the extracellular loops of FecA are not restricted by crystal contacts. The absence of an altered β-barrel conformation upon siderophore binding may also reflect functional differences between FhuA and FecA or the regulation of the *fec* transport operon.
17. D. E. Koshland, *Nature Med*. **4**, 1112 (1998).
18. R. J. Kadner, K. J. Keller, *J. Bacteriol*. **177**, 4829 (1995).
19. E. Schramm, J. Mende, V. Braun, R. M. Kemp, *J. Bacteriol*. **169**, 3350 (1987).
20. N. Cadieux, R. J. Kadner, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10673 (1999).
21. N. Cadieux, C. Bradbeer, R. J. Kadner, *J. Bacteriol*. **182**, 5954 (2000).
22. G. S. Moeck, L. Letellier, *J. Bacteriol*. **183**, 2755 (2001).
23. G. S. Moeck, J. W. Coulton, K. Postle, *J. Biol. Chem*. **272**, 28391 (1997).
24. M. Braun, H. Killmann, V. Braun, *Mol. Microbiol*. **33**, 1037 (1999).
25. R. A. Larsen, M. G. Thomas, K. Postle, *Mol. Microbiol*. **31**, 1809 (1999).
26. K. C. Usher, E. Özkan, K. H. Gardner, J. Deisenhofer, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10676 (2001).
27. B. S. Smith *et al*., *Acta Crystallogr.* **D54**, 697 (1998).
28. S. Doublié, *Methods Enzymol*. **276**, 532 (1997).
29. Z. Otwinowski, W. Minor, *Methods Enzymol*. **276**, 307 (1997).
30. W. A. Hendrickson, *Science* **254**, 51 (1991).
31. C. Weeks, R. Miller, *Acta Crystallogr.* **D55**, 492 (1999).
32. E. de La Fortelle, G. Bricogne, *Methods Enzymol*. **276**, 472 (1997).
33. A. T. Brünger *et al*., *Acta Crystallogr.* **D54**, 905 (1998).
34. T. A. Jones, J.-Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr.* **A47**, 110 (1991).
35. A. D. Ferguson *et al*., data not shown.
36. M. Bonhivers, A. Ghazi, P. Boulanger, L. Letellier, *EMBO J.* **15**, 1850 (1996).
37. L. Letellier, K. P. Locher, L. Plançon, J. Rosenbusch, *J. Biol. Chem*. **272**, 1448 (1997).
38. C. Bös, D. Lorenzen, V. Braun, *J. Bacteriol*. **180**, 605 (1998).
39. C. S. Klug, S. S. Eaton, G. R. Eaton, J. B. Feix, *Biochemistry* **37**, 9016 (1998).
40. D. C. Scott *et al*., *J. Biol. Chem*. **276**, 13025 (2001).
41. X. Jiang *et al*., *Science* **276**, 1261 (1997).
42. R. M. Esnouf, *Acta Crystallogr.* **D55**, 938 (1999).
43. L. Esser, personal communication.
44. A. Nicholls, K. A. Sharp, B. Honig, *Proteins Struct. Funct. Genet*. **11**, 282 (1991).
45. We thank N. Duke, S. Korolev, R. M. Sweet, and C. Ogata for their assistance during data collection; A. Bino for dinuclear ferric citrate; L. Venkatramani and E. Lemke for initial purification and crystallization experiments; Z. Otwinowski and W. Minor for sharing their data reduction expertise; and C. A. Brautigam, C. L. Colbert, E. Goldsmith, and D. Tomchick for critical reading of the manuscript. Supported by grants from the Welch Foundation (J.D.) and NIH (D.v.d.H.) and by postdoctoral fellowships from the Human Frontier Science Program and Canadian Institutes of Health Research (A.D.F.). Crystallographic coordinates and structure factor amplitudes have been deposited in the Protein Data Bank under the accession codes 1KMO (unliganded) and 1KMP (liganded).

22 October 2001; accepted 2 January 2002

Long-Range Interactions Within a Nonnative Protein

Judith Klein-Seetharaman,[1]* Maki Oikawa,[2] Shaun B. Grimshaw,[3] Julia Wirmer,[1]† Elke Duchardt,[1]† Tadashi Ueda,[2] Taiji Imoto,[2] Lorna J. Smith,[3] Christopher M. Dobson,[3]‡ Harald Schwalbe[1]†§

Protein folding and unfolding are coupled to a range of biological phenomena, from the regulation of cellular activity to the onset of neurodegenerative diseases. Defining the nature of the conformations sampled in nonnative proteins is crucial for understanding the origins of such phenomena. We have used a combination of nuclear magnetic resonance (NMR) spectroscopy and site-directed mutagenesis to study unfolded states of the protein lysozyme. Extensive clusters of hydrophobic structure exist within the wild-type protein even under strongly denaturing conditions. These clusters involve distinct regions of the sequence but are all disrupted by a single point mutation that replaced residue Trp^{62} with Gly located at the interface of the two major structural domains in the native state. Thus, nativelike structure in the denatured protein is stabilized by the involvement of Trp^{62} in nonnative and long-range interactions.

Incompletely folded states of proteins are coupled to cellular processes such as protein synthesis, translocation across membranes, and signal transduction [reviewed in (*1*, *2*)]. In addition, intrinsically unstructured proteins have been predicted to be common within the genomes of all organisms (*3*). Unstructured and partially folded conformations of proteins are, however, prone to aggregate and have been implicated in a wide range of diseases (*4*). The structural and dynamic characterization of nonnative states of proteins is therefore crucial for understanding these processes in addition to being fundamental to an understanding of protein folding itself.

Nonnative states of proteins are ensembles of conformers, the individual members of which may differ substantially in their structural and dynamic properties. Conformational sampling of denatured proteins can be significantly restricted, and the existence of "compact states" has been postulated to occur (*5*–*9*). In some cases, specific experimental structural information has been obtained although in general this information is either indirect or highly localized. An important question relating to all nonnative states is the extent to which long-range interactions are important in the stabilization of nonrandom interactions. We use site-directed mutagenesis and NMR spectroscopy to show that long-range nonnative interactions stabilize native-like hydrophobic clusters in lysozyme.

A wide range of approaches has been developed to characterize nonnative states of proteins in atomic detail by NMR spectroscopy (*10*), and evidence for the presence of residual structure even under strongly denaturing conditions has been presented, see, e.g. (*11*–*17*). Residual structure appears to reside predominantly in hydrophobic clusters, in which tryptophan or histidine residues are surrounded by other hydrophobic side chains (*18*–*20*). It has been postulated that hydrophobic clusters are stabilized by long-range interactions and may influence the folding of the protein, for example by acting as nucleation sites around which structure can be formed (*16*–*18*, *21*). Hydrophobic clusters have also been identified in nonnative states of hen lysozyme, in both the oxidized and the reduced form in 8 M urea at pH 2 [in the reduced protein the free sulfhydryl groups are blocked by methylation (*16*)].

Of the measured NMR parameters, chemical shift values of H^N and H_α protons and transverse (R_2) relaxation rates are the most direct indicators of residual structure. Here, we use such parameters to examine the reduced state of hen lysozyme in the absence of urea and then to investigate the structural changes resulting from the replacement of residue Trp^{62} by Gly (W62G). H^N and H_α chemical shifts measured for reduced and methylated wild-type lysozyme (WT-S^{ME}) in water are shown in Fig. 1A along with data for WT-S^{ME} in 8 M urea (*16*). In WT-S^{ME}, significant deviations in chemical shifts of the H^N resonances from random coil values (*22*, *23*) can be seen for Gly^{22}, Trp^{63}, and Cys^{64}, and of the H_α resonances for residues 19 to 32, 58 to 64, 119 to 124, and 106 to 113. The largest differences are observed at positions 106 to 116, a result indicative of an increase in helical character for this region of

[1]Massachusetts Institute of Technology, Department of Chemistry, Francis Bitter Magnet Laboratory, 170 Albany Street, Cambridge, MA 02139, USA. [2]Graduate School of Pharmaceutical Sciences, Kyushu University, Fukuoka 812-8582, Japan. [3]Oxford Centre for Molecular Sciences, New Chemistry Laboratory, University of Oxford, South Parks Road, Oxford, OX1 3QH, UK.

*Present address: Institute for Software Research International, Carnegie Mellon University, Wean Hall 4604, Pittsburgh, PA 15213, USA.

†Present address: Johann Wolfgang Goethe-University, Center for Biological Magnetic Resonance, Institute for Organic Chemistry, Marie-Curie-Strasse 11, D-60439 Frankfurt am Main, Germany.

‡Present address: Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK.

§To whom correspondence should be addressed. E-mail: schwalbe@nmr.uni-frankfurt.de

the sequence for the protein in water compared with it in urea; consistent with this conclusion is the observation from circular dichroism (CD) measurements of an 8% increase in helicity.

Heteronuclear ^{15}N R_2 relaxation rates of backbone amide groups of WT-S^{ME} in water are shown in Fig. 2B. For comparison, ^{15}N R_2 rates of this protein in 8 M urea (pH 2) are reproduced from (*16*) in Fig. 2A. The relaxation rates for the protein in urea reach a plateau in the central region of the protein sequence, with the terminal regions having shorter relaxation rates. Simple models of a polypeptide chain, in which the physical properties of the chain are dominated by unrestrained segmental motion, predict this pattern of behavior (*16*). Thus the ^{15}N relaxation properties of a given amide group are in general not markedly influenced by the identity of its neighbors, but predominantly reflect the motional properties of the polypeptide main chain. The model allows a two-parameter fit of experimental relaxation rates when using an intrinsic relaxation rate $R_{intrinsic}$ of an amide ^{15}N nucleus and a persistence length (λ_0) of the polypeptide chain [Web eq. 1 in supplementary material (*24*)]. The relaxation rates fitted by the model are indicated by the bold line in Fig. 2A. The best fit for the protein in water (Fig. 2B) corresponds to $R_{intrinsic}$ of 0.20 s^{-1} and λ_0 of 7 (in units of number of residues); in urea, the values were found to be 0.25 s^{-1} and 7, respectively. Thus, WT-S^{ME} in water has the same chain stiffness as WT-S^{ME} denatured in urea ($\lambda_0 = 7$). The lower value for $R_{intrinsic}$ of WT-S^{ME} in water compared with that in urea can be attributed to the differences in viscosity of the two solvent systems (*25*).

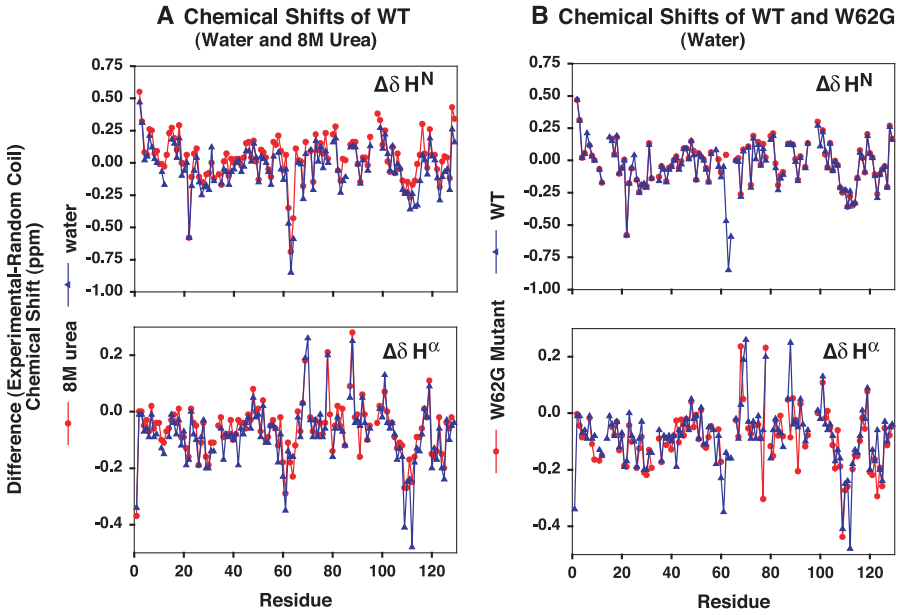

Fig. 1. (**A**) Comparison of perturbations of chemical shift values measured for reduced hen lysozyme (WT-S^{ME}) in water and in 8 M urea from chemical shifts measured in short unstructured peptides ($\Delta\delta = \delta^{expt} - \delta^{random\ coil}$) (*22*, *23*). (Top) $\Delta\delta H^N$ for WT-S^{ME} in water (red line) and in 8 M urea (blue line). (Bottom) $\Delta\delta H_\alpha$ for WT-S^{ME} in water (red line) and in 8 M urea (blue line). (**B**) Comparison of perturbations of chemical shift values measured for reduced hen lysozyme (WT-S^{ME}) and W62G. (Top) $\Delta\delta H^N$ for WT-S^{ME} in water (red line) and for W62G-S^{ME} in water (blue line). (Bottom) $\Delta\delta H_\alpha$ for WT-S^{ME} in water (red line) and for W62G-S^{ME} in water (blue line).

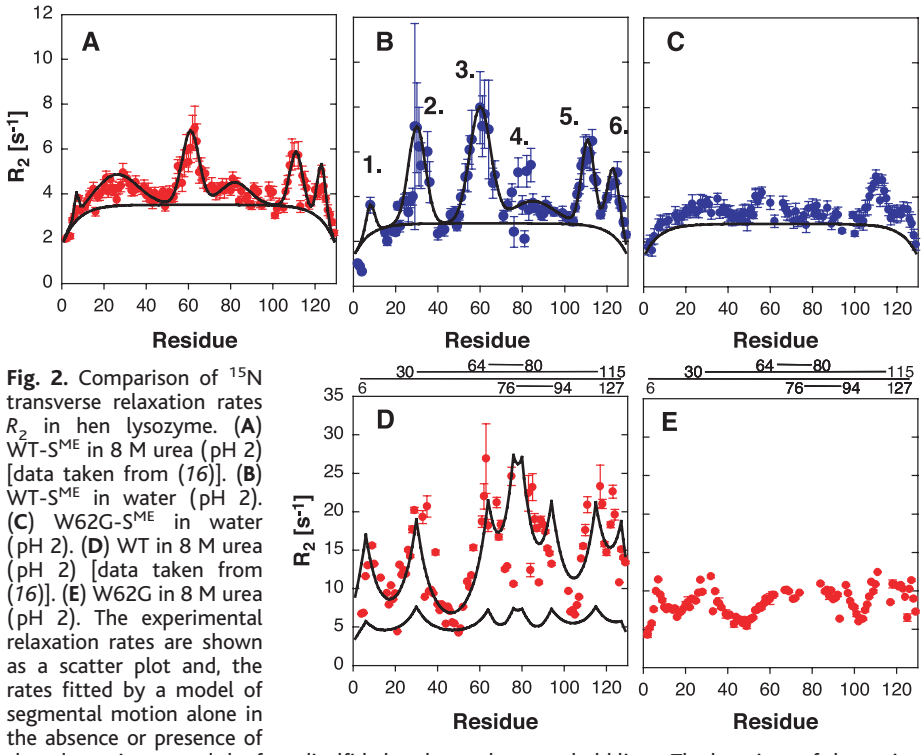

Fig. 2. Comparison of ^{15}N transverse relaxation rates R_2 in hen lysozyme. (**A**) WT-S^{ME} in 8 M urea (pH 2) [data taken from (*16*)]. (**B**) WT-S^{ME} in water (pH 2). (**C**) W62G-S^{ME} in water (pH 2). (**D**) WT in 8 M urea (pH 2) [data taken from (*16*)]. (**E**) W62G in 8 M urea (pH 2). The experimental relaxation rates are shown as a scatter plot and, the rates fitted by a model of segmental motion alone in the absence or presence of slow dynamics around the four disulfide bonds are shown as bold lines. The locations of the native disulfide bridges are indicated. Six clusters of residual structure were identified for HEWL-S^{ME} in water (B). Values obtained by statistical modeling are given in (*24*).

Several regions of the polypeptide chain show larger ^{15}N R_2 relaxation rates than those anticipated from the model. These differences indicate the presence of nonrandom structure in the chain, attributable to the presence of clusters of hydrophobic side chains (*26*). Overall, the distribution of relaxation rates as a function of sequence could be described by a sum of the segmental motion and additional deviations centered around six clusters, which can be modeled as a Gaussian distribution [Web eq. 2 in (*24*)]. Although the locations of the regions showing such deviations are similar for the protein in the two solvents, their intensity and length are very different. In general, the deviations are larger for the protein in water than in urea. In particular, cluster 2 is much more pronounced and defined in water than in urea. This is consistent with the anticipated weakening of hydrophobic interactions in the presence of denaturant.

Lysozyme contains two structural domains, the α domain, involving residues 1 to 35 and 85 to 129, and the β domain, which comprises residues 36 to 84 (*27*, *28*). The clusters are located in sequentially distinct regions along the polypeptide chain; the positions are highlighted in Fig. 3. The largest clusters of residues from the relaxation analysis correlate with the location of hydrophobic residues in the sequence; tryptophan residues are involved in four of the six clusters: Trp^{28} (cluster 2), Trp^{62} and Trp^{63} (cluster 3), Trp^{108} and Trp^{111} (cluster 5), and Trp^{123} (cluster 6).

Elevated values of R_2 relaxation rates are observed toward the NH_2-terminus of the pro-

tein in water, notably for Ala^{9} to Ala^{11} (cluster 1); these residues are located in the region of the sequence that forms the A helix in the native structure. High values of R_2 are also found for Leu^{25} to Glu^{35} residues located in the B helix in the native structure (cluster 2) (Fig. 3B). Side chains from the A and B helices interact with each other in the native structure. Specifically, Ala^{9} is close to Leu^{25} and Val^{29}; both of the latter residues are located on one face of helix B. On the other face of the B helix, there is a common interface with helix D (cluster 5) and the residues of the weak and extended cluster 4. Contacts involve residues Asn^{27} to Trp^{111}, Trp^{28} to Trp^{108}, Phe^{34} to Ala^{110}, Phe^{34} to Arg^{114}, Glu^{35} to Trp^{108}, and Glu^{35} to Ala^{110}. This interface brings into proximity residues that are linked in the native state by the formation of a disulfide bond between Cys^{30} and Cys^{115} in the oxidized protein. Trp^{123} is close to Ala^{9} (cluster 1) and Val^{29} (cluster 2). Thus, a core of residual structure appears to be formed, even in the absence of disulfide bridges in WT-S^{ME}, by clusters 1, 2, 5, and 6 in what is to become the central region of the α domain (Fig. 3). A second core of residual structure is formed by cluster 3 located at the interface between the α and β domains in the native state.

Kinetic refolding studies indicate that both domains achieve their nativelike structure in folding intermediates formed before the development of the extensive tertiary interactions that span the two domains (*29*, *30*). In the hydrogen exchange measurements, on which this conclusion is based, protection of amide hydrogens from exchange with solvent water occurs significantly faster for residues in the α domain than for those in the β domain. An exception is residue Trp^{63}, whose amide hydrogen becomes protected as rapidly as most of the residues in the α domain, despite its location in the β domain of the native structure. This result suggests that the folding of Trp^{63} must be associated in some manner with the folding of the α domain. Rothwarf *et al*. showed in addition that Trp^{62} and Trp^{108} are involved in nonnative tertiary interactions in intermediates populated during the refolding (*31*); replacement of tryptophan residues by tyrosines at either of these two positions results in an increase in the rate of refolding relative to that of the wild-type sequence. Further evidence that Trp^{62} plays an important role in folding arises from the observation that chemical modification of this residue increases the propensity of the protein to misfold (*32*). It has also been shown that Trp^{62} is critical for the correct formation of disulfide bonds in peptide fragments lacking residues 1 to 59 and 105 to 129 (*33*). Furthermore, in a peptide fragment (36 to 105) with native disulfide bonds, tryptophan fluorescence spectra indicate that Trp^{62} and Trp^{63} become exposed upon disruption of helical structure at position 88 to 98 through acetylation of Lys^{96} and Lys^{97}, suggesting that Trp^{62} and Trp^{63} interact with residues 88 to 98 (*34*).

In order to explore the tertiary interactions involving Trp^{62}, ^{15}N R_2 relaxation rates were measured for a mutant protein with a single amino acid replacement, W62G. ^{15}N R_2 relaxation rates for W62G-S^{ME} in solution in the absence of urea are shown in Fig. 2C. Remarkably, the marked deviations (Fig. 2B) from random coil behavior observed in the WT protein are virtually absent, although the parameters defining the underlying random coil behavior are unchanged (bold line in Fig. 2, B and C). This result indicates that all the clusters of residual structure are substantially disrupted by the single amino acid replacement at the domain interface. It also reveals that Trp^{62} must be involved in extensive long-range tertiary interactions in the denatured state of the WT, in order for the effect of its substitution by glycine to be so widespread. As WT-S^{ME} does not contain disulfide bonds, these long-range interactions do not require the cross-linking of different regions of the sequence. In addition, the CD spectra of W62G in water and in 8 M urea at pH 2 do not indicate substantive residual helical character under both conditions.

The presence of disulfide bonds does, however, affect the properties of denatured lysozyme. In the oxidized WT protein (Fig. 2D), the deviations of the rates from random coil values are larger than in the reduced protein with R_2 values of up to 20 s^{-1}; for comparison, the maximum value of R_2 is 7 s^{-1} in WT-S^{ME} (*16*). These increased relaxation rates have been attributed to the intrinsic effect of cross-linking on the conformational dynamics of the polypeptide chain in the vicinity of the disulfide bonds; the simulated effects of such motions are shown in Fig. 2D. This conclusion is, however, challenged by the observation in the present study of the dramatic effects of the W62G mutation on the dynamics of the polypeptide chain in the presence of the disulfide bonds. The variation in R_2 values is substantially diminished as a result of this mutation, although the overall perturbations of the relaxation rates from random coil values still correlate with the location of the disulfide bonds. The relaxation rates of the oxidized W62G mutant are therefore unlikely to reflect significantly the direct contribution of the presence of additional covalent restraints resulting from disulfide cross-linking to the dynamics of the polypeptide chain. Rather, the strikingly larger deviations from random coil behavior observed in oxidized WT compared with WT-S^{ME} can be attributed to the effects of the additional restraints imposed by the presence of disulfide bonds on the hydrophobic clusters present in the denatured protein.

The changes in relaxation rates throughout the protein sequence as a result of the W62G mutation indicate that the deviations from random coil behavior in the WT-S^{ME} protein must be coupled together by long-range tertiary interactions. In order to probe these effects in more detail, analysis of the chemical shifts of the different residues in WT-S^{ME} and W62G-S^{ME} was carried out and is shown in Fig. 1B. There is an excellent correlation ($r^2 = 0.995$) between the H^N chemical shifts in the two proteins when residue 62 is excluded from the analysis (Fig. 1B, top). A similar situation pertains to the H_α chemical shifts (Fig. 1B, bottom), although here the correlation was somewhat lower than for H^N chemical shifts ($r^2 = 0.935$). That the mutation does not alter significantly the chemical shift values of residues other than Trp^{62} and its immediate neighbors indicates that the local structures of the clusters

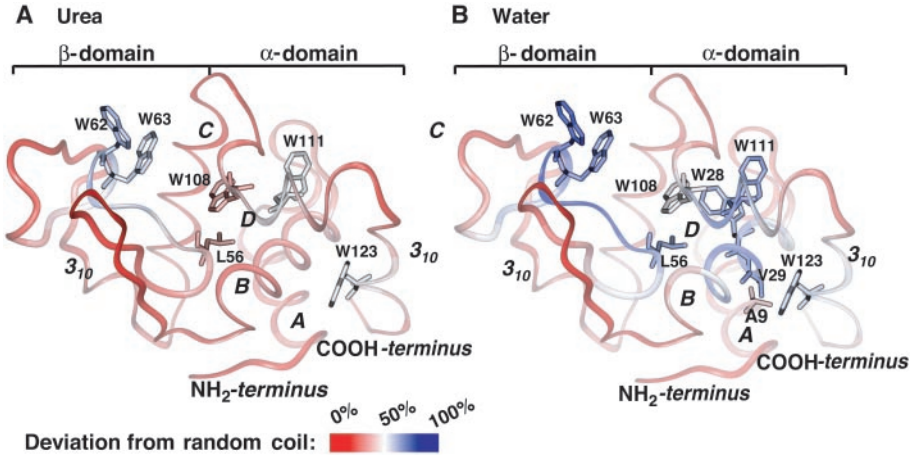

Fig. 3. Mapping of the deviations of the relaxation behavior from those predicted by a random coil model (black line in Fig. 2, A and B) onto the native state structure (Protein Data Bank access code 193L) (*39*). (**A**) ^{15}N relaxation rates measured for WT-S^{ME} in 8 M urea (Fig. 2A). (**B**) ^{15}N relaxation rates measured for WT-S^{ME} in water (Fig. 2 B). Single-letter amino acid abbreviations: A, Ala; L, Leu; V, Val; and W, Trp.

are on average largely unperturbed. We conclude that the observed changes in relaxation rates result primarily from changes in the dynamic rather than structural properties of the various clusters.

In the native structure, Trp^{62} is highly exposed to solvent; the crystal structure shows that its side chain is substantially disordered (*27*), and NMR measurements in solution reveal dynamic behavior (*35*). By contrast, in the denatured states (*36*), and particularly during the early stages of folding (*37*), NMR experiments indicate that this residue, like the other tryptophan residues, is largely inaccessible to solvent. Together with data from the hydrogen-exchange protection experiments, we conclude that in these nonnative states the β-domain residues Trp^{62} and Trp^{63} associate with a nativelike hydrophobic cluster in the α domain involving Trp^{108} and Trp^{111} that is itself strongly linked to the other regions of nonrandom structure. Thus nonnative interactions stabilize a nativelike core. Presumably, the replacement of the large hydrophobic tryptophan residue at position 62 by glycine results in the destabilization of this core. The resulting increase in dynamic flexibility could reflect an increased population of more extended structures in the ensemble of interconverting conformers. A polypeptide chain in such structures will undergo conformational averaging much more rapidly than in compact denatured states, where significant energetic barriers are known to exist.

Our results suggest that, within the ensemble of conformations representing the denatured states of lysozyme, there are long-range interactions that link clusters of residues that are not close together in sequence. These results are consistent with the hypothesis that steps that involve the reorganization of the species formed initially during the refolding of lysozyme are likely to be key determinants of the kinetics of the folding process (*29*, *30*). Although the folding of small proteins is dominated by the search for nativelike contacts, in the case of larger proteins, including those with multiple domains, species with at least some nonnative interactions can be important determinants of the folding process (*8*). Such interactions appear to be located primarily at the interface between the two structural domains, the region associated with the slowest step in the folding of lysozyme (*4*).

The avoidance of misfolding and potential aggregation of nonnative species is a key aspect of the folding and long-term stability of proteins. For example, single point mutations in human lysozyme are responsible for the occurrence of systemic disease in which large quantities of amyloid fibrils are deposited in a variety of internal organs *(38)*. That a single amino acid replacement can perturb the nonnative state of the protein is of particular interest, as the aggregation of partially or completely unfolded species is the essential step in the development of the amyloid structures. Thus, although a residue such as tryptophan may be exposed in the native state for functional reasons, it could be buried in the early stages of folding to reduce the tendency of these transiently populated species to aggregate. Such a conclusion leads to the possibility that the sequence of a protein codes for structural characteristics other than those of the native fold.

References and Notes

1. P. E. Wright, H. J. Dyson, *J. Mol. Biol.* **293**, 321 (1999).
2. S. E. Radford, C. M. Dobson, *Cell* **97**, 291 (1999).
3. P. Romero *et al.*, *Pac. Symp. Biocomput.* **3** (1998).
4. C. M. Dobson, *Philos. Trans. R. Soc. London B. Biol. Sci.* **356**, 133 (2001).
5. B. A. Shoemaker, P. G. Wolynes, *J. Mol. Biol.* **287**, 657 (1999).
6. B. A. Shoemaker, J. Wang, P. G. Wolynes, *J. Mol. Biol.* **287**, 675 (1999).
7. W. F. van Gunsteren, R. Burgi, C. Peter, X. Daura, *Angew. Chem. Int. Ed. Engl.* **40**, 351 (2001).
8. C. M. Dobson, A. Sali, M. Karplus, *Angew. Chem. Int. Ed. Engl.* **37**, 868 (1998).
9. D. Shortle, M. S. Ackerman, *Science* **293**, 487 (2001).
10. H. J. Dyson, P. E. Wright, *Nature Struct. Biol.* **5** (suppl.), 499 (1998).
11. Y. K. Mok, C. M. Kay, L. E. Kay, J. Forman-Kay, *J. Mol. Biol.* **289**, 619 (1999).
12. F. J. Blanco, L. Serrano, J. D. Forman-Kay, *J. Mol. Biol.* **284**, 1153 (1998).
13. D. Shortle, C. Abeygunawardana, *Structure* **1**, 121 (1993).
14. K. M. Fiebig, H. Schwalbe, M. Buck, L. J. Smith, C. M. Dobson, *J. Phys. Chem.* **100**, 2661 (1996).
15. L. J. Smith *et al.*, *J. Mol. Biol.* **255**, 494 (1996).
16. H. Schwalbe *et al.*, *Biochemistry* **36**, 8977 (1997).
17. K. B. Wong, S. M. Freund, A. R. Fersht, *J. Mol. Biol.* **259**, 805 (1996).
18. D. Neri, M. Billeter, G. Wider, K. Wuthrich, *Science* **257**, 1559 (1992).
19. G. Saab-Rincon, P. J. Gualfetti, C. R. Matthews, *Biochemistry* **35**, 1988 (1996).
20. I. J. Ropson, C. Frieden, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7222 (1992).
21. M. E. Hodsdon, C. Frieden, *Biochemistry* **40**, 732 (2001).
22. G. Merutka, H. J. Dyson, P. E. Wright, *J. Biomol. NMR* **5**, 14 (1995).
23. D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, *J. Biomol. NMR* **5**, 67 (1995).
24. Supplementary material is available on *Science* Online at www.sciencemag.org/cgi/content/full/295/5560/1719/DC1.
25. H. A. Sober, R. A. Harte, *Handbook of Biochemistry; Selected Data for Molecular Biology* (Chemical Rubber Co., Cleveland, ed. 2, 1970).
26. D. R. Shortle, *Curr. Opin. Struct. Biol.* **6**, 24 (1996).
27. C. C. Blake *et al.*, *Nature* **206**, 757 (1965).
28. H. Schwalbe *et al.*, *Protein Sci.* **10**, 677 (2001).
29. A. Miranker, C. V. Robinson, S. E. Radford, R. T. Aplin, C. M. Dobson, *Science* **262**, 896 (1993).
30. S. E. Radford, C. M. Dobson, P. A. Evans, *Nature* **358**, 302 (1992).
31. D. M. Rothwarf, H. A. Scheraga, *Biochemistry* **35**, 13797 (1996).
32. T. Ueda, H. Yamada, H. Aoki, T. Imoto, *J. Biochem. (Tokyo)* **108**, 886 (1990).
33. T. Ueda, T. Ohkuri, T. Imoto, *Biochem. Biophys. Res. Commun.* **228**, 203 (1996).
34. T. Ueda *et al.*, *J. Mol. Biol.* **235**, 1312 (1994).
35. M. Buck *et al.*, *Biochemistry* **34**, 4041 (1995).
36. C. Lyon, J. Jones, C. Redfield, C. Dobson, P. Hore, *J. Am. Chem. Soc.* **121**, 6505 (1999).
37. P. Hore, S. Winder, C. Roberts, C. M. Dobson, *J. Am. Chem. Soc.* **119**, 5049 (1997).
38. D. R. Booth *et al.*, *Nature* **385**, 787 (1997).
39. M. C. Vaney, S. Maignan, M. RiesKautt, A. Ducruix, *Acta Crystallogr. D Biol. Crystallogr.* **52** (1996).

40. We thank the Massachusetts Institute of Technology (for start-up funds and Rosenblith stipend), the Karl Winnacker-Foundation, the Fonds der Chemischen Industrie, the Bundesministerium für Bildung und Forschung (BMBF), and the Alfred P. Sloan Foundation for support. We dedicate this report to Gerhard Quinkert on the occasion of his 75th birthday.

2 November 2001; accepted 22 January 2002

A Functional Screen for the Type III (Hrp) Secretome of the Plant Pathogen *Pseudomonas syringae*

David S. Guttman,[1]*† Boris A. Vinatzer,[2]*† Sara F. Sarkar,[1] Max V. Ranall,[2] Gregory Kettler,[2] Jean T. Greenberg[2]

Type III secreted "effector" proteins of bacterial pathogens play central roles in virulence, yet are notoriously difficult to identify. We used an in vivo genetic screen to identify 13 effectors secreted by the type III apparatus (called Hrp, for "hypersensitive response and pathogenicity") of the plant pathogen *Pseudomonas syringae*. Although sharing little overall homology, the amino-terminal regions of these effectors had strikingly similar amino acid compositions. This feature facilitated the bioinformatic prediction of 38 *P. syringae* effectors, including 15 previously unknown proteins. The secretion of two of these putative effectors was shown to be type III–dependent. Effectors showed high interstrain variation, supporting a role for some effectors in adaptation to different hosts.

The bacterial type III secretion system is responsible for some of the most devastating diseases of animals and plants. This remarkable system enables a bacterium to strategically inject proteins directly into the host cytoplasm or its extracellular milieu, and thereby subvert host cellular processes (*1*–*3*). The type III apparatus is required for patho-