Morrison, *Cytokine* **4**, 340 (1992); M. Tewari and V. M. Dixit, *Curr. Opin. Genet. Dev.* **6**, 39 (1996).
14. S. Yamaoka *et al.*, *Cell* **93**, 1231 (1998); D. M. Rothwarf, E. Zandi, G. Natoli, M. Karin, *Nature* **395**, 297 (1998); F. Mercurio *et al.*, *Mol. Cell. Biol.* **19**, 1526 (1999).
15. Three 15-cm plates of MEF cells from IKK2$^{+/+}$ and IKK2$^{-/-}$ mice were untreated or treated with human TNF-α (hTNF-α) (10 ng/ml) for 7 min. Whole-cell lysates from each 15-cm plate were prepared and immunoprecipitated with 10 μl of antibody to IKK1 in 1 ml of immunoprecipitation (IP) buffer (*4*). Twenty microliters of protein A were added and samples were rotated for 2 hours at 4°C. The immunoprecipitates were then washed three times with IP buffer. Samples from all three 15-cm plates were pooled into one tube and washed once with kinase assay (KA) buffer (*4*). Sixty micrograms of synthetic peptide in 140 μl of KA buffer were added to the protein A beads and samples were rotated for 6 hours at 4°C. After a brief spin, the eluates were transferred to new tubes. Twenty microliters of eluates were used for each KA reaction or protein immunoblot analysis.
16. A. A. Beg, W. C. Sha, R. T. Bronson, S. Ghosh, D. Baltimore *Nature* **376**, 167 (1995).
17. M. Hirano *et al.*, *J. Biol. Chem.* **271**, 13234 (1996).
18. V. Imbert *et al.*, *Cell* **86**, 787 (1996).
19. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; E, Glu; G, Gly; L, Leu; M, Met; P, Pro; R, Arg; and W, Trp.
20. We thank Y. Marchuk and B. Dominguez for their excellent technical assistance, A. Israel for NEMO antibody, and B. Coyne for her help with this manuscript. We also thank the following people for their expertise, discussions, and sustained interest in our work: Q. Lu, S. Crone, J. C. Izpisúa Belmonte, members of the Verma laboratory, and M. Hoekstra, D. Anderson, and A. Lewis of Signal Pharmaceuticals. Q.L. is supported by a training grant from NIH, K.-F.L. is a Pew Scholar and is supported by NIH and the March of Dimes Foundation, and I.M.V. is an American Cancer Society Professor of Molecular Biology and is supported by the Valley Foundation and grants from NIH.

15 December 1998; accepted 12 March 1999

Evolution of a Protein Fold in Vitro

Matthew H. J. Cordes,[1] Nathan P. Walsh,[1] C. James McKnight,[2] Robert T. Sauer[1]*

A "switch" mutant of the Arc repressor homodimer was constructed by interchanging the sequence positions of a hydrophobic core residue, leucine 12, and an adjacent surface polar residue, asparagine 11, in each strand of an intersubunit β sheet. The mutant protein adopts a fold in which each β strand is replaced by a right-handed helix and side chains in this region undergo significant repacking. The observed structural changes allow the protein to maintain solvent exposure of polar side chains and optimal burial of hydrophobic side chains. These results suggest that new protein folds can evolve from existing folds without drastic or large-scale mutagenesis.

Protein sequences in biological systems evolve by random mutation, including substitutions and en bloc changes resulting from frameshifts or large insertions and deletions. Such genetic changes can result, at least occasionally, in structural evolution to a new or dramatically different three-dimensional (3D) fold (*1*). Little is known, however, about how many or what kind of sequence changes might lead to significant structural changes. Mutagenesis experiments show that limited changes in sequence can have large effects on stability and activity, but generally do not lead to large shifts in structure. For example, highly disruptive mutations such as insertions in elements of regular secondary structure or hydrophobic-to-charged substitutions at core positions lead to only minor structural differences in bacteriophage T4 lysozyme and staphylococcal nuclease, pointing to a strong drive to preserve the basic native fold (*2*, *3*). Here, by contrast, we show that mutations at adjacent positions in the antiparallel β sheet of Arc repressor are sufficient to change the local secondary structure to a right-handed helix without loss of global protein stability or folding cooperativity. This suggests that it is plausible to evolve a new protein fold from an existing fold by the accumulation of simple substitution mutations.

The wild-type Arc repressor homodimer (*4*) contains an antiparallel β sheet consisting of a single strand of sequence Gln^{9}-Phe^{10}-Asn^{11}-Leu^{12}-Arg^{13}-Trp^{14} from each monomer. The odd-numbered side chains are polar, solvent-exposed, and form the surface of Arc that binds operator DNA. The even-numbered side chains are hydrophobic, buried in the protein core, and are crucial for Arc folding and stability (*5*). By interchanging Asn^{11} and Leu^{12}, a surface residue and an adjacent core residue, respectively, we constructed "switch" Arc, a mutant with the same amino acid composition as the wild type but with a different binary pattern of polar and hydrophobic side chains in the β-sheet region.

The purified wild-type and switch Arc proteins differ in their far-ultraviolet circular dichroism (CD) spectra (Fig. 1A), near-ultraviolet CD spectra (Fig. 1B), and fluorescence spectra (Fig. 1C), suggesting that the switch mutations alter the normal Arc fold (*6*). In

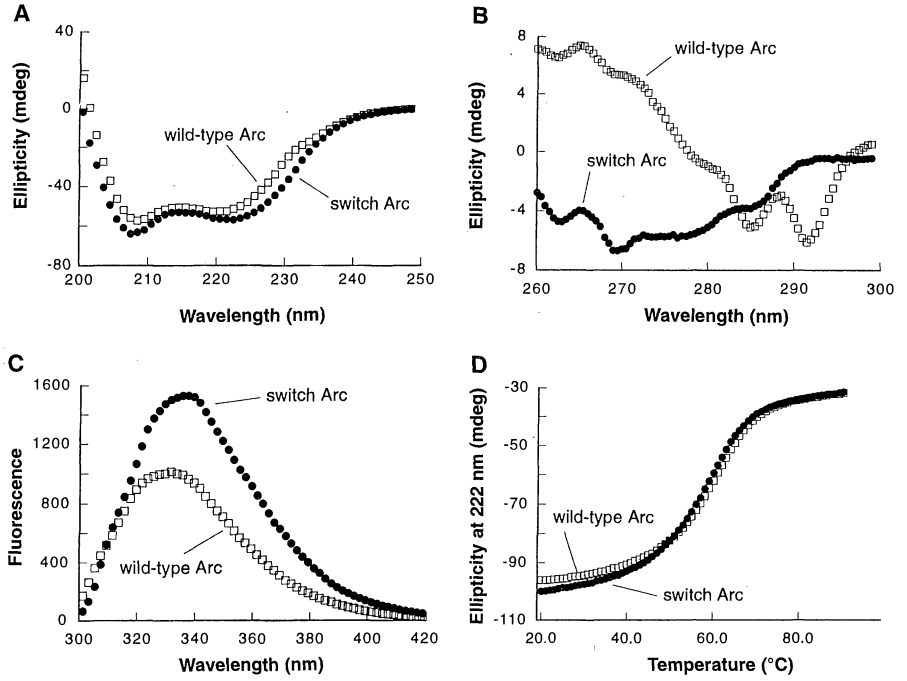


Fig. 1. Biophysical properties of wild-type and switch Arc. (**A**) Far-ultraviolet CD spectra (50 μM protein, 15°C), (**B**) near-ultraviolet CD spectra (100 μM protein, 15°C), (**C**) tryptophan fluorescence emission spectra (50 μM protein, 25°C), and (**D**) CD thermal denaturation curves (10 μM protein). The CD signal is expressed in millidegrees of rotation.

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]Department of Biophysics, Boston University School of Medicine, Boston, MA 02118, USA.

*To whom correspondence should be addressed. E-mail: bobsauer@mit.edu

addition, thermal denaturation experiments (Fig. 1D) show that switch Arc undergoes a cooperative unfolding transition with a midpoint and shape comparable to that of wild-type Arc. The near wild-type stability of switch Arc is also consistent with a change in structure, because native protein folds are strongly destabilized by mutations that replace a hydrophobic core residue with a polar side chain (*3*, *7*).

^{1}H and ^{15}N multidimensional nuclear magnetic resonance (NMR) experiments (*8*) enabled us to characterize the structure of switch Arc in greater detail. The residue-by-residue differences between wild-type and switch Arc H_α chemical shifts are small for most residues (Fig. 2), indicating that much of the wild-type structure is maintained in switch Arc. By contrast, residues 9 to 14 (corresponding to the wild-type β sheet) show large differences, suggesting that the β sheet is replaced with another type of secondary structure in switch Arc. Moreover, backbone and side-chain connectivities from NOESY (nuclear Overhauser enhancement and exchange spectroscopy) spectra of switch Arc reflect the expected secondary structure and tertiary packing of the α-helical portions of the Arc fold, but are strongly inconsistent with the presence of the wild-type β sheet. For example, Fig. 3 shows aliphatic-aromatic side-chain crosspeaks between Leu^{21} and Phe^{45}, which are close in the wild-type structure, but also crosspeaks between Leu^{11} and both Phe^{10} and Trp^{14}, which are not in close contact in the wild-type β sheet.

A structural model of switch Arc was calculated with 99 NOE (nuclear Overhauser enhancement) distance restraints involving residues 7 to 14 (*9*). The minimized average of 14 structures (Fig. 4, C and D) had pairwise root mean square deviations of 0.6 Å for backbone atoms and 1.6 Å for heavy atoms (computed for residues 7 to 14 in both monomers simultaneously). In switch Arc, each strand of the wild-type β sheet is replaced by a right-handed helix spanning residues 9 to 13. The COOH-terminus of this structure is a type I β-turn, with a hydrogen bond between the carbonyl oxygen of Leu^{11} and the amide proton of Trp^{14}, as well as backbone dihedral angles for residues 12 and 13 which fall near (−60, −30) and (−90, 0), respectively. For residues 9 to 11, the resolution of the model is insufficient to distinguish between an α helix, a 3_{10} helix, or additional

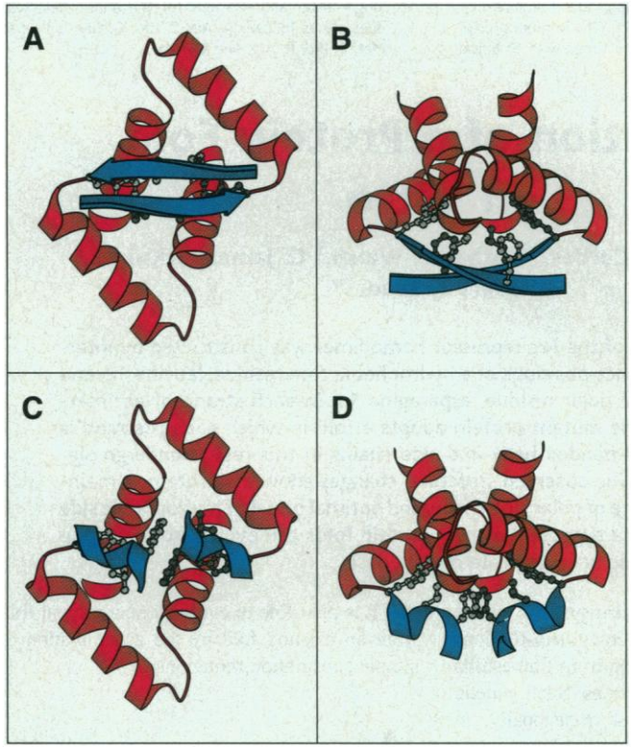
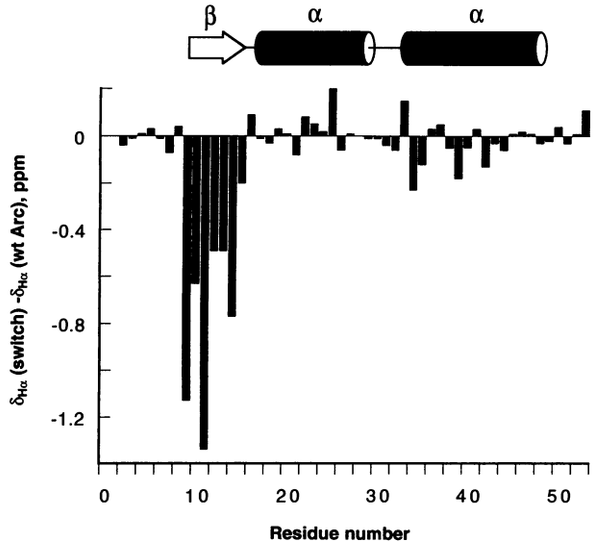


Fig. 4. MOLSCRIPT (*16*) representations of two views of residues 8 to 46 in wild-type Arc (**A** and **B**) and in the minimized average switch Arc model (**C** and **D**). The portion of the fold common to wild-type and switch Arc (residues 15 to 46) is shown in red, whereas the region of structural difference (residues 8 to 14) is shown in cyan. The six hydrophobic core side chains in the region 8 to 14 are shaded in light or dark gray for different monomers.



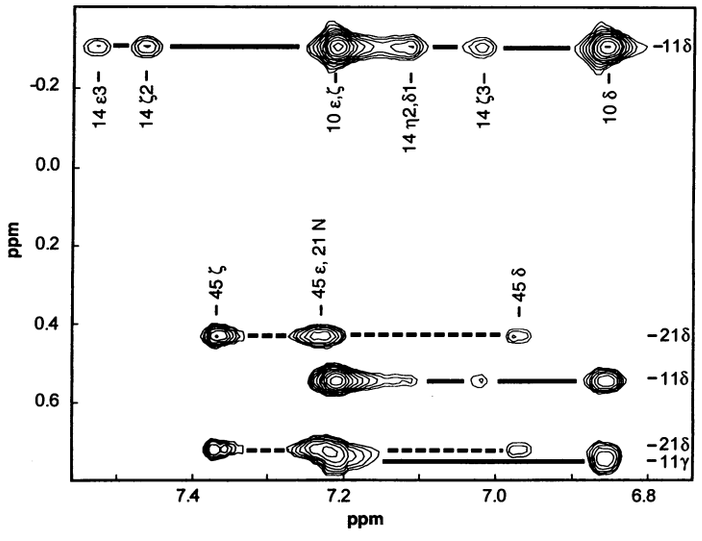


Fig. 2 (left). Residue-by-residue differences in H_α chemical shift for wild-type and switch Arc, suggesting structural changes involving residues 9 to 14 and structural similarity in other regions. **Fig. 3 (right).** Portion of a 200-ms NOESY spectrum of switch Arc, showing crosspeaks between aromatic side-chain resonances and high-field aliphatic resonances. Crosspeaks involving Leu^{21} side-chain protons (inconsistent with structure of residues 9 to 14 in wild type) are connected by solid lines. Crosspeaks involving Leu^{21} side-chain protons (consistent with structure of wild-type helical regions) are connected by dashed lines.

turns. It is clear, however, that the switch Arc backbone structure in this region is radically unlike the β sheet of wild type, in which residues 9 to 14 exist in an extended conformation.

The changes in secondary structure in switch Arc are accompanied by significant alterations in tertiary packing. As shown in Fig. 4, hydrophobic core packing interactions involving residues 9 to 14 of wild-type and switch Arc both involve three side chains from each monomer, but these residues pack in a different spatial arrangement, and switch Arc uses Leu^{11} in its core whereas wild-type Arc uses Leu^{12}. The approximate side-chain positions of each Leu^{12} in wild type are occupied by Phe^{10} from the opposite monomer in switch Arc, and those of each Phe^{10} in the wild type are replaced by Leu^{11} from the same monomer in switch Arc. Despite this extensive repacking, the core of switch Arc appears tightly and specifically structured, as judged by large differences in chemical shifts for the two δ-methyl groups of Leu^{11} (−0.31 and +0.53 parts per million) (Fig. 3) and in NOE crosspeaks from these methyls to other residues.

The results presented above show that interchanging the identities of two adjacent residues in Arc results in a stably folded mutant protein with a well-defined but different local fold in the region of the sequence changes. What is the driving force behind this structural change, and how is it specified by the sequence change? Burial of hydrophobic groups is widely acknowledged as a principal source of protein-folding stability (*10*), whereas burial of polar groups inevitably decreases stability (*3*, *7*). From this perspective, it is notable that despite marked differences in secondary and tertiary structure, both wild-type and switch Arc bury each of the hydrophobic side chains present in residues 9 to 14 and expose each of the polar side chains in this region to solvent. It seems likely, therefore, that the observed changes in secondary structure occur to maintain this optimal partitioning of hydrophobic and polar groups. Protein folds are often constructed from amphipathic elements of regular secondary structure whose local binary patterns of apolar and polar groups reflect their inherent structural periodicity (*11*). In wild-type Arc, the sequence from 9 to 14 has the alternating pattern PHPHPH (where P is polar and H is apolar), which in a β strand leads to partitioning of apolar and polar residues on opposite faces. Switching the positions of Asn^{11} and Leu^{12} changes this pattern to PHHPPH, which reflects the structural periodicity of a helix with three to four residues per turn. Simple changes in binary pattern, subject to conformational and packing constraints, may provide one mechanism for the evolution of protein folds.

Are there features of Arc's β sheet that might help identify regions of other proteins that are also likely to be susceptible to mutationally induced structural changes? The amide protons of Arc's β sheet exchange with solvent at a rate faster than those in Arc's helices (*12*), and this region undergoes modest packing and backbone adjustments upon binding to operator DNA (*4*). These properties suggest that this part of Arc's fold, although crucial for folding stability, is reasonably dynamic and flexible. The β sheet also forms late in the folding of Arc (*13*) and is located near the NH_2-terminus of Arc as the first secondary structure element. An alternative local structure in this region of Arc is therefore only constrained to connect to the remainder of the fold at one end. Replacement of an internal secondary structure element in Arc or another protein might be topologically disallowed if the ends of the new helix or strand cannot make unstrained connections to the rest of the fold. It will be interesting to see whether, in other proteins, COOH- or NH_2-terminal regions that form late in folding or are flexible are also prone to mutationally induced structural changes.

Structural plasticity, the ability of some proteins to adopt alternate folds in response to small changes in sequence or environment, may help permit new and different protein structures to arise through evolution. It may also lead to certain diseases. Inherited prion disorders, such as Creutzfeldt-Jakob's disease (CJD), are linked to familial mutations in the prion protein PrP (*14*). The pathogenesis of CJD is not completely understood, but according to one model (*15*), these sequence changes promote adoption of an alternate fold by PrP in which two α-helical secondary structural elements are converted into a four-stranded antiparallel β sheet. Our work suggests that such structural plasticity may be more common than hitherto assumed and offers one mechanism for a mutationally induced change in a protein fold.

References and Notes

1. D. H. Ohlendorf, W. F. Anderson, M. Lewis, C. O. Pabo, B. W. Matthews, *J. Mol. Biol.* **169**, 757 (1983); A. G. Murzin, *Curr. Opin. Struct. Biol.* **8**, 380 (1998).
2. S. Dao-pin, D. E. Anderson, W. A. Baase, F. W. Dahlquist, B. W. Matthews, *Biochemistry* **30**, 11521 (1991); L. J. Keefe, J. Sondek, D. Shortle, E. E. Lattman, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 3275 (1993); L. J. Keefe, S. Quirk, A. Gittis, J. Sondek, E. E. Lattman, *Protein Sci.* **3**, 391 (1994); D. W. Heinz *et al.*, *J. Mol. Biol.* **236**, 869 (1994); I. R. Vetter *et al.*, *Protein Sci.* **5**, 2399 (1996).
3. W. E. Stites, A. G. Gittes, E. E. Lattman, D. Shortle, *J. Mol. Biol.* **221**, 7 (1991).
4. J. N. Breg, J. H. van Opheusden, M. J. Burgering, R. Boelens, R. Kaptein, *Nature* **346**, 586 (1990); B. E. Raumann, M. A. Rould, C. O. Pabo, R. T. Sauer, *ibid.* **367**, 754 (1994).
5. J. U. Bowie and R. T. Sauer, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2152 (1989); M. E. Milla, B. M. Brown, R. T. Sauer, *Nature Struct. Biol.* **1**, 518 (1994).
6. Construction of switch Arc involved cassette mutagenesis of the *arc-st11* gene of plasmid pET800. Switch Arc-st11 and wild-type Arc-st11 proteins were overexpressed from the *Escherichia coli* strains BL21(λDE3)/pET800-switch and BL21(λDE3)-pLysS/pET800, respectively, and purified to greater than 95% homogeneity by chromatography on nickel-affinity and SP-Sephadex columns [M. E. Milla, B. M. Brown, R. T. Sauer, *Protein Sci.* **2**, 2198 (1993)]. Circular dichroism and fluorescence experiments were performed in 50 mM tris(hydroxymethyl)aminomethane, 250 mM KCl, and 0.2 mM EDTA (pH 7.5) as described [M. E. Milla and R. T. Sauer, *Biochemistry* **34**, 3344 (1995)].
7. W. A. Lim, D. C. Farruggio, R. T. Sauer, *Biochemistry* **31**, 4324 (1992).
8. Uniform ^{15}N-labeled switch Arc-st11 protein was overexpressed from pET800/BL21(λDE3) in M9T minimal medium containing $^{15}NH_4Cl$ (0.8 g/liter) as the sole nitrogen source and purified in the same manner as unlabeled switch Arc-st11. NMR experiments were conducted at 303 K on a sample containing ~4 mM uniform ^{15}N-labeled switch-st11, 20 mM sodium phosphate (pH 4.9), and 10% D_20. ^1H resonance assignments were obtained from analysis of a 3D NOESY (150 ms) spectrum and of 2D TOCSY (total correlation spectroscopy; 50 and 110 ms), DQF-COSY (double quantum-filtered correlated spectroscopy), and NOESY (50 and 200 ms) spectra. Chemical shifts in Fig. 2B were referenced to internal 3-(trimethylsilyl)-propionic acid (TMSP) at 0.0 ppm. For Fig. 2A, the H_α chemical shifts for switch Arc were corrected by −0.12 ppm, because of a systematic difference of +0.12 ppm between our ^1H chemical shift values and those reported for wild-type Arc [J. N. Breg, R. Boelens, A. V. E. George, R. Kaptein, *Biochemistry* **28**, 9826 (1989)], notably in the disordered COOH- and NH-terminal, regions, that is, residues 2 to 6 and 49 to 52 [A. M. J. J. Bonvin *et al.*, *J. Mol. Biol.* **236**, 328 (1994)].
9. Structure calculations were performed with X-PLOR 3.1 [A. T. Brünger, *X-PLOR v3.1 Manual* (Yale Univ. Press, New Haven, CT, 1987)]. For residues 7 to 14, 81 NOE crosspeaks were obtained from a 50-ms 2D NOESY spectrum and were translated into strong (1.8 to 2.8 Å), medium (1.8 to 3.3 Å), medium-weak (1.8 to 3.8 Å), or weak (1.8 to 4.3 Å) distance restraints. Eighteen additional crosspeaks were obtained from a 150-ms 3D NOESY spectrum and were assigned looser restraints (1.8 to 4.0 Å and 1.8 to 6.0 Å) because of the longer mixing time used in this experiment. To simulate the structure of residues 14 to 53 (residue 14 was described by both actual and computed restraints), a list of 1635 distances was generated between pairs of heavy atoms within 4 Å of each other in the wild-type Arc crystal structure, and the interatomic distances were constrained to be within ±1.0 Å of this value. Twenty-eight initial structures with random configurations for residues 1 to 13 were generated by performing a set of calculations with only the simulated restraints for residues 14 to 53. Then, 28 model structures were generated by including the experimental restraints for residues 7 to 14. No restraints were used for residues 1 to 6, which appear to be disordered. Ambiguities in intra- and intermolecular NOE assignment resulting from Arc's twofold symmetry were resolved computationally as described [M. Nilges, *Protein Struct. Funct. Genet.* **17**, 297 (1993)]. Nineteen of 28 structures were accepted with no NOE violations greater than 0.3 Å. The five structures with the highest energy, which also contained backbone dihedral angles in disallowed regions of Ramachandran space, were discarded. The remaining 14 structures were used to generate a minimized average model (see text). Supplementary information (input files for X-PLOR calculations, NOE distant restraints, chemical shift assignments, and the minimized averaged model structure of switch Arc) can be found at www.sciencemag.org/feature/data/986149.shl. PDB files for the 14 model structures are available by anonymous FTP to rosa.mit.edu.
10. K. A. Dill, *Biochemistry* **29**, 7133 (1990).
11. M. F. Perutz, J. C. Kendrew, H. C. Watson, *J. Mol. Biol.* **13**, 669 (1965); M. W. West and M. H. Hecht, *Protein Sci.* **4**, 2032 (1995).
12. M. J. Burgering, M. Hald, R. Boelens, J. N. Breg, R. Kaptein, *Biopolymers* **35**, 217 (1995).
13. J. F. Schildbach, M. E. Milla, P. D. Jeffrey, B. E. Raumann, R. T. Sauer, *Biochemistry* **34**, 1405 (1995).
14. S. B. Prusiner, *Philos. Trans. R. Soc. London Ser. B.* **343**, 447 (1994).
15. Z. Huang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7139 (1994).
16. P. J. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).

17. Supported by NIH grant AI-15706 and by Helen Hay Whitney and Merck-MIT postdoctoral fellowships (M.H.J.C.). We thank G. Rose, G. Tucker-Kellogg, and M. Nohaile for helpful comments and discussion.

28 October 1998; accepted 9 March 1999