/97/9413409-5$2.00/0 PNAS is available online at http://www.pnas.org.

the results presented here, and detailed comparison of the two ensembles will be presented elsewhere. Here, we invoke the ergodic hypothesis: the time average of a property for a single system (a single molecular dynamics trajectory) is the same as the average property taken over the members of an ensemble at a particular time (experiment).

RESULTS

We present a 4-ns denaturation simulation, beginning from the average NMR structure, of barnase at 498 K in water. This is a continuation of a simulation and full descriptions of the unfolding pathway, a control simulation at 298 K, and comparisons with experiments for the transition and intermediate states are presented elsewhere (26). Here, we concentrate on the properties of the protein late in the simulation in an attempt to model the denatured state of barnase.

The Denatured Ensemble. Unfolding was identified by the loss in secondary structure, tertiary contacts, and the overall \( \alpha \)-carbon (C\(_\alpha \)) root-mean-square deviation from the initial NMR structure (Table 1). All properties monitored showed little change in their average behavior after 1.5–2 ns. Therefore, we use the 2- to 4-ns structures to represent the denatured state, which corresponds to an ensemble of \( 10^4 \) structures. The denatured state was compact with an average radius of gyration of 16 Å compared with an estimated value of 34 Å for a random coil (30) (Table 1, Fig. 2). The average solvent-accessible surface area for the structures making up the denatured state was \( \approx 50\% \) greater than the native state. The denatured ensemble was heterogeneous with an average root-mean-square deviation spread of \( >6 \) Å (Table 1). The loss of secondary structure was not evenly distributed among the \( \alpha \)-helices and \( \beta \)-sheet. The overall helical content was approximately constant with 13 residues (or 12% of the protein), on average, adopting the appropriate backbone geometries (31) (Table 1). In contrast, only \( \approx 5 \) residues retained \( \beta \)-structure, and protracted \( \beta \)-strands were not observed. The backbone was dynamic, with average fluctuations in the main chain dihedral angles of \( \approx 40^\circ \) (Table 1).

Residual Structure in the Denatured State. \( \alpha 1 \) and \( \alpha 2 \) retained a significant amount of helical structure in the denatured ensemble. The average helix length was 6.4 (±1.0) and 5.8 (±0.5) residues for \( \alpha 1 \) and \( \alpha 2 \), respectively, compared with 11 and 7 for the native structure. Interestingly, \( \alpha 1 \) was partly unfolded at 2 ns, primarily at the N terminus (residues 7–12), but it regained most of its native hydrogen bonding network with time. The packing of hydrophobic residues between \( \alpha 1 \) and residues 88–89 increased as residues 7–13 were added to the helix (Fig. 3). The C terminus of the helix (residues 14–17) maintained a helical turn throughout the ensemble as judged by both backbone orientation and hydrogen bonding. Unlike \( \alpha 1 \), \( \alpha 2 \) was more ordered and yet made few contacts with other elements of structure (Fig. 2).

Disruption of native secondary structure was extensive in the \( \beta \)-sheet region of denatured barnase (residues 50–110). The backbone dihedral angles of these residues sampled a large \((\phi, \psi)\) space, exhibiting some preference for locally extended conformations but not strict repeating \( \beta \)-sheet structures. Many transient turns formed and punctuated the extended portions of the structure. In particular, residues in \( \beta 3 \) and in a turn between \( \beta 3 \) and \( \beta 4 \) formed a nonnative cluster state, which corresponds to an ensemble of \( 10^4 \) structures. The denatured state was compact with an average radius of gyration of 16 Å compared with an estimated value of 34 Å for a random coil (30) (Table 1, Fig. 2). The average solvent-accessible surface area for the structures making up the denatured state was \( \approx 50\% \) greater than the native state. The denatured ensemble was heterogeneous with an average root-mean-square deviation spread of \( >6 \) Å (Table 1). The loss of secondary structure was not evenly distributed among the \( \alpha \)-helices and \( \beta \)-sheet. The overall helical content was approximately constant with 13 residues (or 12% of the protein), on average, adopting the appropriate backbone geometries (31) (Table 1). In contrast, only \( \approx 5 \) residues retained \( \beta \)-structure, and protracted \( \beta \)-strands were not observed. The backbone was dynamic, with average fluctuations in the main chain dihedral angles of \( \approx 40^\circ \) (Table 1).

\[ \text{Residual Structure in the Denatured State. } \alpha 1 \text{ and } \alpha 2 \text{ retained a significant amount of helical structure in the denatured ensemble. The average helix length was 6.4 (±1.0) and 5.8 (±0.5) residues for } \alpha 1 \text{ and } \alpha 2, \text{ respectively, compared with 11 and 7 for the native structure. Interestingly, } \alpha 1 \text{ was partly unfolded at 2 ns, primarily at the N terminus (residues 7–12), but it regained most of its native hydrogen bonding network with time. The packing of hydrophobic residues between } \alpha 1 \text{ and residues 88–89 increased as residues 7–13 were added to the helix (Fig. 3). The C terminus of the helix (residues 14–17) maintained a helical turn throughout the ensemble as judged by both backbone orientation and hydrogen bonding. Unlike } \alpha 1, \alpha 2 \text{ was more ordered and yet made few contacts with other elements of structure (Fig. 2).} \]

\[ \text{Disruption of native secondary structure was extensive in the } \beta \text{-sheet region of denatured barnase (residues 50–110). The backbone dihedral angles of these residues sampled a large } (\phi, \psi) \text{ space, exhibiting some preference for locally extended conformations but not strict repeating } \beta \text{-sheet structures. Many transient turns formed and punctuated the extended portions of the structure. In particular, residues in } \beta 3 \text{ and in a turn between } \beta 3 \text{ and } \beta 4 \text{ formed a nonnative cluster state, which corresponds to an ensemble of } 10^4 \text{ structures.} \]

\[ \text{The results presented here, and detailed comparison of the two ensembles will be presented elsewhere. Here, we invoke the ergodic hypothesis: the time average of a property for a single system (a single molecular dynamics trajectory) is the same as the average property taken over the members of an ensemble at a particular time (experiment).} \]

Table 1. Overall properties of the denatured ensemble

<table>
<thead>
<tr>
<th>Property</th>
<th>NMR structure</th>
<th>Denatured state (MD)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radius of gyration, Å</td>
<td>13.4</td>
<td>15.9 (0.7)</td>
</tr>
<tr>
<td>( \langle C^\alpha \rangle ) RMSD, Å</td>
<td>11.3 (0.7)</td>
<td>6.4 (1.6)</td>
</tr>
<tr>
<td>( \langle \text{Pairwise RMSD} \rangle ), Å</td>
<td>36</td>
<td>5 (3)</td>
</tr>
<tr>
<td>Total % ( \alpha )</td>
<td>19</td>
<td>12 (4)</td>
</tr>
<tr>
<td>Total % ( \beta )</td>
<td>38 (18)</td>
<td>42 (19)</td>
</tr>
<tr>
<td>( \langle \phi \rangle ), °</td>
<td>—</td>
<td>38 (18)</td>
</tr>
<tr>
<td>( \langle \psi \rangle ), °</td>
<td>—</td>
<td>42 (19)</td>
</tr>
<tr>
<td>(Trp-35 SASA), Å²</td>
<td>6357</td>
<td>9513 (370)</td>
</tr>
<tr>
<td>(Trp-71 SASA), Å²</td>
<td>1</td>
<td>162 (35)</td>
</tr>
<tr>
<td>(Trp-94 SASA), Å²</td>
<td>1</td>
<td>142 (37)</td>
</tr>
<tr>
<td>(Trp-94 SASA), Å²</td>
<td>76</td>
<td>94 (41)</td>
</tr>
</tbody>
</table>

*The values for the denatured ensemble were calculated for the 2 to 4-ns time period of the 498 K simulation. Standard deviations are given in parentheses. MD, molecular dynamics.

†Only the \( \alpha \)-carbons were used for the calculation of the radius of gyration.

‡The \( \langle C^\alpha \rangle \) root-mean-square deviation (RMSD) from the starting structure was calculated after removal of translational and rotational motion. The average pairwise RMSD is calculated by comparing all structures in the denatured ensemble in a pairwise manner.

§The secondary structure content is the fraction of residues (of a total of 107) that adopted repeating secondary structure (3 or more consecutive residues) based on whether the \((\phi, \psi)\) values fall in the \( \alpha \)- or \( \beta \)-region of conformational space (31).

¶The solvent-accessible surface area (SASA) was calculated by using the program NACCESS (32) using a 1.4-Å probe radius.

\[ \text{FIG. 1. The NMR solution structure of barnase (20) with segments of secondary structure labeled. The solvent-accessible surface area (SASA) was calculated by using a 1.4-Å probe radius.} \]

\[ \text{FIG. 2. Representative hydrophobic clusters and electrostatic interactions in the denatured state of barnase. (A) The 3-ns structure is displayed. The backbone of the } \beta(3–4) \text{ turn is colored yellow. (B) Representative side chain interactions in the } \beta(3–4) \text{ turn are displayed for the 2- and 4-ns structures.} \]

\[ \text{Proc. Natl. Acad. Sci. USA 94 (1997)} \]
The most persistent interactions can be summarized as follows: (i) helical structure in the C-terminal portion of α1 (residues 13–17) and in α2; (ii) a turn and hydrophobic clustering between β3 and β4; and (iii) tertiary contacts between residues in α1 and the C-terminal portion of the β-sheet. Some of these features are illustrated in representative structures in Figs. 2 and 3.

The available structural experimental information for the denatured state of barnase, as summarized in Fig. 5, correlates with the regions of structure observed in the simulation. Significant deviations from random coil values for the amide and 15N chemical shifts occur in α1 and α2 and residues between β3 and β4 (Fig. 5), as do the strong sequential NH—NH (i → i + 1) nuclear Overhauser effect crosspeaks, 6, 8, 14). These helical regions are observed directly in the simulation (Figs. 2 and 3). α1 underwent unfolding and refolding in the denatured state, although the C-terminal portion of the helix was stable throughout (Fig. 3). Interestingly, docking of side chain residues of the helix with residues in the β-sheet region of the protein aided in the folding of the helix even though the interactions were a mix of native and nonnative contacts because of the disorder of the β-sheet structure. The main chain of β(3-4) appears to be the most structured portion of β-sheet based on the NMR data (6, 8, 14). In addition, the side chains in this region appear to be partially structured, given the deviations of their side chain chemical shifts from random coil values (Fig. 5) and the strong sequential NH—NH nuclear Overhauser effect crosspeaks (6, 8, 14), yet interestingly these data are not consistent with native structure (8). The simulation provides a structural framework for interpretation of these results: residues 90–94 participate in a turn between β3 and β4 that is stabilized by a dynamic hydrophobic cluster (including many residues spanning β3 and β4, Fig. 2). Further, the cluster involved residues separated by >17 Å in the native state. Also, the native salt bridge between Asp-93 and Arg-69 aided in maintenance of the turn (Fig. 2). Experimental studies show the pKa of Asp-93 is decreased significantly in the denatured state, suggesting it is involved in a favorable electrostatic interaction (33). Despite these tertiary interactions, extension of the turn to form a β-hairpin did not occur, and the main chain adopted predominantly nonnative conformations, leading to low S values (discussed further below) (Fig. 5).

Residual Structure: Relevance to the Folding/Unfolding Pathway. Both the NMR studies and the simulation provide evidence for residual structure in the denatured state of barnase. We have outlined the specifics of that structure, which in some cases is typified by tertiary interactions in the absence of secondary structure and in others is because of persistent but fluctuating secondary structure. Next, we consider the residual structure in light of the folding pathway of barnase. Both the major intermediate (I) and transition state (TS) have been investigated by φ value analysis (34–36). In this, the degree of formation of structure at a particular residue is inferred from changes in the energetics of kinetics and thermodynamics on mutation of that residue or surrounding ones. A φ value of 0 implies denatured-like structure and φ = 1, native-like structure at that position. Although structural attributes of the intermediate states during folding/unfolding are inferred just from the energetics by using this approach, there is a good correlation with the results of simulations that probe structure directly (24–26). Each residue in the putative

Fig. 3. Side chain interactions between residues in α1 and the β-sheet region of the native protein that aid in the coil → helix transition from 2 to 2.5 ns.

Fig. 4. Snapshots from the denatured ensemble depicting motions of the protein backbone. The structures are colored from red at the N terminus (looking down the helix axis) to blue at the C terminus. The 2-, 2.5-, 3-, 3.5-, and 4-ns structures are displayed. The 6-Å separation labeled in the figure is shown merely to provide perspective.
models for these states may be assigned a structure index ($S$) for semi-quantitative comparison with $\phi$ (37). Thus, the approaches are quite complementary, and when the ensembles representing the transition or intermediate states in a simulation are in agreement with experiment, we assume that the structural models can be used in the interpretation of the experimental data. Therefore, for the purposes of discussing the possible role of residual structure in the folding/unfolding pathway of barnase, we consider the experimental and simulation results for the transition and intermediate states of barnase together.

The regions of residual structure in denatured barnase correspond to the moderately and highly structured portions of the intermediate and transition states fairly well, as determined by the protein engineering method (compare regions of residual structure in $D$ to the magenta and blue $\phi$ values in Fig. 5). The C-terminal portion of $\alpha 1$, portions of the loop between $\beta 1$ and $\beta 2$, and several residues within $\beta 3$ have $\phi$ values close to 1. In contrast, mutations throughout $\alpha 2$ do not appear to affect the stability of the intermediate or the transition state (23). Yet, the conclusion from the NMR studies is that residues within $\alpha 2$ are helical in the denatured state.

What then is the relationship between structure observed in denatured proteins and the deduced structural interactions in kinetic folding intermediates? It is not necessary for all structure observed in a denatured protein to lead to productive interactions along the folding pathway. Nonnative interactions that are sufficiently fluid can decrease the conformational search without excessive stabilization of partially folded or unfolded forms of the protein. Excessive stabilization of nonnative interactions slows down folding (36), as will excessive stabilization of native structure that is not formed in the transition state (38). This may be true even for the native structure seen in $\alpha 2$. In this regard, it is worth noting that while $\alpha 2$ is intact in the simulated denatured state, it is not present in the simulated intermediate state (Fig. 5). Whereas having $\alpha 1$ and $\alpha 2$ in the helical conformation in the denatured state presumably decreases the conformational search during folding, another important factor was contacts between loop 1 and $\beta 1$ that served to “pinch off” the second domain, containing $\alpha 2$, $\alpha 3$, and loop 2 (Fig. 5). These interactions prevented gross extension of the first half of the protein. This region also displays lower mobility and/or residual structure by NMR (Fig. 5), in support of the idea that it is partially segregated from the rest of the protein.

The persistent structure seen in the denatured ensemble and intermediate and transition states, as suggested by both experiment and simulation, supports a folding pathway in which regions of local structure, primarily in the C terminus of $\alpha 1$ and the $\beta (3–4)$ turn region, initiate early events in folding. This is in agreement with predictions of potential nucleation sites based on burial of hydrophobic surface area (39) and NMR studies of barnase peptide fragments (40). In other words, $\alpha 1$, core 1, and portions of the $\beta$-sheet fold about persistent initiation sites in the denatured ensemble until enough stabilizing interactions involving residues distant in sequence form a nucleation site (refs. 14 and 41; Fig. 5). Following initiation of folding, there is further consolidation of both secondary and tertiary structure, as reflected in the increase in the interaction energy, or $\phi$ values, and as observed in the simulation (Fig. 5).

An example of how such nonlocal interactions may not only stabilize but also induce the formation of secondary structure is illustrated in Fig. 3, which shows a coil to helix transition of $\alpha 1$ in the denatured state. An invariant one and one-half turns of helix were present in all of the denatured structures. However, the N-terminal portion of $\alpha 1$ only adopted significant amounts of helical structure in the presence of tertiary contacts with the $\beta$-sheet, in particular with the $\beta (3–4)$ turn. Whereas this docking was not stable or persistent enough to lead to global refolding under these conditions, it did lead to
repair of α. Further, this process provides a description of a plausible structural mechanism for nucleation leading to consolidation of structure. In this regard, it is worth emphasizing that α and β-(3–4) are structured in the denatured state but undergo conformational exchange. They are well established in the simulated intermediate state and are loosely coupled to neighboring strands. The packing of the other β-strands around this nucleus is essentially complete in the transition state (Fig. 5), but β-(3–4) appears to be more important than α1 in propagation of the β-structure (Fig. 5), which is supported by fragment studies in which native-like structure persists on removal of α1 (42).

Figs. 1–3 were made by using MOLSCRIPT (43) and Figs. 4 and 5 with University of California, San Francisco MIDASPLUS (29). This work was supported by National Institutes of Health Grant GM 50789 (to V.D.).