Crystal Structure, Folding, and Operator Binding of the Hyperstable Arc Repressor Mutant PL8[†,‡]

Joel F. Schildbach,[§] Marcos E. Milla,[§,||] Philip D. Jeffrey,[⊥] Brigitte E. Raumann,[§] and Robert T. Sauer*[,§]

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10021

Received October 6, 1994; Revised Manuscript Received November 28, 1994[⊗]

ABSTRACT: Arc repressor is a small, dimeric DNA-binding protein that belongs to the ribbon−helix−helix family of transcription factors. Replacing Pro8 at the N-terminal end of the β-sheet with leucine increases the stability of the mutant protein by 2.5 kcal/mol of dimer. However, this enhanced stability is achieved at the expense of significantly reduced DNA binding affinity. The structure of the PL8 mutant dimer has been determined to 2.4-Å resolution by X-ray crystallography. The overall structure of the mutant is very similar to wild type, but Leu8 makes an additional interstrand hydrogen bond at each end of the β-sheet of the mutant, increasing the total number of β-sheet hydrogen bonds from six to eight. Comparison of the refolding and unfolding kinetics of the PL8 mutant and wild-type Arc shows that the enhanced stability of the mutant is accounted for by a decrease in the rate of protein unfolding, suggesting that the mutation acts to stabilize the native state and that the β-sheet forms after the rate-limiting step in folding. The reduced operator affinity of the PL8 dimer appears to arise because the mutant cannot make the new interstrand hydrogen bonds and simultaneously make the wild-type set of contacts with operator DNA.

Increasing the stability of proteins for clinical or industrial applications is one of the goals of protein engineering. Some success has been achieved in stabilizing model proteins through rational design strategies, including introduction of disulfide bonds (Villafranca et al., 1987; Matsumura et al., 1989a,b; Kanaya et al., 1991), charged groups to neutralize helix dipoles (Nicholson et al., 1988, 1991; Kaarsholm et al., 1993), and residues that reduce the torsional freedom of the main chain (Matthews et al., 1987; Ishikawa et al., 1993). Many single amino acid substitutions that cause large increases in protein stability have not been designed, however, but rather have been identified following genetic selections or screens (Shortle & Lin, 1985; Vershon et al., 1986a; Pakula & Sauer, 1989; Das et al., 1989). Studying the structural basis of the enhanced stability of these latter proteins can increase our understanding of the determinants of protein stability and also of the relationship between stability and protein function.

The subject of this paper is a hyperstable variant of the Arc repressor of bacteriophage P22. Arc is a small protein (53 amino acids/monomer) that belongs to the ribbon−helix−helix family of DNA-binding proteins [for a review, see Phillips (1991) and Raumann et al. (1994a)]. During lytic growth of P22, Arc is needed to repress transcription of the *antirepressor* gene, which initiates from the P_{ant} promoter (Susskind & Youderian, 1983). Arc is a dimer in solution (Vershon et al., 1985), but it binds to operator DNA as a tetramer (Brown et al., 1990). In the solution and crystal structures of the Arc dimer, residues 1−6 form a disordered N-terminal arm, residues 8−14 form a two-stranded, antiparallel β-sheet, and residues 15−28 and 32−48 form α-helices (Figure 1a; Breg et al., 1990; Bonvin et al., 1994; C. Kissinger, U. Obeysekare, L. Keefe, B. Raumann, R. Sauer, and C. Pabo, manuscript in preparation). In the crystal structure of the protein−DNA complex, Arc dimers bind to each half of the operator, using β-sheet residues to make base contacts in the major groove and using residues in the N-terminal arm, which folds along the DNA backbone, to make many of the key phosphate contacts (Raumann et al., 1994b). The functional importance of these arm and β-sheet residues of Arc has been verified by mutational and biochemical analyses (Vershon et al., 1986a; Brown et al., 1994).

In this paper, we describe the structure and detailed properties of a mutant (PL8) in which Pro8, the first residue of the Arc β-sheet, is changed to leucine. The PL8 mutant was initially isolated as an inactive variant and was shown to be markedly more stable than wild type in thermal and chemical denaturation studies (Vershon et al., 1986a). Here, we show that (i) the crystal structure of the PL8 mutant protein contains two additional interstrand hydrogen bonds in the β-sheet but is otherwise extremely similar to the wild-type structure, (ii) the stability of the PL8 protein is increased by 2.5 kcal/mol of dimer relative to wild type, (iii) the PL8 mutant refolds at the same rate as wild type but unfolds approximately 100-fold more slowly, and (iv) the tetrameric PL8−operator complex is 4.9 kcal/mol less stable than the wild-type complex. The kinetic and structural results suggest that the formation of the new β-sheet hydrogen bonds occurs after the rate-limiting step in protein folding. The DNA binding results and molecular modeling studies suggest that

[†]Supported by NIH Grants AI-15706 and AI-16892 and by postdoctoral fellowships from the American Cancer Society (J. F. S.) and the Jane Coffin Childs Fund for Medical Research (M.E.M.).

[‡] Atomic coordinates have been deposited in the Brookhaven Protein Data Bank (identification code 1MYK).

* Author to whom correspondence should be addressed.

[§] Massachusetts Institute of Technology.

[||] Present address: Department of Biochemistry, Duke University Medical Center, Durham, NC 27710.

[⊥] Memorial Sloan-Kettering Cancer Center.

[⊗] Abstract published in *Advance ACS Abstracts,* January 15, 1995.

0006-2960/95/0434-1405$09.00/0

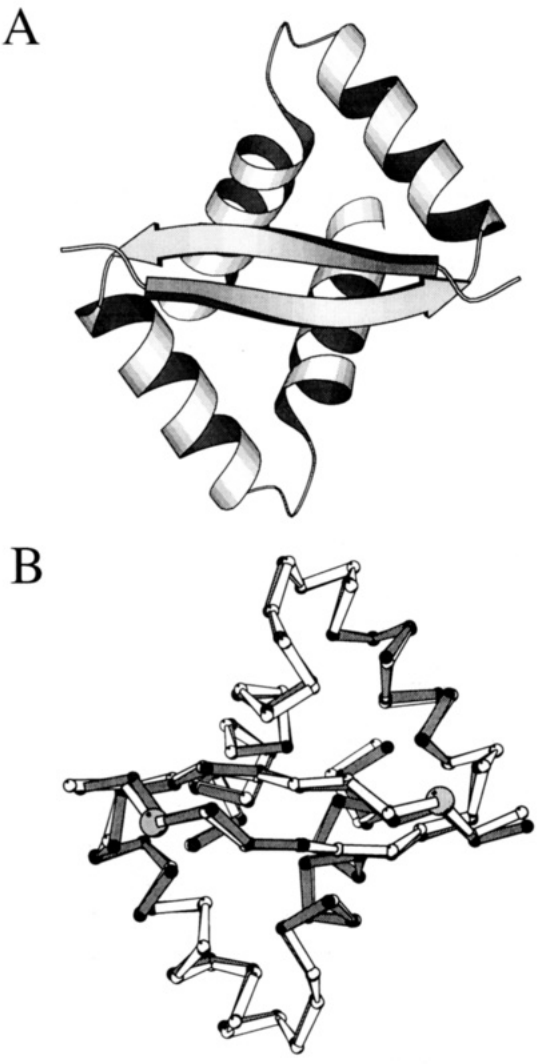

FIGURE 1: (A) Ribbon trace of the polypeptide backbone for the Arc dimer. Each monomer contains a disordered N-terminal arm, a β-strand, and two α-helices. (B) Superposition of α-carbons of residues 6−45 and 6′−45′ of the wild-type Arc dimer (light) and the PL8 mutant dimer (dark). The α-carbons of residue 8 are indicated by larger spheres. Coordinates for wild-type Arc are for dimer 1 from the 2.2-Å crystal structure (C. Kissinger, U. Obeysekare, L. Keefe, B. Raumann, R. Sauer, and C. Pabo, manuscript in preparation). This figure and Figure 2 were prepared using MOLSCRIPT (Kraulis, 1991).

formation of the new hydrogen bonds is incompatible with maintaining the wild-type protein−DNA contacts. As a result, binding of the PL8 mutant to DNA appears to require either that the new hydrogen bonds be broken or that the protein−DNA contacts be distorted.

MATERIALS AND METHODS

Crystallization and Structure Determination. The PL8 mutant was overexpressed and purified from *Escherichia coli* strain UA2F/pTA200-PL8 as described (Vershon et al., 1986a; Bowie & Sauer, 1989a). Although crystals of the PL8 protein were obtained by using the wild-type crystallization conditions (Jordan et al., 1985), they did not form in the wild-type space group and were not well suited for structure determination. Alternative conditions for PL8 crystallization were identified using the sparse matrix screening procedure (Jancarik & Kim, 1991). The best PL8 crystals were obtained by mixing a solution of PL8 protein [15 mg/mL in 30 mM bis-tris propane−HCl (pH 7.4) and 1 mM NaN_3] with an equal volume of precipitant buffer [22−24% PEG-3400 and 7−9% 2-propanol in 100 mM NaHEPES (pH 7.4)] and equilibrating against the precipitant buffer using the hanging drop method. Rod-shaped crystals (0.2 × 0.2 × 0.6 mm) grew within a week at room temperature. The crystals were tetragonal (space group $P4_12_12$; $a = b = 51.8$ Å, $c = 110.7$ Å) and diffracted to a resolution of better than 2.4 Å. Data were collected at 22 °C from a single crystal by using a Rigaku R-AXIS II detector, and were reduced and scaled by using the programs DENZO and SCALEPACK (Z. Otwinowski).

The PL8 structure was solved by molecular replacement, using a portion of the wild-type Arc dimer (residues 9−46 of one monomer and 9′−53′ of the other monomer; C. Kissinger, U. Obeysekare, L. J. Keefe, B. E. Raumann, R. T. Sauer, and C. O. Pabo, manuscript in preparation) as a search model. Identical solutions were obtained by using MERLOT (Fitzgerald, 1988) for rotation and BRUTE (Fujinaga & Read, 1987) for translation searches, and by using X-PLOR (Brünger) for both rotation and translation searches. For rotation searches, data from 6.0−3.5 Å (with MERLOT) or 25.0−2.5 Å (with X-PLOR) were used. For translation searches, data from 8−4 Å (with BRUTE) or 25.0−2.5 Å (with X-PLOR) were used. Following the translation search, the correlation coefficient was 0.59 and the crystallographic R-factor (R_{cryst}) was 0.43 (data from 8.0−3.5 Å). The asymmetric unit contains a single PL8 dimer. Prior to refinement, 10% of the data set was set aside for cross-validation (calculation of R_{free}; Brünger, 1992). Throughout refinement, decisions to accept or reject structural changes were made on the basis of the effects on R_{free}. Structure refinement was performed using X-PLOR and utilized rigid-body refinement, Powell energy minimization, simulated-annealing refinement, and restrained individual atomic B-factor refinement. In the early stages of refinement, data from 6.0−2.5 Å were used; in the later stages, data from 6.0−2.4 Å were used. A series of 20 overlapping simulated-annealing omit maps (each deleting five residues from the model) were calculated with X-PLOR and used to direct rebuilding of the model. A total of 16 water molecules were added to the structure after examination of $F_o - F_c$ maps. Manual rebuilding and addition of water molecules was performed using the program O (T. A. Jones and M. Kjeldgaard). In the final step of refinement, the values of F_o were scaled locally to F_c using MAXSCALE (M. Rould) and the model was subjected to Powell minimization. This reduced R_{cryst} from 0.227 to 0.217 and R_{free} from 0.289 to 0.277. Atomic coordinates have been deposited in the Brookhaven Protein Data Bank (identification code 1MYK).

Protein Stability and Folding Kinetics. Protein stability was measured by GuHCl denaturation and thermal denaturation. Unless otherwise noted, equilibrium denaturation experiments and kinetic experiments were performed in a standard buffer containing 50 mM Tris-HCl (pH 7.5), 100 mM KCl, and 0.2 mM EDTA. For thermal melts, protein at a concentration of 10 μM was heated in 1 °C steps with a 1-min equilibration at each temperature using a Hewlett-Packard 89100A temperature controller. The CD ellipticity at 222 nm was averaged for 30 s at each temperature using an AVIV 60DS CD spectrometer. GuHCl denaturation studies were performed at 20 or 25 °C using 5 μM protein. The extent of denaturation at different denaturant concentrations was monitored by changes in fluorescence emission

(325 nm) or circular dichroism (230 nm). Denaturation reactions were greater than 95% reversible and were fit by a two-state model in which the native dimer is in equilibrium with two denatured monomers [for evidence supporting the two-state model for wild-type Arc unfolding and equations and details of the fitting procedures, see Bowie and Sauer (1989b) and Milla et al. (1993)]. For the PL8 mutant, similar ΔG_u values were calculated in denaturation experiments monitored by either circular dichroism or fluorescence (Table 3), providing support for the assumption that the equilibrium denaturation of the PL8 protein can be described by a two-state model. The stability (ΔG_u) of wild-type Arc or the PL8 protein at a temperature (T) other than the T_m was calculated using the equation

$$\Delta G_u = \Delta H_u - (T/T_m)(\Delta H_u + RT_m \ln(P_t)) + \Delta C_p(T - T_m - T\ln(T/T_m))$$

where ΔH_u is the enthalpy of unfolding/dissociation for the protein at its T_m, P_t is the total protein concentration in monomer equivalents, and ΔC_p has a value of 1.31 kcal/mol (Becktel & Schellman, 1987; Milla et al., 1993, 1994).

Rates of protein refolding/association and unfolding/dissociation were determined in GuHCl-jump or pH-jump experiments, monitored by changes in tryptophan fluorescence using an Applied Photophysics DX17-MV stop-flow instrument. pH-jumps were initiated by mixing acid-denatured Arc [20 μM in 10 mM phosphoric acid (pH 2.0), 100 mM KCl, and 0.2 mM EDTA] with an equal volume of refolding buffer [100 mM Tris-HCl (pH 8.0), 100 mM KCl, and 0.2 mM EDTA] containing 0.15−0.75 M GuHCl. Following mixing, the pH was 7.5. GuHCl-jump refolding experiments were initiated by mixing the PL8 mutant (20 μM in standard buffer plus 5 M GuHCl) with 5 vol of buffer containing 0−1.5 M GuHCl. Unfolding reactions were initiated by mixing Arc or PL8 in standard buffer with buffer containing sufficient GuHCl to achieve a final concentration of 2−5 M. Kinetic data for unfolding were fit as single exponentials; refolding data were fit as relaxations with explicit inclusion of the unfolding rate as described (Milla & Sauer, 1994). The ΔG_u values expected on the basis of the observed folding and unfolding rate constants [calculated as $-RT\ln(2k_u/k_f)$] were close to the ΔG_u values observed in equilibrium experiments [12.35 vs 12.37 kcal/mol of dimer for PL8; 9.81 vs 9.97 kcal/mol of dimer for wild type]. This correspondence of kinetic and equilibrium values provides further evidence that denaturation of both proteins can be well described as a two-state reaction.

Linear and nonlinear least-squares fitting were performed using the programs Excel 3.0 (Microsoft) and NonLin (Johnson & Frasier, 1985; Brenstein, 1989), respectively. Standard errors for fitted values were propagated in subsequent calculations as described by Bevington (1969).

Operator-Binding Assays. The affinities of wild-type Arc and the PL8 mutant for the 27 base pair *arc* operator DNA fragment O1 were determined by gel shift assays, at 20 °C, in a buffer containing 10 mM Tris-HCl (pH 7.5), 100 mM KCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 0.02% Nonidet P-40, and 100 μg/mL bovine serum albumin as described (Brown et al., 1990; Brown & Sauer, 1993). Three to five sets of independent binding curves were determined for each protein and were averaged. In fitting the DNA binding curves, the concentrations of free Arc monomers [U] or free Arc dimers [A_2] were calculated from the total protein concentration (P_t) and the equilibrium constant for dissociation/unfolding of the dimer (K_u) using the following equations:

$$[U] = (-K_u + \sqrt{K_u^2 + 8K_u \cdot P_t})/4$$

$$[A_2] = [U]^2/K_u$$

These calculations ignore the concentration of the protein−DNA complex, but sufficiently small quantities of radiolabeled operator DNA were used in each assay (≈20 pM) to justify this simplification. The rates of dissociation of the tetrameric PL8−DNA or wild-type Arc−DNA complexes were measured by gel shift assays using the same buffer and conditions used for the equilibrium binding assays (Brown et al., 1990). Reassociation of the dissociated protein to labeled operator DNA was blocked by addition of a large excess of unlabeled operator DNA at time zero.

Repression assays *in vivo* were performed in *Escherichia coli* strain UA2F, which contains a chromosomal fusion of the P_{ant} promoter to the *lacZ* gene encoding β-galactosidase (Vershon et al., 1986b). Arc was expressed from the P_{tac} promoter of plasmid pTA200, and the PL8 mutant was expressed from an otherwise isogenic derivative of this plasmid. A derivative of pTA200 in which the *arc* gene and a portion of the *tet* gene were deleted was constructed to serve as a negative control. UA2F cells containing pTA200, pTA200-PL8, or the control plasmid were grown at 25, 30, or 37 °C in LB broth plus 100 μg/mL ampicillin and 50 μg/mL kanamycin to an A_{600} between 0.4 and 0.6. At this point the cells were chilled on ice, 0.1 mL of cells was added to 0.9 mL of Z buffer (Miller, 1972), and 50 μL of chloroform and 20 μL of 0.1% SDS were added to lyse the cells. β-Galactosidase activity was then assayed by hydrolysis of *o*-nitrophenyl-β-D-galactoside as described by Miller (1972). A second set of experiments was performed under identical conditions, but when cells reached an A_{600} of approximately 0.1, IPTG was added to a final concentration of 100 μg/mL to induce expression of Arc or PL8 from the P_{tac} promoter. Growth was then continued until A_{600} reached 0.4−0.6 (≥2 h). For each set of growth conditions, five independent determinations of enzyme activity were performed, and the resulting activities were averaged.

RESULTS

Determination of the PL8 Mutant Structure. To examine the structural basis for the observed stability and DNA-binding phenotypes of the PL8 mutant [see below and Vershon et al. (1986a)], we crystallized the protein and determined its structure by molecular replacement. The structure was refined at a resolution of 2.4 Å to an R_{cryst} of 0.217 and an R_{free} of 0.277. The statistics of the data processing and refinement are shown in Table 1. The final structure has good geometry for bond angles and bond lengths and has no ϕ,ψ angles in unfavorable regions of the Ramachandran plot.

As shown by backbone superposition (Figure 1B), there are no large differences between the PL8 dimer structure and the wild-type Arc dimer structure. The rms deviation in the positions of the main-chain atoms of residues 8−45 and 8′−45′ is 0.48 Å. Examination of the PL8 structure around the sites of the substitution (Figure 2) reveals that

Table 1: Statistics for Crystallographic Data Processing and Refinement

measured reflns (6.0−2.4 Å)	26 922
unique reflns (6.0−2.4 Å)	5340
data completeness (6.0−2.4 Å)	90.4%
data completeness (2.5−2.4 Å)	56.6%
R_{merge}[a]	0.094
R_{cryst}[b] (90% of all data from 6.0−2.4 Å)	0.217
R_{free}[c] (10% of all data from 6.0−2.4 Å)	0.277
total number of protein atoms	764
no. of waters	16
residues in model	6−52, 6′−50′
residues not in model	1−5, 53, 1′−5′, 51′−53′
side chains built as alanine	6, 6′, 50′
RMSD from ideal bond lengths	0.013 Å
RMSD from ideal bond angles	1.492°
average *B*-factor	33.5 Å^2
RMSD in *B*-factors between	
bonded main-chain atoms	1.50 Å^2
all bonded atoms	3.19 Å^2

[a] $R_{merge} = \sum|I - \langle I\rangle|/\sum I$. [b] $R_{cryst} = \sum_{h,k,l}|F_o(h,k,l) - F_c(k,h,l)|/\sum_{h,k,l} F_o(h,k,l)$. [c] R_{free} is calculated as R_{cryst}, except using the 10% of the data against which the model has not been refined.

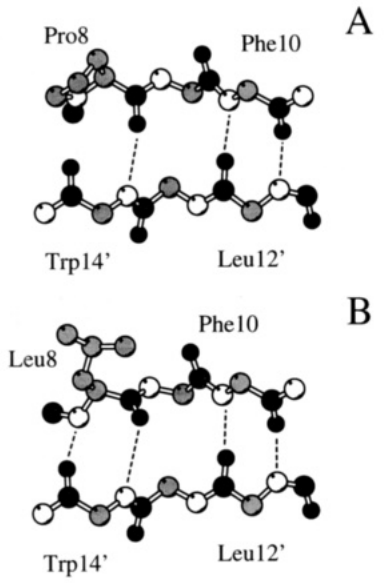

FIGURE 2: (**A**) Interstrand β-sheet hydrogen bonds (dashed lines) near Pro8 in the wild-type structure. (**B**) Interstrand β-sheet hydrogen bonds near Leu8 in the PL8 structure. Note that the mutant structure contains an additional hydrogen bond between the amide of Leu8 and the carbonyl oxygen of Trp14′, compared to wild type.

the amide nitrogen of Leu8 is an average of 2.8 Å from the carbonyl oxygen of Trp14, on the opposite strand of the antiparallel β-sheet. This distance and the N−H−O angle of 149° are consistent with formation of a hydrogen bond between the two groups (Table 2). By contrast, in the wild-type structure, the imide nitrogen of Pro8 is an average of 3.4 Å from the carbonyl oxygen of Trp14, and there is no hydrogen to participate in a hydrogen bond. As a result, although the β-sheets in the mutant and wild-type structures contain the same number of residues, there are eight backbone hydrogen bonds in the PL8 dimer structure and only six in the wild-type dimer structure. We presume that the additional hydrogen bonds in the mutant structure account, at least in part, for its enhanced stability to denaturation. Other structural changes in the mutant include minor alterations in the geometries of other hydrogen bonds in the β-sheet (e.g., the hydrogen bond between the backbone

Table 2: Geometry of β-Sheet Hydrogen Bonds in the PL8 and Wild-Type Arc Dimer Structures[a]

donor	acceptor	av distance (Å)	av N−H−O angle (°)
PL8			
Leu8 N	Trp14′ O	2.78 ± 0.13	149 ± 23
Leu8′ N	Trp 14 O		
Phe10 N	Leu12′ O	3.07 ± 0.01	157 ± 6
Phe10′ N	Leu12 O		
Leu12 N	Phe10′ O	2.90 ± 0.07	167 ± 2
Leu12′ N	Phe10 O		
Trp14 N	Leu8′ O	3.50 ± 0.03	160 ± 2
Trp14′ N	Leu8 O		
wild type			
Pro8 N	Trp14′ O	3.41 ± 0.11	
Pro8′ N	Trp14 O		
Phe10 N	L12′ O	2.96 ± 0.10	159 ± 2
Phe10′ N	Leu12 O		
Leu12 N	Phe10′ O	2.85 ± 0.09	166 ± 3
Leu12′ N	Phe10 O		
Trp14 N	Pro8′ O	3.21 ± 0.09	146 ± 5
Trp14′ N	Pro8 O		

[a] The hydrogen bond distances and angles for PL8 are averages of the two noncrystallographic symmetry-related interactions in the dimer (e.g., Leu8 N → Trp14′ O and Leu8′ N → Trp14 O). For wild type, the distances and angles are averages of the four interactions observed in the two dimers in the wild-type asymmetric unit. Errors given are calculated as

$$\sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n}}$$

where x values are the experimental values, and n is the number of values.

amide of Trp14 and the carbonyl oxygen of residue 8 is an average of 0.29 Å longer in the PL8 mutant; Table 2) and some differences in the packing of the mutant Leu8 side chain compared with the wild-type Pro8 side chain. However, the side-chain atoms of Leu8 have *B*-factors roughly 80% higher than average for the PL8 structure, suggesting that these atoms do not engage in tight packing interactions that might be expected to stabilize the protein. There are also changes in the conformations of several surface side chains between the mutant and wild-type structures, but these appear to arise from changes in crystal packing contacts and not because of the mutation. Moreover, mutant studies indicate that none of these surface side chains contribute significantly to Arc stability (Milla et al., 1994).

Stability and Folding Kinetics. The Arc unfolding reaction (shown below) is well described as a concerted transition between native dimers and denatured monomers (Bowie & Sauer, 1989b; Milla & Sauer, 1994):

$$A_2 \overset{K_u}{\rightleftharpoons} 2U$$

In thermal denaturation experiments performed at pH 7.5, 100 mM KCl, and a protein concentration of 10 μM, the T_m of wild-type Arc was 53.4 °C, while the T_m of the PL8 protein was 67.3 °C (Table 3; Figure 3A). By calculating the stability of each protein at the other protein's T_m, the PL8 protein was found to be 2.55 kcal/mol of dimer more stable than wild type at 53.4 °C and 2.78 kcal/mol of dimer more stable than wild type at 67.3 °C. In GuHCl denaturation studies (see, for example, Figure 3B), the PL8 mutant was found to be more stable than wild type by an average of 2.5 kcal/mol of dimer at 25 °C and 2.4 kcal/mol of dimer

Table 3: Parameters for Thermal and GuHCl denaturation of Arc and the PL8 Mutant [50 mM Tris-HCl (pH 7.5) and 100 mM KCl][a]

	concn (μM)	T_m °C	ΔH_u (kcal/mol of dimer)	$\Delta\Delta G_u$[b] (kcal/mol of dimer)
Thermal Denaturation				
Arc	10	53.4(±0.2)	55.2(±1.4)	
PL8	10	67.3(±0.3)	71.8(±3.0)	2.78 (67.3°); 2.55 (53.4°)

	T (°C)	ΔG_u (kcal/mol of dimer)	m [kcal/(mol M)]	
GuHCl Denaturation				
Arc (F)	25	10.07(±0.35)	3.03(±0.24)	
Arc (CD)	25	10.21(±0.34)	3.01(±0.25)	
Arc (F)	20	9.97(±0.01)	2.98(±0.01)	
PL8 (F)	25	12.56(±0.09)	3.10(±0.04)	2.49 (25 °C)
PL8 (CD)	25	12.75(±0.18)	3.16(±0.13)	2.54 (25 °C)
PL8 (F)	20	12.37(±0.11)	2.96(±0.05)	2.40 (20 °C)

[a] T_m and ΔH_u values are the average of two independent determinations. GuHCl denaturation was performed at 20 or 25 °C and was assayed by fluorescence (F) or circular dichroism (CD) as noted. m is the slope of a plot of $-\Delta G_u$ vs GuHCl concentration. Errors for GuHCl-denaturation parameters are 67% confidence limits from nonlinear least squares fitting of single denaturation curves. Errors for thermal parameters are the larger of the 67% confidence limits calculated for each fit or the error of the replicates. [b] $\Delta\Delta G_u = \Delta G_u$(PL8) $- \Delta G_u$(wild type).

at 20 °C (Table 3). Hence, the stability difference between wild-type Arc and the PL8 mutant is relatively constant over a wide temperature range.

In principle, the enhanced stability of the PL8 protein relative to wild type could be caused by an increase in the rate of protein folding, a decrease in the rate of protein unfolding, or a combination of the two. The second-order rate constants ($\approx 8 \times 10^6$ M^{-1} s^{-1}) for refolding/association of wild-type Arc and the PL8 protein at 20 °C in the absence of denaturant were found to be the same within experimental error (Figure 4A; Table 4). As a result, faster folding does not contribute to the enhanced stability of the PL8 mutant. By contrast, at 20 °C the PL8 dimer unfolds/dissociates at a rate almost 100 times slower than wild type (Figure 4B; Table 4). This large difference in unfolding rates indicates that the enhanced stability of PL8 is caused by an increase in the energy barrier between the native state and the transition state for protein folding/unfolding. The observed decrease in the unfolding rate for PL8 predicts a change in $\Delta\Delta G_u$ of 2.6 kcal/mol of dimer, very close to the values measured in equilibrium experiments. Thus, the simplest model that is consistent with both the thermodynamic and the kinetic data is that the PL8 mutation stabilizes the native protein by roughly 2.5 kcal/mol of dimer, without significantly changing the energies of the transition state or the denatured state.

Binding to Operator DNA. The binding reaction of Arc to operator DNA is thermodynamically linked to the protein folding/dimerization reaction:

$$A_4O \underset{\Delta G_d}{\overset{K_2}{\rightleftharpoons}} 2A_2 + O \underset{2\Delta G_u}{\overset{K_u^2}{\rightleftharpoons}} 4U + O$$

To assess the effects of the PL8 mutation on DNA binding *in vitro*, gel shift experiments were performed. Figure 5 shows DNA binding curves plotted as a function of both the free dimer concentration and the free monomer concentration. Inspection of Figure 5A shows that the PL8 dimer

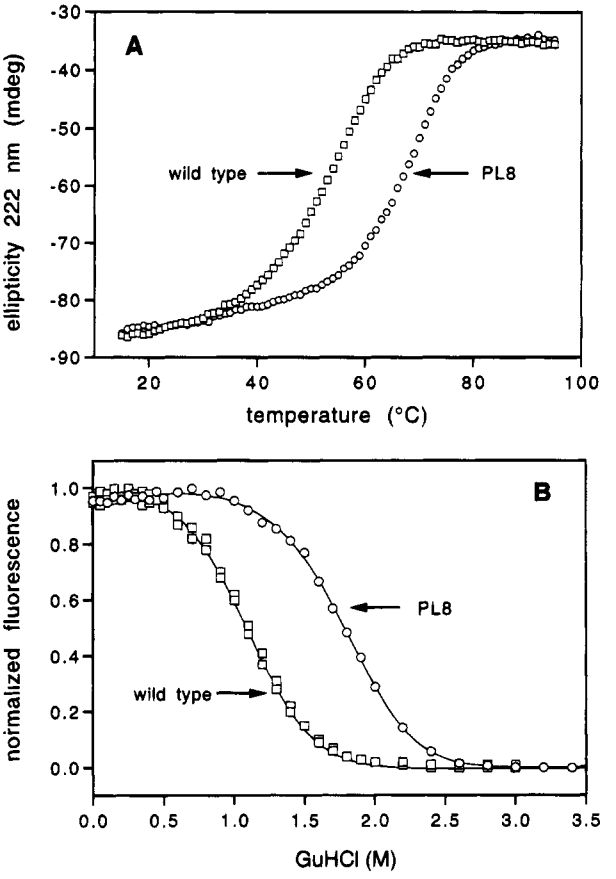

FIGURE 3: Assays of equilibrium stability of wild-type Arc and the PL8 mutant by thermal denaturation (A) and by GuHCl denaturation at 25 °C (B). The protein concentration is 10 μM in panel A and 5 μM in panel B. The normalized fluorescence (325 nm) was calculated as $(F - F_{min})/(F_{max} - F_{min})$, where F is the fluorescence at each denaturant concentration, and F_{max} and F_{min} are the maximal and minimal fluorescence, respectively. The solid lines in panel B are the expected denaturation curves calculated using the ΔG_u and m values listed in Table 3.

has significantly reduced affinity for operator DNA compared to the wild-type dimer. K_2, the equilibrium constant for dissociation of the protein tetramer–operator complex to free operator and two dimers, changes from $(3.6 \pm 2.8) \times 10^{-23}$ M^2 for wild type to $(1.7 \pm 0.7) \times 10^{-19}$ M^2 for PL8. This increase in K_2 corresponds to a 4.9 kcal/mol reduction in the stability of the tetrameric PL8–operator complex compared to the wild-type complex. However, as shown in Figure 5B, PL8 monomers and wild-type monomers bind operator DNA equally well. Thus, the enhanced stability of two PL8 dimers relative to two wild-type dimers [$2\Delta\Delta G_u$ = 2 × 2.5 kcal/mol of dimer = 5.0 kcal/mol of tetramer] compensates almost exactly for the reduction in the stability of the PL8–operator complex [$\Delta\Delta G_d = -4.9$ kcal/mol of tetramer]. In kinetic experiments, the PL8–operator complex was found to be greater than 90% dissociated 30 s after the addition of competitor DNA, suggesting that the complex half-life is less than 10 s. The wild-type complex, with a half-life of approximately 75 min (Brown et al., 1990), is far more stable than the PL8 complex.

To test the activities of Arc and the PL8 mutant in the cell, we monitored repression of a P_{ant}–β-galactosidase fusion at different temperatures and at different protein expression levels (± induction of the P_{tac} promoter with IPTG). At low levels of expression (cells grown without

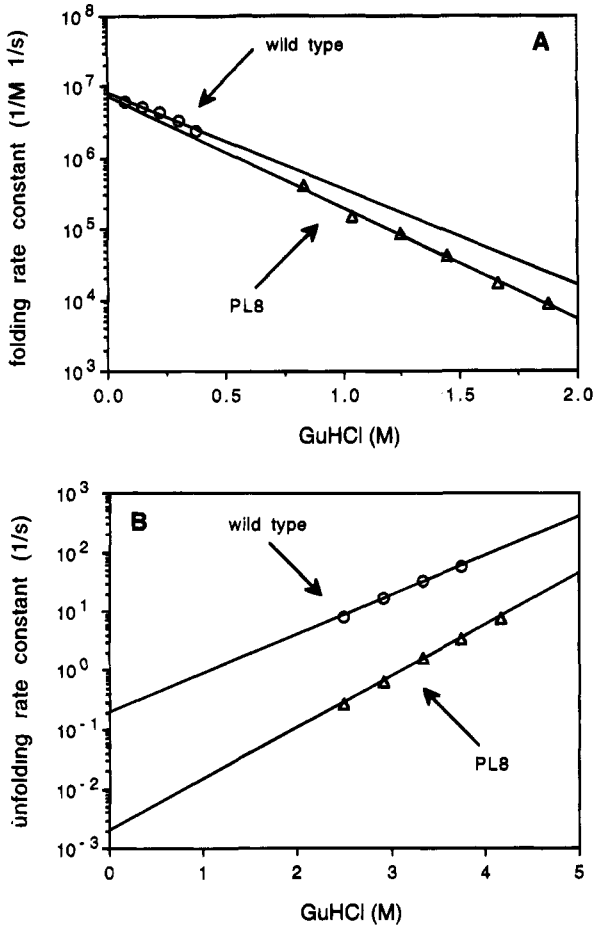

FIGURE 4: Dependence of the rate constants for folding/association (A) and unfolding/dissociation (B) on the concentration of GuHCl for wild-type Arc and the PL8 mutant. Experiments were performed at 20 °C, pH 7.5, and 100 mM KCl.

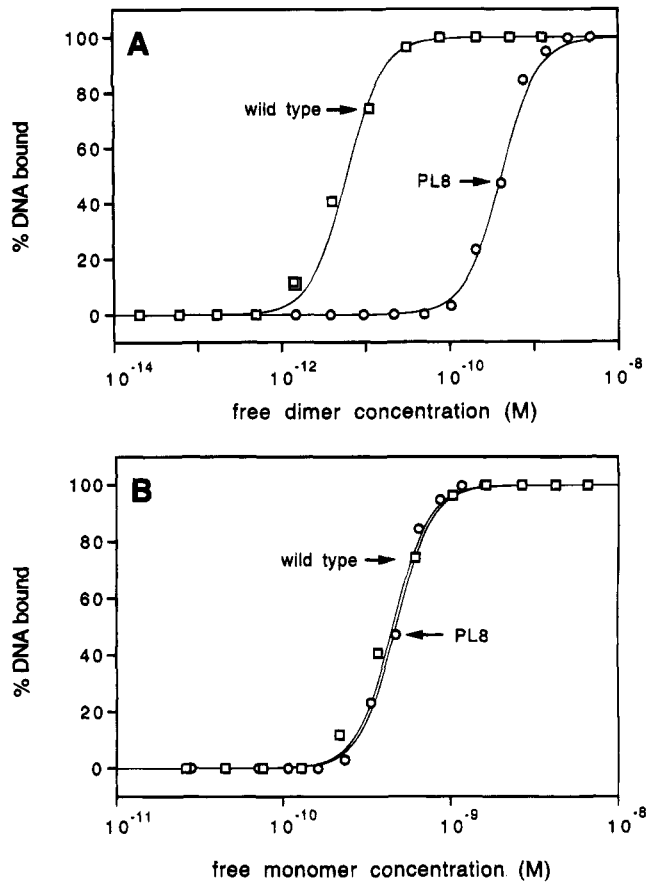

FIGURE 5: Binding of wild-type Arc and the PL8 mutant to *arc* operator DNA. The same data are shown in both panels. The fraction of DNA bound is plotted against the free dimer concentration in (A) and against the free monomer concentration in (B). The solid lines are theoretical curves for the reactions $2A_2 + O \rightleftharpoons A_4O$ ($K_2 = [A_2]^2[O]/[A_4O]$) in (A) and $4U + O \rightleftharpoons A_4O$ ($K_u^2K_2 = [U]^4[O]/[A_4O]$) in (B) ($K_2 = 3.6 \times 10^{-23}$ M^2 and $K_u = 3.4 \times 10^{-8}$ M for wild type; $K_2 = 1.7 \times 10^{-19}$ M^2 and $K_u = 5.4 \times 10^{-10}$ M for PL8). The experimental data shown for wild type are averages of three independent experiments; the data shown for PL8 are averages of five independent experiments.

IPTG), PL8 shows no repressor activity since as much β-galactosidase is expressed as in control cells containing no Arc repressor (Table 5). Induction of PL8 expression by IPTG results in approximately 50% repression of β-galactosidase transcription. Wild-type Arc, however, gives approximately 70% repression when uninduced, and 85% repression when induced, and is therefore far more active than PL8. The PL8 mutant and wild-type Arc were expressed at the same level as assayed by SDS gel electrophoresis of induced cell lysates. Moreover, IPTG induction increases expression of both proteins at least 10-fold, and yet Arc at uninduced levels is as effective a repressor as PL8 at induced levels. Taken together, the activity and expression data indicate that the PL8 mutant has significantly reduced activity in the cell.

We were not successful in obtaining cocrystals of the PL8–operator complex and consequently have no direct information concerning possible changes in the structure of the mutant complex that might explain its reduced stability. We did, however, model the changes observed in the crystal structure of the PL8 dimer into the wild-type tetramer–operator complex. These studies suggest that the additional hydrogen bonds at the ends of the PL8 β-sheet cannot be maintained without distortions in the protein–DNA contacts. In fact, the average distance between the imide nitrogen of Pro8 and the carbonyl oxygen of Trp14 increases from 3.4 Å in the Arc dimer structure to 3.9 Å in the wild-type protein–DNA complex (Raumann et al., 1994b). As a result, the new hydrogen bonds between Leu8 and Trp14 may need to be broken in the complex to allow formation of a protein–

Table 4: Unfolding (k_u) and Refolding (k_f) Rate Constants for Wild-Type Arc and the PL8 Mutant (20 °C, pH 7.5, 100 mM KCl)[a]

	k_u (s^{-1})	m_u [kcal/(mol M)]	k_f (M^{-1} s^{-1})	m_f [kcal/(mol M)]
Arc	$(1.9 \pm 0.7) \times 10^{-1}$	0.88(±0.06)	$(8.4 \pm 0.8) \times 10^6$	−1.83(±0.21)
PL8	$(2.1 \pm 0.7) \times 10^{-3}$	1.14(±0.05)	$(7.4 \pm 1.3) \times 10^6$	−2.10(±0.07)

[a] Determined from the experiments shown in Figure 5. m_f is the slope of a plot of $RT \ln(k_f)$ vs [GuHCl], and m_u is the slope of a plot of $RT \ln(k_u)$ vs [GuHCl]. The errors shown are for linear extrapolation of the data to 0 M GuHCl and should be taken as minimum estimates since there is also error associated with determination of k_u and k_f values at each GuHCl concentration. The zero-denaturant values of k_u and k_f reported here for wild-type Arc are within experimental error of values determined previously from urea-jump and pH-jump experiments ($k_f = 9.0 \times 10^6$ M^{-1} s^{-1}; $k_u = 0.2$ s^{-1}; Milla & Sauer, 1994).

Table 5: Repression of the P_{ant}–β-Galactosidase Fusion in Strain UA2F by Wild-Type Arc and the PL8 Mutant[a]

protein	temp (°C)	IPTG	β-gal activity (Miller units)[b]
Arc	37	no	473(±29)
PL8	37	no	1616(±172)
control	37	no	1392(±13)
Arc	37	yes	207(±39)
PL8	37	yes	873(±24)
control	37	yes	1560(±41)
Arc	30	no	652(±23)
PL8	30	no	2091(±107)
control	30	no	2061(±100)
Arc	30	yes	320(±151)
PL8	30	yes	1335(±58)
control	30	yes	1902(±88)
Arc	25	no	1266(±81)
PL8	25	no	3941(±253)
control	25	no	4212(±177)

[a] Wild-type Arc was expressed from pTA200, and the PL8 mutant was expressed from pTA200-PL8. The control cells contain a derivative of pTA200 in which the *arc* gene is deleted. At 25 °C, the cells grew too slowly in the presence of IPTG to allow reliable assays. [b] Miller (1972).

DNA interface similar to wild type. Alternatively, these hydrogen bonds may be maintained but at the expense of optimal protein–DNA contacts.

DISCUSSION

Vershon et al. (1986a) originally isolated the PL8 mutant and found that the purified protein was more stable to denaturation than wild-type Arc and had reduced affinity for operator DNA. The studies performed here confirm these results and also extend them in several ways, including determining the free energy changes in the stability of the dimer and the tetramer–DNA complex. Thus, we find that the PL8 dimer is approximately 2.5 kcal/mol of dimer more stable than wild type, but the tetrameric PL8-operator complex is almost 4.9 kcal/mol of tetramer less stable than the wild-type complex. Clearly, the enhanced stability of the mutant PL8 dimer is achieved at the expense of reduced DNA binding affinity.

Comparison of the crystal structures of the PL8 mutant and wild-type Arc reveals that the backbone amide of Leu8 makes a new hydrogen bond at each end of the β-sheet; the NMR secondary structure assignments for the PL8 mutant are also consistent with the presence of an additional hydrogen bond (Zagorski et al., 1989). The geometries of some other β-sheet hydrogen bonds are also altered in the PL8 structure (Table 2), but most of these changes would be expected to result in weaker rather than stronger interactions. Similarly, the Pro8→Leu substitution would be expected to increase the entropy of unfolding because the Φ angle of proline is more highly restricted than that of leucine, but this effect should decrease rather than increase the stability of the PL8 mutant (Matthews et al., 1987). It might be argued that improved packing and/or burial of hydrophobic surface by the Leu8 side chain causes some of the observed stabilization. However, the high *B*-factors of the Leu8 side chain atoms in the mutant structure make this seem unlikely. Moreover, Ala8, a residue whose side chain would have fewer packing interactions and bury less surface, but whose main chain could make the same new hydrogen bonds, stabilizes Arc as much as Leu8 [$\Delta\Delta G_u$ = 2.9 kcal/mol of dimer for PA8; Milla et al., 1994]. Thus, we believe that the new Leu8-mediated hydrogen bonds are the most likely cause of the enhanced stability of the PL8 mutant. If the stabilization of the PL8 mutant were due solely to these additional hydrogen bonds, then each interaction in the dimer would contribute an average of 1.25 kcal/mol, a value similar to estimates of 1.3 (±0.6) kcal/mol for hydrogen-bonding interactions in other systems (Shirley et al., 1992, and references therein). It is not possible to determine whether the PL8 mutant is more stable because it contains additional hydrogen bonds or whether wild-type Arc is less stable because it contains unpaired polar groups in solvent-inaccessible environments (i.e., the carbonyl oxygen of Trp14 and the imide nitrogen of Pro8). The presence of such unpaired polar groups within the protein interior is almost certainly destabilizing, since hydrogen bonds between these groups and solvent in the denatured state are not replaced in the native protein.

The slow step in Arc folding is bimolecular, indicating that the transition state is dimeric (Milla & Sauer, 1994). We have found that the enhanced equilibrium stability of the PL8 variant arises because the mutant protein folds at the same rate as wild type but unfolds roughly 100-fold more slowly. This finding suggests that the new hydrogen bonds in the PL8 mutant form late in the folding process (after the transition state or rate-limiting step) and break early in the unfolding process (before the transition state or rate-limiting step). Moreover, because we assume that formation of the interstrand hydrogen bonds in the β-sheet is cooperative, it appears that the β-sheet between Arc monomers is not formed in the transition-state dimer.

Pro8 occupies an important position in the wild-type protein–DNA complex, linking the the N-terminal arm that makes DNA backbone contacts with the β-sheet that contacts the major groove (Raumann et al., 1994b). However, none of the atoms of the Pro8 side chain are within 5 Å of the DNA, and thus the reduced operator affinity of the PL8 protein is likely to arise from indirect effects of the mutation. Modeling experiments show that if the PL8 mutant were to adopt the wild-type conformation when bound to DNA, then it would require breaking the Leu8-mediated hydrogen bonds. We presume that, upon binding of the PL8 protein to the operator, either the mutant hydrogen bonds are broken to allow formation of an otherwise wild-type set of protein–DNA interactions or the mutant hydrogen bonds are maintained but at the cost of energetically unfavorable conformational alterations in the protein–DNA complex. In this regard, it is intriguing that the loss in the stability of the PL8–operator complex is almost exactly equal to the gain in stability for two PL8 dimers, as might be expected if the additional interactions that stabilize PL8 need to be broken in the protein–DNA complex.

Although the changes in the structure of the PL8 dimer are small compared to wild type, the mutant protein shows greatly reduced activity as a repressor *in vivo* and displays significantly reduced affinity for operator DNA *in vitro*. Obviously, small changes in protein structure can have large effects on biological function. We also note that our DNA binding results *in vitro* indicate that PL8 monomers are as active in DNA binding as wild-type monomers (Figure 5B). Hence, if most Arc and PL8 molecules were monomeric in the cell, the two proteins would be expected to have similar activities. Since this is clearly not the case, it appears that

the proteins are largely dimeric *in vivo*. In strain UA2F/pTA200 grown under inducing conditions, we estimate that the total intracellular concentrations of wild-type Arc and the PL8 mutant are approximately 25 μM. These total concentrations are significantly greater than the equilibrium constants for folding/dimerization of these proteins (3.4×10^{-8} M for wild type; 5.4×10^{-10} M for PL8), and thus it would not be surprising if both proteins are mostly dimeric in the cell.

The degree of protein stabilization conferred by the PL8 mutation is substantial and in the same range as stabilization achieved by introducing disulfide bonds (Villafranca et al., 1987; Kanaya et al., 1991; Matsumura et al., 1989a,b), introducing good helix-capping residues (Nicholson et al., 1988,1991; Kaarsholm et al., 1993), and replacing buried polar groups with hydrophobic groups (Hickey et al., 1991). In many of these cases, however, the function of the stabilized protein is compromised to some degree even when the changes in protein structure are known to be small. In addition, for DNA-binding proteins such as Arc, λ repressor, and λ Cro, we have found that enhanced-stability mutants occur surprisingly frequently ($\approx$10%) in collections of functionally defective proteins with changes in or near the DNA-binding surface (Hecht et al., 1984; Pakula & Sauer, 1989; Vershon et al., 1986a; Brown et al., 1994). Taken together, these observations suggest that the evolutionary optimization of protein sequence and structure for biological function may be achieved at the expense of protein stability in many cases.

ACKNOWLEDGMENT

We thank Mark Rould, Tracy Smith, and Bronwen Brown for help and advice and Carl Pabo for the use of X-ray equipment.

REFERENCES

Becktel, W. J., & Schellman, J. A. (1987) *Biopolymers 26*, 1859–1877.

Bevington, P. R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York.

Bonvin, A. M. J. J., Vis, H., Breg, J. N., Burgering, M. J. M., Boelens, R., & Kaptein, R. (1994) *J. Mol. Biol. 236*, 328–341.

Bowie, J. U., & Sauer, R. T. (1989a) *J. Biol. Chem. 264*, 7596–7602.

Bowie, J. U., & Sauer, R. T. (1989b) *Biochemistry 28*, 7139–7143.

Breg, J. N., van Ophesden, J. H. J., Burgering, M. J., Boelens, R., & Kaptein, R. (1990) *Nature 346*, 586–589.

Brenstein, R. J. (1989) Robelko Software, version 0.9 8b5, Carbondale, IL.

Brown, B. M., & Sauer, R. T. (1993) *Biochemistry 32*, 1354–1363.

Brown, B. M., Bowie, J. U., & Sauer, R. T. (1990) *Biochemistry 29*, 11189–11195.

Brown, B. M., Milla, M. E., Smith, T. L., & Sauer, R. T. (1994) *Nature Struct. Biol. 1*, 164–168.

Brünger, A. T. *X-PLOR v3.1 Manual*, Yale University Press, New Haven, CT.

Brünger, A. T. (1992) *Nature 355*, 472–474.

Das, G., Hickey, D. R., McLendon, D., McLendon, G., & Sherman, F. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 496–499.

Fitzgerald, P. M. D. (1988) *J. Appl. Crystallogr. 21*, 273–278.

Fujinaga, M., & Read, R. J. (1987) *J. Appl. Crystallogr. 20*, 517–521.

Hecht, M. H., Sturtevant, J. M., & Sauer, R. T. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5685–5689.

Hickey, D. R., Berghuis, A. M., Lafond, G., Jaeger, J. A., Cardillo, T. S., McLendon, D., Das, G., Sherman, F., Brayer, G. D., & McLendon, G. (1991) *J. Biol. Chem. 266*, 11686–11694.

Ishikawa, K., Kimura, S., Kanaya, S., Morikawa, K., & Nakamura, H. (1993) *Protein Eng. 6*, 85–91.

Jancarik, J., & Kim, S. H. (1991) *J. Appl. Crystallogr. 24*, 409–411.

Johnson, M., & Frasier, S. (1985) *Methods Enzymol. 117*, 301–342.

Jordan, S. R., Pabo, C. O., Vershon, A. K., & Sauer, R. T. (1985) *J. Mol. Biol. 185*, 445–446.

Kaarsholm, N. C., Norris, K., Jørgensen, R. J., Mikkelsen, J., Ludvigsen, S., Olsen, O. H., Sørensen, A. R., & Havelund, S. (1993) *Biochemistry 32*, 10773–10778.

Kanaya, S., Katsuda, C., Kimura, S., Nakai, T., Kitakuni, E., Nakamura, H., Katayanagi, K., Moridawa, K., & Ikehara, M. (1991) *J. Biol. Chem. 266*, 6038–6044.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.

Matsumura, M., Becktel, W. J., Levitt, M., & Matthews, B. W. (1989a) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6562–6566.

Matsumura, M., Signor, G., & Matthews, B. W. (1989b) *Nature 342*, 291–293.

Matthews, B. W., Nicholson, H., & Becktel, W. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6663–6667.

Milla, M. E., & Sauer, R. T. (1994) *Biochemistry 33*, 1125–1133.

Milla, M. E., Brown, B. M., & Sauer, R. T. (1993) *Protein Sci. 2*, 2198–2205.

Milla, M. E., Brown, B. M., & Sauer, R. T. (1994) *Nature Struct. Biol. 1*, 518–523.

Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Nicholson, H., Becktel, W. J., & Matthews, B. W. (1988) *Nature 336*, 651–655.

Nicholson, H., Anderson, D. E., Dao-pin, S., & Matthews, B. W. (1991) *Biochemistry 30*, 9816–9828.

Pakula, A. A., & Sauer, R. T. (1989) *Proteins: Struct., Funct., Genet. 5*, 202–210.

Phillips, S. E. V. (1991) *Curr. Opin. Struct. Biol. 1*, 89–98.

Raumann, B. E., Brown, B. M., & Sauer, R. T. (1994a) *Curr. Opin. Struct. Biol. 4*, 36–43.

Raumann, B. E., Rould, M. A., Pabo, C. O., & Sauer, R. T. (1994b) *Nature 367*, 754–757.

Shirley, B. A., Stanssens, P., Hahn, U., & Pace, C. N. (1992) *Biochemistry 31*, 725–732.

Shortle, D., & Lin, B. (1985) *Genetics 110*, 539–555.

Susskind, M. M., & Youderian, P. (1983) in *Lambda II* (Hendrix, R. W., Roberts, J. W., Stahl, F. W., & Weisberg, R. A., Eds.) pp 347–364, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Vershon, A. K., Youderian, P., Susskind, M. M., & Sauer, R. T. (1985) *J. Biol. Chem. 260*, 12124–12129.

Vershon, A. K., Bowie, J. U., Karplus, T. M., & Sauer, R. T. (1986a) *Proteins: Struct., Funct., Genet. 1*, 302–311.

Vershon, A. K., Blackmer, K., & Sauer, R. T. (1986b) in *Protein Engineering: Applications in Science, Medicine, and Industry* (Inouye, M., & Sarma, R., Eds.) pp 243–256, Academic Press, Orlando, FL.

Villafranca, J. E., Howell, E. E., Oatley, S. J., Xuong, N., & Kraut, J. (1987) *Biochemistry 26*, 2182–2189.

Zagorski, M. G., Bowie, J. U., Vershon, A. K., Sauer, R. T., & Patel, D. J. (1989) *Biochemistry 28*, 7139–7143.

BI9423493