Protein Science (1993), *2*, 2198-2205. Cambridge University Press. Printed in the USA.

P22 Arc repressor: Enhanced expression of unstable mutants by addition of polar C-terminal sequences

MARCOS E. MILLA, BRONWEN M. BROWN, AND ROBERT T. SAUER
Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(RECEIVED August 25, 1993; ACCEPTED September 17, 1993)

Abstract

Many mutant variants of the P22 Arc repressor are subject to intracellular proteolysis in *Escherichia coli*, which precludes their expression at levels sufficient for purification and subsequent biochemical characterization. Here we examine the effects of several different C-terminal extension sequences on the expression and activity of a set of Arc mutants. We show that two tail sequences, KNQHE (st5) and H_6KNQHE (st11), increase the expression levels of most mutants from 10- to 20-fold and, in some cases, result in restoration of biological activity in the cell. A third tail sequence, HHHHHH (st6), was not as effective in increasing mutant expression levels. All three tail sequences are functionally and structurally silent, as judged by their lack of effects on the DNA binding activity and stability of otherwise wild-type Arc. The properties of the st11 tail sequence make it an efficient system for the expression and purification of mutant Arc proteins, both because mutant expression levels are increased and because the proteins can be rapidly purified using nickel-chelate affinity chromatography. Arc mutants containing the EA28, RL31, and SA32 mutations were purified in the st11 background. The thermodynamic stability of the EA28 mutant ($\Delta\Delta G_u \approx -0.4$ kcal/mol) is reduced modestly compared to the st11 parent, whereas the RL31 mutant ($\Delta\Delta G_u \approx -3.0$ kcal/mol) and SA32 mutant ($\Delta\Delta G_u \approx -3.3$ kcal/mol) are substantially less stable.

Keywords: affinity purification; C-terminal tails; DNA binding protein; intracellular degradation; protein stability; protein structure-function

Biochemical and biophysical studies of mutant variants that contain single amino acid substitutions have become the primary method for evaluating the contributions of individual residues to the folding, stability, and function of proteins. We are studying P22 Arc, a 53-residue dimeric phage repressor (Sauer et al., 1983; Vershon et al., 1985; Knight et al., 1989). The solution structure of the Arc dimer has been determined (Breg et al., 1990), as have the crystal structures of the DNA bound protein and the unbound protein (B. Raumann, M. Rould, C. Pabo, & R. Sauer, submitted for publ.; C. Kissenger, U. Obeysekare, B. Raumann, R. Sauer, & C. Pabo, in prep.). Unfortunately, many Arc mutants are sensitive to proteolysis in *Escherichia coli* and are difficult to purify because they do not accumulate to high steady-state levels and/or are degraded during purification (Vershon et al., 1986; Bowie & Sauer, 1989c). The expression of proteolytically unstable mutants in other systems has been increased by using strains defective in specific proteases or the heat-shock response or by altering growth conditions (Gottesman & Zipser, 1978; Kohno & Roth, 1979; Baker et al., 1984; Parsell & Sauer, 1989; Reidhaar-Olson et al., 1990), but these approaches have not helped in the Arc system (Vershon et al., 1986; our unpublished results).

Another potential method of increasing mutant expression levels involves altering the C-terminus of Arc, since addition of a 26-residue tail has been shown to protect some Arc mutants from intracellular degradation (Bowie & Sauer, 1989b; Parsell & Sauer, 1989). In this paper, we examine the effects of three different C-terminal tail sequences, ranging from 5 to 11 residues, on the intracellular expression of a set of mutant Arc proteins. The tail sequences examined were KNQHE (designated st5), HHHHHH (H_6; designated st6), and H_6KNQHE (designated st11). Tails including the KNQHE sequence were investigated because this sequence occurs at the end of a 26-residue extension that was previously shown to prevent proteolysis of some Arc mutants (Bowie & Sauer, 1989b). Tails including the H_6 sequence were investigated because of the possibility of affinity purification of the variant proteins using nickel-chelate columns (Hochuli et al., 1987). Addition of the st5 or st11 tails to *arc* genes encod-

Reprint requests to: Robert T. Sauer, Department of Biology, MIT 16-843, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.

ing proteolysis-sensitive mutants results in a significant increase in their steady-state levels of expression and, in some cases, leads to restoration of Arc activity in vivo. The st6 tail was found to be less effective in increasing expression levels of unstable Arc mutants. None of the three C-terminal extensions tested affects the DNA binding activity of Arc. In addition, otherwise wild-type variants bearing these tails have the same thermodynamic stability as wild-type Arc. Of the three tail sequences tested, the st11 tail is potentially the most useful because it both increases mutant expression levels and allows affinity purification.

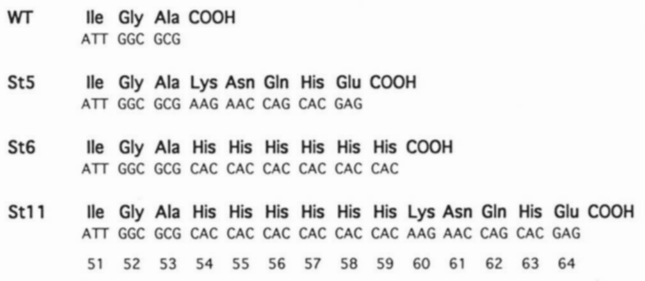

Fig. 1. Sequences of the C-terminal residues of wild-type (WT) Arc and the Arc-st5, Arc-st6, and Arc-st11 variants.

Results

Arc mutants containing the PL15, PS15, RC23, EA28, NK29, NT29, RL31, RW31, and SF32 amino acid substitutions were chosen for expression studies (Vershon et al., 1986). The steady-state levels of these mutants range from 50% to less than 10% of the wild-type level under inducing conditions in strain UA2F.

Effects of C-terminal tails on intracellular expression

The nine Arc mutants described above were recloned into genes encoding the three different C-terminal extensions shown in Figure 1. We refer to these tail sequences as st5, st6, and st11. Figure 2 shows the relative expression levels for each of the nine mutants and wild-type Arc in the different tail backgrounds, as determined by densitometry of Coomassie-stained sodium dodecyl sulfate (SDS) gels. The st6 tail has only a modest effect on expression of the mutant proteins. In contrast, in most cases, expression of the mutant proteins in the st5 or st11 tail backgrounds increases 10–20-fold relative to the levels observed for the mutants without tails. The st5 and st11 tails also increase the steady-state levels of otherwise wild-type Arc from 1.5- to 3-fold.

Increased expression can suppress some but not all mutant defects

For inactive Arc mutants that are expressed poorly, it is important to determine whether the defective phenotypes arise solely from their reduced intracellular levels or also occur because of defects in stability or DNA binding. To address this issue, we assayed the activities of the nine mutants described above and an additional seven mutants from the collection of Vershon et al. (1986) in the st5 tail background. As described in Materials and methods, an Arc^+ phenotype in strain UA2F results in chloramphenicol sensitivity, whereas an Arc^- phenotype results in a chloramphenicol resistance. For example, we know that the Arc-st5, Arc-st6, and Arc-st11 variants are all active in vivo because, as shown in Table 1, UA2F cells containing these proteins fail to grow in the presence of chloramphenicol.

Table 1 also shows the chloramphenicol phenotypes of the mutants with and without the st5 tail. Each of the 16 mutants is inactive without the tail whether or not expression is induced with isopropyl-β-D-thiogalactoside (IPTG). By contrast, the EA17-st5, RH23-st5, EA28-st5, and SA32-st5 mutants are active under both inducing and noninducing conditions, and the PS15-st5, RC23-st5, EK28-st5, and NK29-st5 mutants are active under inducing conditions. Thus, the st5 tail rescues biological activity for eight of the previously inactive mutants. The st5 tail does not rescue activity for the remaining eight mutants (PL15-st5, NT29-st5, NY29-st5, RL31-st5, RW31-st5, SF32-st5, EA36-st5, and RQ40-st5), which are inactive even under inducing conditions.

The st5, st6, and st11 tails do not affect the activity or stability of otherwise wild-type Arc

Variants containing the st5, st6, and st11 tails were purified for biochemical studies of activity and stability in

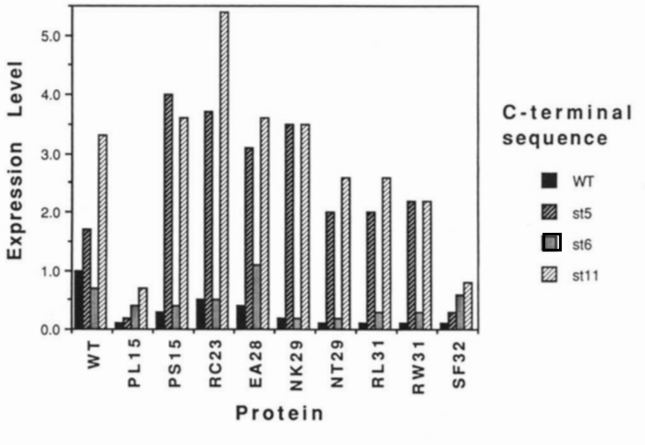

Fig. 2. Steady-state expression levels of wild-type and mutant Arc proteins with different C-terminal sequences in strain UA2F. Levels were measured by densitometry of SDS gels (see Materials and methods) following IPTG induction for 2.5 h at 37 °C. Expression levels are reported in arbitrary units relative to a value of 1.0 for wild-type Arc.

Table 1. *Effects of Arc variants on growth in the presence of chloramphenicol*[a]

Without tail			With tail		
Protein	No IPTG	Plus IPTG	Protein	No IPTG	Plus IPTG
Arc	–	–	Arc-st5	–	–
			Arc-st6	–	–
			Arc-st11	–	–
EA17	+	+	EA17-st5	–	–
RH23	+	+	RH23-st5	–	–
EA28	+	+	EA28-st5	–	–
SA32	+	+	SA32-st5	–	–
PS15	+	+	PS15-st5	+	–
RC23	+	+	RC23-st5	+	–
EK28	+	+	EK28-st5	+	–
NK29	+	+	NK29-st5	+	–
PL15	+	+	PL15-st5	+	+
NT29	+	+	NT29-st5	+	+
NY29	+	+	NY29-st5	+	+
RL31	+	+	RL31-st5	+	+
RW31	+	+	RW31-st5	+	+
SF32	+	+	SF32-st5	+	+
EA36	+	+	EA36-st5	+	+
RQ40	+	+	RQ40-st5	+	+

[a] + Indicates cell growth; – indicates no growth.

vitro. For Arc-st6 and Arc-st11, the adjacent histidine sequences in the C-terminal tails permit purification on nickel-chelate columns (Hochuli et al., 1987). Figure 3 shows fractions from the purification of Arc-st11; the protein is substantially free of contaminants after a single affinity chromatography step (lane 5). The circular dichroism and fluorescence spectra of wild-type Arc, Arc-st5, Arc-st6, and Arc-st11 were virtually identical (data not shown), suggesting that the tail sequences do not cause any substantial changes in protein structure.

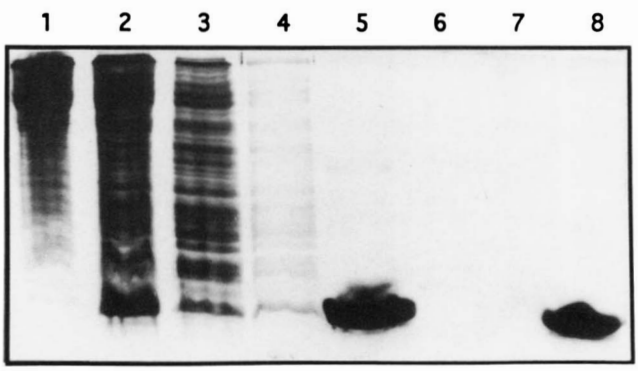

Fig. 3. SDS-polyacrylamide gel analysis of fractions from the purification of Arc-st11. Lanes 1, 2, pellet and supernatant fractions following GuHCl lysis of cells; lanes 3–5, flow through, pH 8.0 wash, and 0.2 M acetic acid elution of Ni^{2+}-NTA agarose column; lanes 6–8, flow through, 0.1 M NaCl wash, and 2 M NaCl elution of SP-Sephadex column.

In gel mobility shift experiments, Arc, Arc-st5, Arc-st6, and Arc-st11 all bind operator DNA with approximately the same affinity, as shown by the concentration of free protein (ca. 0.2–0.3 nM) required to bind half of the operator DNA (Table 2). This result indicates that the tails do not cause significant changes in the ability of the Arc variants to bind operator DNA. Figure 4 shows a DNA binding isotherm for Arc-st5. Like wild-type Arc, the Arc-st5 binding curve shows an approximate fourth-order dependence upon protein concentration, as expected for a reaction in which the protein binds the operator as a tetramer but monomers are the predominant solution species at the subnanomolar protein concentrations at which the binding reaction occurs (Brown et al., 1990).

The thermodynamic stabilities of Arc, Arc-st5, Arc-st6, and Arc-st11 were measured in equilibrium denaturation experiments. Thermal melts of the proteins at concentrations of 5 μM are shown in Figure 5A. The melting curves for each of the proteins are similar. Wild-type Arc has a t_m of 55.6 (±0.3, SD) °C and a van't Hoff enthalpy of unfolding of 55 (±2) kcal/mol under these conditions. As shown in Table 2, the t_m's and ΔH_u's of the st5, st6, and st11 variants are only marginally different from those of wild type.

Figure 5B shows urea denaturation experiments for Arc, Arc-st5, Arc-st6, and Arc-st11 at protein concentrations of 5 μM. The four proteins all display extremely similar denaturation profiles. Table 3 lists values for ΔG_u, m, and C_m (the concentration of urea required for 50% denaturation) for these experiments as well as additional experiments performed using Arc-st6 and Arc-st11. These

Table 2. *DNA binding activity and thermal stability*[a]

	[Protein] at 50% DNA binding (nM)	t_m (°C)	ΔH_u (kcal/mol)	$\Delta\Delta G_u$ (kcal/mol)
Arc	0.21	55.6 (±0.3)	54.9 (±2.0)	0.00
Arc-st5	0.20	54.7 (±0.3)	53.7 (±2.2)	−0.15
Arc-st6	0.16	57.1 (±0.3)	57.0 (±2.2)	0.25
Arc-st11	0.30	56.9 (±0.3)	55.4 (±2.5)	0.15
EA28-st11	ND	53.2 (±0.5)	53.7 (±2.0)	−0.46
RL31-st11	ND	35.0 (±0.5)	29.9 (±2.5)	−3.08
SA32-st11	ND	33.8 (±0.2)	26.0 (±2.5)	−3.31

[a] Operator DNA binding was measured at 20 °C in a buffer containing 10 mM Tris (pH 7.5), 100 mM KCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 100 μg/mL bovine serum albumin, and 0.02% NP40. Thermal stability was measured at a protein concentration of 5 μM in a buffer containing 50 mM Tris (pH 7.5), 250 mM KCl, and 0.1 mM EDTA. ΔH_u values are calculated at the t_m for the protein. $\Delta\Delta G_u$ values for Arc-st5, Arc-st6, and Arc-st11 are calculated relative to wild-type Arc at its t_m (55.6 °C). $\Delta\Delta G_u$ values for EA28-st11, RL31-st11, and SA32-st11 are calculated relative to Arc-st11 at 55.6 °C. The error values shown in parentheses are 67% confidence limits. When the upper and lower limits were asymmetric, the larger of the two errors is cited. ND, not determined.

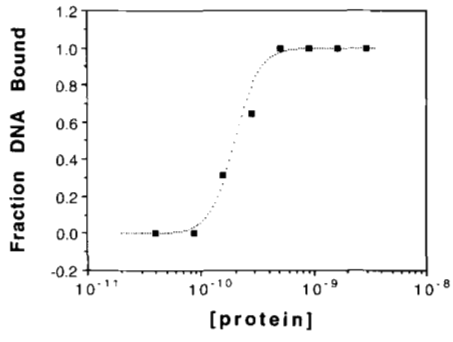

Fig. 4. Binding of Arc-st5 to the O1 *arc* operator DNA fragment. The dotted line is a theoretical curve for a fourth-order binding reaction.

additional experiments show that similar values of ΔG_u and m are measured in denaturation experiments performed in duplicate, in experiments performed at different protein concentrations, and in experiments monitored by CD or fluorescence. In the case of Arc-st11, the average values (±SD) of ΔG_u and m calculated from seven independent experiments are 10.57 (±0.21) kcal/mol and 1.39 (±0.06) kcal/mol-M. Unlike the values of ΔG_u or m, which should not vary with protein concentration, the value of C_m does increase with protein concentration as expected for a bimolecular folding reaction (Bowie & Sauer, 1989a). For example, Arc-st11 at a concentration of 2 μM is 50% denatured at roughly 2 M urea, whereas the same protein at a concentration of 20 μM is 50% denatured at roughly 3 M urea (Table 3).

The kinetics of refolding of the Arc-st5 protein were examined by stop flow experiments monitored by fluorescence. The bimolecular refolding rate constant (k_f) was found to be 6×10^6 M^{-1} s^{-1} at 20 °C, 0.1 M KCl, and pH 7.5 (data not shown). Under comparable conditions, k_f for wild-type Arc is 9×10^6 M^{-1} s^{-1} (Milla & Sauer, 1993). Thus, the st5 tail has very little effect on the refolding rate.

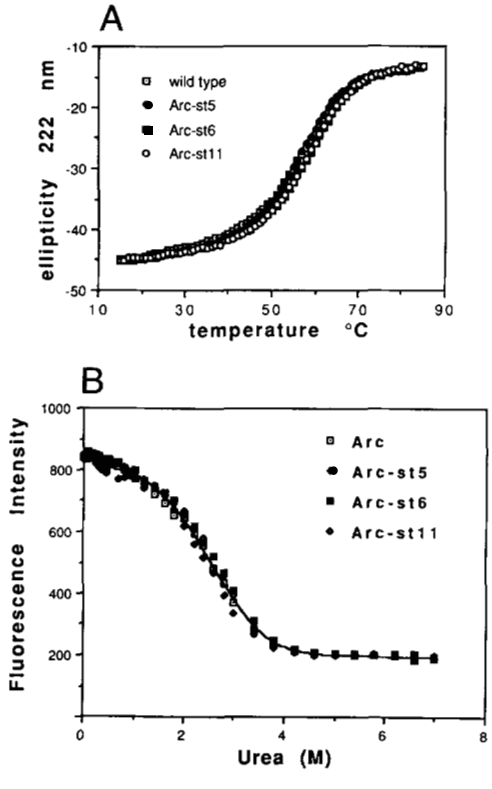

Fig. 5. A: Thermal melts of wild-type Arc and the st5, st6, and st11 variants. **B:** Urea denaturation of wild-type Arc and the st5, st6, and st11 variants. The solid curve was calculated using $\Delta G_u = 10.56$ kcal/mol and $m = 1.39$ kcal/mol-M. In both cases, the protein concentration is 5 μM, and the buffer contains 50 mM Tris-HCl (pH 7.5), 250 mM KCl, and 0.2 M EDTA. The urea denaturation experiments were performed at 25 °C.

Purification and stabilities of mutant proteins

To test the usefulness of the st11 tail for expression and purification of Arc mutants, we cloned the EA28, RL31, and SA32 substitutions into genes containing the st11 extension and affinity purified the mutant proteins. Thermal melts (Table 2) and urea denaturation (Fig. 6; Table 3) of the three mutant proteins were performed. The EA28-st11 mutant is slightly less thermally stable than Arc-st11 ($\Delta t_m = -3.7$ °C). In contrast, the RL31-st11 ($\Delta t_m = -21.9$ °C) and SA32-st11 ($\Delta t_m = -23.1$ °C) mutants are significantly less thermally stable. Similar changes in stability are seen in urea denaturation experiments. The RL31-st11 and SA32-st11 proteins are roughly 2.5–3.0 kcal/mol less stable than Arc-st11 to urea denaturation, whereas EA28-st11 is only about 0.3 kcal/mol less stable (Table 3).

Discussion

Two factors, thermodynamic stability and C-terminal sequence, appear to determine the susceptibility of Arc and its variants to intracellular degradation (Vershon et al., 1986; Bowie & Sauer, 1989b; Parsell & Sauer, 1989; Parsell et al., 1990). Protein stability is thought to be impor-

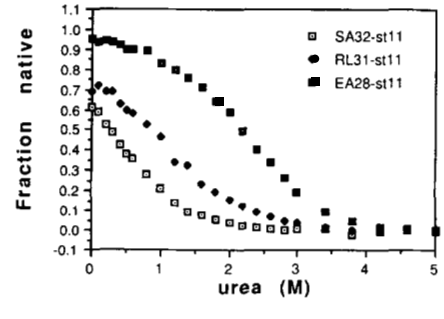

Fig. 6. Urea denaturation curves for the EA28-st11, RL31-st11, and SA32-st11 mutants. Conditions: 5 μM protein, 25 °C, 50 mM Tris-HCl (pH 7.5), 250 mM KCl, and 0.2 M EDTA.

Table 3. *Urea denaturation parameters*[a]

	P_t (μM)	ΔG_u (kcal/mol)	m (kcal/mol-M)	C_m (M)
Arc	5	10.41	1.34	2.39
Arc-st5	5	10.82	1.53	2.36
Arc-st6	5	10.65	1.35	2.55
Arc-st6	8.7	10.64	1.37	2.75
Arc-st6	8.7	10.26	1.29	2.62
Arc-st6†	8.7	10.50	1.32	2.75
Arc-st6†	8.7	10.79	1.40	2.80
Arc-st6 mean		10.57 ± 0.20	1.35 ± 0.04	
Arc-st11	2	10.58	1.41	2.01
Arc-st11	5	10.94	1.47	2.54
Arc-st11	10	10.22	1.28	2.68
Arc-st11	20	10.61	1.40	3.02
Arc-st11	20	10.49	1.39	2.95
Arc-st11†	20	10.59	1.37	3.07
Arc-st11†	20	10.50	1.38	2.98
Arc-st11 mean		10.56 ± 0.21	1.39 ± 0.06	
EA28-st11	5	10.26	1.39	2.18
RL31-st11	5	8.15	1.13	0.84
SA32-st11	5	7.58	1.46	0.26

[a] P_t, Protein concentration in monomer equivalents; ΔG_u, free energy change associated with unfolding of Arc dimer at standard state concentration of 1 M, 25 °C, pH 7.5, 250 mM KCl calculated from urea denaturation experiments as described in Materials and methods. Denaturation experiments marked by † were monitored by circular dichroism; all others were monitored by fluorescence.

tant because many mutations that reduce thermodynamic stability also lead to increased degradation in the cell. The C-terminal sequence is thought to be important because extensions with charged and polar residues at or near the C-terminus have been shown to increase the intracellular half-life of wild-type Arc and some variants. These extensions or tails are thought to act by preventing the binding of a protease with a preference for relatively nonpolar sequences like GRIGA, which is the C-terminal sequence of wild-type Arc (Parsell et al., 1990). The intracellular protease or proteases that are responsible for degrading Arc mutants in the cell are currently unknown.[1]

We have shown that the st5 and st11 C-terminal tails, both of which end with the sequence KNQHE, increase the intracellular expression of most of the mutant Arc variants tested. The KNQHE sequence was originally found as the terminal pentapeptide of the *lt*1 tail, a 26-residue extension that increased the intracellular half-life of three different Arc mutants (Bowie & Sauer, 1989b). The KNQHE tail appears to be as effective as the *lt*1 tail in increasing mutant expression levels, providing support for the studies of Parsell et al. (1990) that suggested that the stabilizing effect of the *lt*1 tail was due to the polar nature of its terminal five residues. These studies also showed that a variety of pentapeptide sequences consisting of charged and polar residues were effective in blocking intracellular degradation of another protein, the N-terminal domain of phage λ repressor. For this reason, we also assayed the effects of the st6 tail, which adds six histidines to the C-terminus. Surprisingly, the st6 tail was found to be far less effective than either st5 or st11 in increasing the expression levels of the Arc mutants. This result indicates that we do not yet fully understand the rules that relate C-terminal sequences and susceptibility to intracellular degradation.

The st5 and st11 tails are not universally effective at increasing mutant expression, and a few mutants, such as PL15 and SF32, are still expressed at relatively low levels even with these tail sequences. It is possible that these mutations affect Arc expression via changes in message stability or translational efficiency or introduce new proteolytic target sites and make Arc subject to degradation by tail-independent proteases. Alternatively, the st5 and st11 tails may not provide complete protection from tail-dependent proteases for mutants whose native structures are severely destabilized. It is intriguing that both Pro 15 and Ser 32 act as N-caps for α-helices in Arc (Breg et al., 1990; Kissenger, Obeysekare, Raumann, Sauer, & Pabo, in prep.), since mutations that affect residues at the beginning of α-helices in λ repressor and λ Cro have also been found to cause large changes in the proteolytic susceptibility of these proteins (Pakula & Sauer, 1989; Reidhaar-Olson et al., 1990).

Addition of the st5, st6, and st11 tails to otherwise wild-type Arc did not alter the operator DNA binding activity or structural stability of the protein. These findings are consistent with the idea that the tail sequences are unstructured and do not interact with the folded portions of the fusion proteins. Several C-terminal residues in the wild-type Arc protein are known to be disordered in both the solution structure (Breg et al., 1990) and the crystal structure (Kissenger, Obeysekare, Raumann, Sauer, & Pabo, in prep.), and NMR studies of the *lt*1 extension have shown that the 26 residues of this tail sequence are also unstructured (Bowie & Sauer, 1989b). In the case of the st5 tail, we also found that this C-terminal extension did not significantly affect the kinetics of Arc folding, suggesting that the C-terminal tails do not suppress degradation by increasing the folding rate.

Three Arc mutants—EA28-st11, RL31-st11, and SA32-st11—were purified for biochemical studies. Each of these substitutions results in loss of biological activity in the wild-type tail background (Vershon et al., 1986), but the EA28 and SA32 mutants are active in the st11 tail background. This result shows that proteolysis makes a substantial contribution to the mutant phenotypes observed for these two mutants in the wild-type tail background.

[1] Tsp, a periplasmic *E. coli* protease, has the proper substrate specificity and degrades the appropriate Arc variants in vitro but is not responsible for the observed degradation of the Arc mutants in vivo (Silber et al., 1992; K. Silber, K. Keiler, & M. Milla, unpubl.).

Arg 31 is a partially buried residue in the turn between the A and B helices of Arc. In the Arc structure (Kissenger, Obeysekare, Raumann, Sauer, & Pabo, in prep.), the side chain nitrogens of Arg 31 make hydrogen bond/salt bridge interactions with the side chain oxygens of the buried Glu 36 side chain and also make a hydrogen bond with the side chain oxygen of Asn 29. Ser 32 is the first residue of helix B, and its side chain hydroxyl hydrogen bonds to its own amide nitrogen and to the amide nitrogen and side chain hydroxyl of Ser 35. Because the SA32 and RL31 mutations cause large decreases in both the thermal stability and the urea stability of the protein, it appears that some or all of these hydrogen bonding and/or salt bridge interactions contribute significantly to the stability of Arc. The destabilization caused by the RL31 and SA32 substitutions results in more than 50% of the mutant proteins being unfolded at 37 °C under the conditions of our denaturation assays. This large degree of unfolding could easily account for the observed susceptibility of the RL31 and SA32 mutant proteins to degradation in the cell.

Glu 28 is a surface residue in helix A (Breg et al., 1990), and the EA28 mutation causes only a small decrease in thermodynamic stability. At 37 °C under the conditions of our assays in vitro, the EA28 mutation would only increase the amount of unfolded protein modestly compared with wild-type Arc, and yet in vivo the EA28 mutant appears to be degraded two to four times as rapidly as wild type, as evidenced by its reduced steady-state levels (Vershon et al., 1986; Fig. 2). Apparently, even small reductions in thermodynamic stability can have measurable effects on proteolytic sensitivity of Arc in the cell.

The biochemical studies presented here show that the st5, st6, and st11 tail sequences are structurally and functionally silent. Any effects of these tails on function in vivo must therefore be due to changes in steady-state expression levels. Both the st5 and st11 tails increase the expression of many Arc mutants to the point where their purification is readily feasible. We believe that an expression system using the st11 tail (H_6KHQNE) will prove to be the most useful in studying unstable Arc mutants both because this tail increases expression levels and because it allows the mutant proteins to be rapidly purified using affinity chromatography. We anticipate that approaches similar to those described here may also be useful for expressing mutant proteins in other systems for which C-terminal-dependent degradation is a problem.

Materials and methods

Strains and plasmids

E. coli strain UA2F is *thi*$^-$ *his*$^-$ *lacZ*$^-$ *lacY*$^+$ *sup*^{0} *recA*$^-$ F′(*lacI*^{Q} *lacZ*::Tn5 [*kan*^{R}] *pro*$^+$) and contains an integrated λ*imm*21 prophage bearing a transcriptional fusion of the Arc-repressible P_{ant} promoter to the *cat* gene (Vershon et al., 1986).

Arc genes encoding C-terminal extensions were constructed by inserting synthetic oligonucleotide cassettes between the *Hin*d III and *Cla* I sites in the *arc* gene of plasmid pSA300 (Bowie & Sauer, 1989c). In each case, the extensions were added immediately following the codon for the last wild-type residue, Ala 53, as shown in Figure 1. Following ligation of the appropriate cassettes to the gel-purified *Hin*d III–*Cla* I backbone of pSA300, constructs were transformed into strain UA2F. Plasmid DNAs from individual transformants were purified, and the expected plasmid structures were verified by restriction mapping and DNA sequencing. Plasmids pSA500, pSA600, and pSA700 encode Arc variants containing the st5, st6, and st11 tails, respectively.

Proteolysis-sensitive Arc mutations were recloned from the collection of Vershon et al. (1986) into different tail backgrounds. The original mutations were isolated in plasmid pTA200. To transfer these mutations, variants of pTA200 were cut with *Pvu* I and *Hin*d III, and the 1.1-kb fragment (containing DNA from the *amp*^{R} gene upstream of *arc* to codon 44 of *arc*) was purified and ligated to the 3.3-kb gel-purified *Pvu* I–*Hin*d III backbone of either pSA300, pSA500, pSA600, or pSA700. Constructs were transformed into strain UA2F by electroporation, and plasmid identities were verified by DNA sequencing.

Expression assays

To assay expression of mutant variants of Arc, fresh colonies bearing variants of plasmid pTA200, pSA500, pSA600, or pSA700 were placed in a culture tube in 2 mL of Luria-Bertani (LB) broth containing 150 μg/mL ampicillin and grown overnight in a roller drum at 37 °C. The next day, 30 μL of each culture was used to inoculate 2 mL of fresh medium. Cells were grown at 37 °C to an absorbance of 0.8 at 600 nm, and IPTG was then added to a final concentration of 0.1 mg/mL. A duplicate set of cultures did not receive IPTG. After 2.5 h, all of the cultures were diluted with LB to an absorbance of 0.7 at 600 nm. Next, 1.5 mL of each culture was transferred to a microfuge tube, and cells were pelleted and resuspended in 0.2 mL of Laemmli loading buffer (Laemmli, 1970). Samples were boiled for 2 min, insoluble material was removed with a toothpick, and samples were electrophoresed on SDS-polyacrylamide gels using the Tris-tricine buffer system (Schägger & von Jagow, 1987). Gels were stained with 0.02% Coomassie brilliant blue R-250 in 50% methanol and 10% acetic acid and dried between cellophane sheets using a Hoeffer Easy Breeze dryer. The intensities of bands of Arc protein and gel background were quantified using a Molecular Dynamics Image Quant computing densitometer.

Protein purification

Wild-type Arc and the variant containing the KNQHE tail (designated Arc-st5) were purified from strains UA2F/pSA300 and UA2F/pSA500, respectively, by chromatography on Sephadex G50 and CM-Accell as previously described (Bowie & Sauer, 1989b).

Variants containing the H_6 tail (designated Arc-st6) and H_6KHQNE tail (designated Arc-st11) were purified by affinity chromatography following the basic procedure of Hochuli et al. (1987). Strains UA2F/pSA600 or UA2F/pSA700 were grown at 37 °C with aeration in 6 L of LB broth to an absorbance of 1.0 at 600 nm, and Arc expression from the P_{tac} promoter was induced by addition of IPTG to a concentration of 0.1 mg/mL. After 3 h, cells were harvested, washed, and resuspended in 100 mL of buffer A (100 mM potassium phosphate, 10 mM Tris-HCl, and 6 M guanidine hydrochloride, adjusted to a final pH of 8 at room temperature). Cells were lysed by gentle stirring for 1 h at room temperature, and cellular debris was removed by centrifugation at 12,000 × *g* (Sorvall SS-34 rotor) for 30 min. The supernatant was loaded onto a 3-mL Qiagen Ni^{2+}-NTA metal-chelating chromatography resin equilibrated in buffer A. The column was washed with 100 mL of buffer A plus 3 mM imidazole, and the Arc-st6 or Arc-st11 variants were eluted with 5 mL of 0.2 M acetic acid containing 6 M guanidine hydrochloride. Following dialysis into buffer B (10 mM Tris-HCl [pH 7.5], 0.2 mM EDTA), the protein was loaded onto a 2-mL SP-Sephadex column equilibrated with buffer B, washed with 100 mL of buffer B plus 100 mM KCl, and eluted with 5 mL of buffer B containing 2 M KCl. The eluate was dialyzed twice against 6 L of storage buffer, which contains 50 mM Tris-HCl (pH 7.5), 250 mM KCl, and 0.2 mM EDTA.

Following purification, Arc, Arc-st5, Arc-st6, and Arc-st11 were found to be greater than 95% pure as judged by Coomassie blue staining of samples following SDS gel electrophoresis (Laemmli, 1970). Typical yields of purified protein from 6-L preparations were 20–40 mg. The concentration of wild-type Arc was calculated using an extinction coefficient of 6,756 M^{-1} cm^{-1} at 280 nm (Brown et al., 1990). Because the tail sequences do not contain any additional tyrosine or tryptophan residues, the same extinction coefficient was used to calculate molar concentrations for Arc-st5, Arc-st6, and Arc-st11.

The variants EA28-st11, RL31-st11, and SA32-st11 were purified from 2-L cultures using chromatography on Ni^{2+}-NTA and SP-Sephadex columns as described above. In each case, the final recovery of the mutant protein was 5–6 mg.

Activity assays

The activity in vivo of Arc variants was assayed by their ability to repress transcription of the gene for chloramphenicol acetyl transferase gene (*cat*) initiated from an Arc-repressible promoter located in the bacterial chromosome of strain UA2F (Vershon et al., 1986). Active Arc variants repress *cat* synthesis, and the host cell cannot grow in the presence of chloramphenicol. For the experiments presented here, cells were plated at 37 °C on LB plates containing 75 μg/mL chloramphenicol, 2 μg/mL IPTG, 100 μg/mL ampicillin, and 25 μg/mL kanamycin.

The binding of purified Arc or its variants to operator DNA was assayed in vitro by gel mobility shift assays as described (Brown et al., 1990; Brown & Sauer, 1993). The operator DNA used for these assays was the 27-bp O1 synthetic fragment (Brown & Sauer, 1993), which had been 5′-end-labeled with adenosine 5′-[γ^{32}P]triphosphate using polynucleotide kinase.

Denaturation assays

The stability against denaturation of wild-type Arc and variants was assayed by monitoring changes in tryptophan fluorescence or CD as a function of urea concentration or temperature (Bowie & Sauer, 1989c). For urea denaturation, Arc was diluted to a final concentration of 2–20 μM and equilibrated at 25 °C for 1 h in buffer containing 50 mM Tris-HCl (pH 7.5), 250 mM KCl, 0.2 mM EDTA, and different concentrations of urea. Denaturation was monitored either by fluorescence emission intensity at 320 nm (excitation 280 nm) using a Perkin-Elmer LS-50 luminescence spectrometer or by CD ellipticity (228–232 nm) using an Aviv 60DS spectropolarimeter. In both cases, readings were taken in a thermostatted cell at 25 °C For thermal melts, Arc was diluted to 5 μM in a buffer containing 50 mM Tris-HCl (pH 7.5), 250 mM KCl, and 0.2 mM EDTA, and the temperature was increased in 1 °C steps. The sample was allowed to equilibrate for 60 s at each temperature point, and the ellipticity at 222 nm was recorded for 30 s. Both the thermal and urea denaturation reactions were reversible.

Denaturation curves were fitted by a nonlinear least-squares procedure using a Macintosh version of the program NonLin (Johnson & Frasier, 1985; Brenstein, 1989). Because the Arc dimer denatures to two unfolded monomers, equations for a bimolecular reaction were used for all fitting procedures (Bowie & Sauer, 1989c). For thermal denaturation, curves were fitted using the following equations:

$$\epsilon_{obs} = f_u * (\epsilon_D + ds * T) + (1 - f_u) * (\epsilon_N + ns * T)$$

$$f_u = [-K_u + (K_u^2 + 8 * P_t * K_u)^{1/2}]/(4 * P_t)$$

$$K_u = \exp(-\Delta G_u/R * T)$$

$$\Delta G_u = \Delta H_u - (T/T_m) * [\Delta H_u + R * T_m * \ln(P_t)] + \Delta C_p * [T - T_m - T * \ln(T/T_m)],$$

where ϵ_{obs} is the observed ellipticity; ϵ_D and ϵ_N are the intercept ellipticities of the denatured and native proteins; *ds* and *ns* are the slopes of the denatured and native baselines; f_u is the fraction of protein denatured; K_u is the equilibrium constant for Arc unfolding and dissociation; ΔG_u, ΔH_u, and ΔC_p are the free energy, enthalpy, and heat capacity changes associated with unfolding and dissociation; T is the temperature (K); T_m is the temperature where $f_u = 0.5$; R is the universal gas constant; and P_t is the protein concentration in monomer equivalents. For each melting curve, ΔH_u, T_m, and the slopes and intercepts of the native and denatured baselines were determined. ΔC_p is not well determined in these experiments and, for fitting purposes, was fixed at 1.6 kcal/mol, the value for wild-type Arc (Bowie & Sauer, 1989c).

Control experiments showed that the fluorescence intensities of the native (I_N) and denatured (I_D) proteins did not change with urea concentration, as has been observed for other proteins (Horovitz et al., 1990). Urea denaturation curves monitored by fluorescence were fit using the following equations:

$$I_{obs} = f_u * I_D + (1 - f_u) * I_N,$$

where f_u is defined as described above, and

$$K_u = \exp\{(-\Delta G_u + m * [urea])/R * T\}.$$

Here, m is the negative slope of the plot of the free energy of unfolding versus urea, and ΔG_u is the free energy of unfolding in the absence of urea. Urea denaturation curves monitored by CD were fit in a similar fashion, except that the slopes of the native and denatured baselines were also allowed to vary.

Acknowledgments

We thank Karen R. Silber for helpful discussions and Olivia Siu for help with expression studies. We also thank Chuck Kissenger, Upel Obeysekare, and Carl Pabo for the coordinates of the Arc dimer; Paul Schimmel for the use of the fluorescence luminescence spectrometer; and Harvey Lodish for the use of the computerized laser densitometer. M.E.M. is a postdoctoral fellow of the Jane Coffin Childs Fund for Medical Research. This work was supported in part by NIH grant AI-15706.

References

Baker, T., Grossman, A., & Gross, C. (1984). A gene regulating the heat shock response in *Escherichia coli* also affects proteolysis. *Proc. Natl. Acad. Sci. USA 81*, 6779–6783.

Bowie, J.U. & Sauer, R.T. (1989a). Equilibrium dissociation and unfolding of the Arc repressor dimer. *Biochemistry 28*, 7139–7143.

Bowie, J.U. & Sauer, R.T. (1989b). Identification of C-terminal extensions that protect proteins from intracellular proteolysis. *J. Biol. Chem. 264*, 7596–7602.

Bowie, J.U. & Sauer, R.T. (1989c). Identifying determinants of folding and activity for a protein of unknown structure. *Proc. Natl. Acad. Sci. USA 86*, 2152–2156.

Breg, J.N., Opheusden, J.H.J., Burgering, M.J.M., Boelens, R., & Kaptein, R. (1990). Structure of Arc repressor in solution: Evidence for a family of β-sheet DNA-binding proteins. *Nature 346*, 586–589.

Brenstein, R.J. (1989). *NonLin for Macintosh.* Robelko Software, Carbondale, Illinois.

Brown, B.M., Bowie, J.U., & Sauer, R.T. (1990). Arc repressor is tetrameric when bound to operator DNA. *Biochemistry 29*, 11189–11195.

Brown, B.M. & Sauer, R.T. (1993). Assembly of the Arc tetramer-operator complex: Cooperative interactions between DNA-bound dimers. *Biochemistry 32*, 1354–1363.

Gottesman, S. & Zipser, D. (1978). Deg phenotype of *Escherichia coli lon* mutants. *J. Bacteriol. 133*, 844–851.

Hochuli, E., Dobeli, H., & Schachter, A. (1987). New metal chelate adsorbent selective for proteins and peptides containing neighboring histidine residues. *J. Chromatogr. 411*, 177–184.

Horovitz, A., Serrano, L., Avron, B., Bycroft, M., & Fersht, A.R. (1990). Strength and co-operativity of contributions of surface salt bridges to protein stability. *J. Mol. Biol. 216*, 1031–1044.

Johnson, M. & Frasier, S. (1985). Nonlinear least-squares analysis. *Methods Enzymol. 117*, 301–342.

Knight, K.L., Bowie, J.U., Vershon, A.K., Kelley, R.D., & Sauer, R.T. (1989). The Arc and Mnt repressors. A new class of sequence-specific DNA-binding protein. *J. Biol. Chem. 264*, 3639–3642.

Kohno, T. & Roth, J. (1979). Electrolyte effects on the activity of mutant enzymes *in vivo* and *in vitro*. *Biochemistry 18*, 1386–1392.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature 227*, 680–685.

Milla, M.E. & Sauer, R.T. (1993). P22 Arc repressor: Folding kinetics of a single domain, dimeric protein. *Biochemistry*, in press.

Pakula, A.A. & Sauer, R.T. (1989). Amino acid substitutions that increase the thermal stability of the lambda Cro protein. *Proteins Struct. Funct. Genet. 5*, 202–210.

Parsell, D.A. & Sauer, R.T. (1989). The structural stability of a protein is an important determinant of its proteolytic susceptibility in *Escherichia coli*. *J. Biol. Chem. 264*, 7590–7595.

Parsell, D.A., Silber, K.R., & Sauer, R.T. (1990). Carboxy-terminal determinants of intracellular protein degradation. *Genes Dev. 4*, 277–286.

Reidhaar-Olson, J.F., Parsell, D.A., & Sauer, R.T. (1990). An essential proline in lambda repressor is required for resistance to intracellular proteolysis. *Biochemistry 29*, 7563–7571.

Sauer, R.T., Krovatin, W., DeAnda, J., Youderian, P., & Susskind, M.M. (1983). Primary structure of the imml immunity region of bacteriophage P22. *J. Mol. Biol. 168*, 699–713.

Schägger, H. & von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kd. *Anal. Biochem. 166*, 368–379.

Silber, K.R., Keiler, K.C., & Sauer, R.T. (1992). Tsp: A tail-specific protease that selectively degrades proteins with nonpolar C termini. *Proc. Natl. Acad. Sci. USA 89*, 295–299.

Vershon, A.K., Bowie, J.U., Karplus, T.M., & Sauer, R.T. (1986). Isolation and analysis of Arc repressor mutants: Evidence for an unusual mechanism of DNA binding. *Proteins Struct. Funct. Genet. 1*, 302–311.

Vershon, A.K., Youderian, P., Susskind, M.M., & Sauer, R.T. (1985). The bacteriophage P22 Arc and Mnt repressors. Overproduction, purification, and properties. *J. Biol. Chem. 260*, 12124–12129.