JOURNAL OF BACTERIOLOGY, Dec. 1996, p. 7025–7030 Vol. 178, No. 24
0021-9193/96/$04.00+0
Copyright © 1996, American Society for Microbiology

The Linker Region of AraC Protein

REBECCA J. EUSTANCE† AND ROBERT F. SCHLEIF*
Biology Department, Johns Hopkins University, Baltimore, Maryland

Received 22 July 1996/Accepted 2 October 1996

AraC protein, a transcriptional regulator of the L-arabinose operon in *Escherichia coli*, is dimeric. Each monomer consists of a domain for DNA binding plus transcription activation and a domain for dimerization plus arabinose binding. These are connected to one another by a linker region of at least 5 amino acids. Here we have addressed the question of whether any of the amino acids in the linker region play active, specific, and crucial structural roles or whether these amino acids merely serve as passive spacers between the functional domains. We found that all but one of the linker amino acids can be changed to other amino acids individually and in small groups without substantially affecting the ability of AraC protein to activate transcription when arabinose is present. When, however, the entire linker region is replaced with linker sequences from other proteins, the functioning of AraC is impaired.

Most larger proteins appear to fold into and to function with domains of 100 to 150 amino acids. While in some cases, the roles of the domains in a protein's function are independent, in others, like DNA-binding regulatory proteins, the domains must be connected for a protein to possess biological activity. A connection between domains may play an important structural role in positioning domains with respect to one another, or the connection may merely tether two domains within a certain distance of one another. In a number of DNA-binding proteins, including AraC, linker regions are known to restrain domains of DNA half-sites within a certain distance apart (2, 5, 9, 10, 19, 20), but in other cases interdomain linkers in DNA-binding proteins play additional roles in protein regulation (13, 24, 30). In the work described here, we have addressed the roles of the linker amino acids in the gene regulatory protein AraC.

The presence of L-arabinose induces *Escherichia coli* to synthesize the gene products necessary for the uptake and catabolism of the sugar (7). AraC, a positive- and negative-acting transcriptional regulator, controls this response (8). In the absence of arabinose, AraC represses expression of the *ara* p_{BAD} and p_C promoters by formation of a DNA loop (6, 15, 16), and in the presence of arabinose, AraC unloops and activates transcription from the p_{BAD} promoter as well as activating transcription from the *ara* p_{FGH} and p_E promoters (12, 28).

AraC protein possesses four important functions relative to the work to be described here. They are dimerization, binding of arabinose, binding to DNA, and activation of transcription. Three lines of experiments have led to the identification of the domain structure of AraC and assignment of the protein's functions to the domains. The first was the isolation of insertions within AraC that left the protein functional (9). One of these insertion mutations not only retained full activity, but it and mutations containing larger insertions at the same point enabled the DNA binding domains of the protein to contact half-sites on the DNA more widely separated than normal. The inference from these experiments was that the insertions were located in a nonessential region of the protein and probably in a linker region connecting the DNA binding domain to the dimerization domain.

The second line of experiments was the behavior of chimeric proteins consisting of portions of AraC fused to the DNA binding domain of LexA or a leucine zipper dimerization domain (4, 9). These experiments located one domain of the protein to amino acids 1 to 170 and a second domain to amino acids 178 to 291. The experiments also showed that the N-terminal domain possessed dimerization and arabinose binding properties and that the C-terminal domain possessed DNA binding and transcription activation properties.

The third line of experiments has been the crystallization and structure determination of a peptide containing amino acids 1 to 177 of AraC (26). The structure shows that amino acids 1 to 166 specify a compact domain that binds arabinose and dimerizes the protein. Beyond residue 166 lies an unstructured region that serves as the linker between the protein's two domains.

The domain structure thus determined for AraC is consistent with the information that can be gleaned from examination of amino acid sequence homologs. Aside from proteins that induce arabinose operons in other bacteria, currently the sequence databases contain no homologs to the dimerization domain of AraC with sufficiently high similarity to assure structural similarity (23). The situation for the DNA binding domain of AraC is different. The first homologs discovered to this region were the regulators of the rhamnose operon in *E. coli*, RhaR and RhaS (29). Since then, the databases have grown to contain more than 70 proteins possessing significant sequence homology to the DNA binding domain of AraC. The fact that AraC protein's DNA binding and transcription activation module is so widely used in nature further raises the question of whether the segment that links this domain to the rest of the protein plays an active or passive role in the function of the protein both in AraC and in the other homologs.

To date, none of the studies on AraC have answered the question of whether the amino acid residues connecting the dimerization and DNA binding domains of the protein play a direct role in generating its response to arabinose. Since we now know the identities of the linker amino acids in AraC, we have specifically altered them in order to test their roles in the function of the protein.

* Corresponding author. Mailing address: Biology Department, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218. Phone: (410) 516-5206. Fax: (410) 516-5213.
† Present address: Program in Molecular and Cellular Biology, Oklahoma Medical Research Foundation, Oklahoma City, OK.

TABLE 1. Oligonucleotides used in linker mutagenesis[a]

Use	Sequence
Mutagenize linker region with long-way-around PCR	1. 5'CCG CCC GAA GAC AGT CGC TC C ATC CAC CGA TGG ATA ATC GG 3' *BbsI* binding site X XXX XXX XXX XXX XXX XXX XX Mutagenize AraC amino acids 171-177
	2. 5'CGC GCG GAA GAC GAG CGA CTC GTT AAT CG 3' *BbsI* binding site Hybridizes with *araC*
Insert *PvuII* in *araC*	1. 5'CCG CCC GAA GAC CAG CTG TTA CTG CGG CGC ATG GAA GCG ATT 3' *BbsI* site *PvuII* site Codes for AraC amino acids 157-166
	2. 5'CGC GCG GAA GAC AAC AGC TGC TCA AGC AGA TTT ATC GCC AGC AG 3' *BbsI* site *PvuII* site Codes for AraC amino acids 149-158
Insert *SphI* in *araC*	1. 5'CCG CCC GAA GAC GCA TGC CAG TAC ATC AGC GAT CAC CTG GCA 3' *BbsI* site *SphI* site Codes for AraC amino acids 181-190
	2. 5'CGC GCG GAA GAC TGC CAT GCC TCG CGT ACC CGA TTA TCC ATC GGT GG 3' *BbsI* site *SphI* site Codes for AraC amino acids 172-182
Mutagenize *araC* residues 171-178 between *SphI* and *PvuII* sites	1. 5'CTG TTA CTG CGG CGC ATG GAA GCG ATT AAC GAG TCG CTC CAT CCA CCG ATG *PvuII* site XXX XXX XXX XXX GAT AAT CGG GTA CGC GAG GCA TGC CAG TAC 3' XXX XXX XXX XXX *SphI* site
	2. 5'GTA CTG GCA TGC CTC GCG 3' *SphI* site Codes for AraC amino acids 179-184

[a] Oligonucleotide pairs used to mutagenize the AraC linker region are shown. Shaded rectangles represent the bases changed from the *araC* sequence to the sequence of the restriction endonuclease sites.

MATERIALS AND METHODS

Strains, plasmids, and media. All constructs used in arabinose isomerase assays were transformed into strain RE1 (Δ*araC-leu-1022 araB*$^+$*A*$^+$*D*$^+$Δ*lac*-74 *galK* Str^r *recA938*::*cat*) (11). For β-galactosidase assays, constructs were transformed into strain BS1, RV Δ*lac-74*, Δ*ara-leu-498 B1 Su* F$^-$ Str^r(λ *ara*I_1-I_1-*lacZ*) and RE5 (Δ*araC-leu-1022 araB*$^+$*A*$^+$*D*$^+$ Δ*lac-74 galK* Str^r [λ *ara*I_1-I_2-*lacZ*]). BS1 is a derivative of strain RS817, and RE5 is a derivative of strain SH321 (11). Both strains were constructed by λ phage-mediated transduction (17). Wild-type AraC protein was expressed from plasmid pAB1003 (3). Indicator plates to distinguish AraC$^-$ from AraC$^+$ constructs were made as described by Schleif and Wensink (25).

Randomizing the linker. Two PCR methods were used to randomize the linker of AraC. First, using the technique of Stemmer and Morris (27), we amplified the entire plasmid with long-way-around PCR using two oligonucleotides that incorporated *Bbs*I class IIS restriction endonuclease sites adjacent to the DNA coding for the linker region (Table 1). A class IIS restriction enzyme cuts the DNA to one side of its asymmetric binding site. Thus, primers can be designed such that the restriction enzyme cleavage removes the nucleotides constituting the enzyme's cleavage site and leaves an overhanging end (Fig. 1). The same can be done at the other end of the DNA, and the single-stranded overhanging regions can be made complementary. This method permits complete freedom in positioning PCR primers used for mutagenesis or gene construction. The AraC linker region was mutagenized by using an oligonucleotide with degenerate sequence over the region coding for amino acids 171 to 177. About one change in the 21-base mutagenized region was made per oligonucleotide (Table 1).

In the second method, we used the same *Bbs*I-long-way-around PCR technique to introduce two unique restriction endonuclease sites flanking the *araC* sequence coding for the interdomain linker. The DNA coding for AraC amino acids Q157 and L158 was changed from CAATTG to CAGCTG. This makes a *Pvu*II restriction endonuclease binding site. DNA coding for A181 and C182 was changed from GCTTGT to GCATGC. This forms an *Sph*I restriction endonuclease site. In both cases, the resulting amino acid sequence of AraC was unaltered. Following the insertion of the unique sites, the entire *araC* gene was sequenced by double-stranded sequencing (14). An oligonucleotide was synthesized spanning the region between the unique sites with degenerate DNA sequence in the region coding for the linker. An average of 1 base change per oligonucleotide was made in the 21 bases coding for the linker. A second oligonucleotide was synthesized (Table 1) that was hybridized to the 3′ nonran-

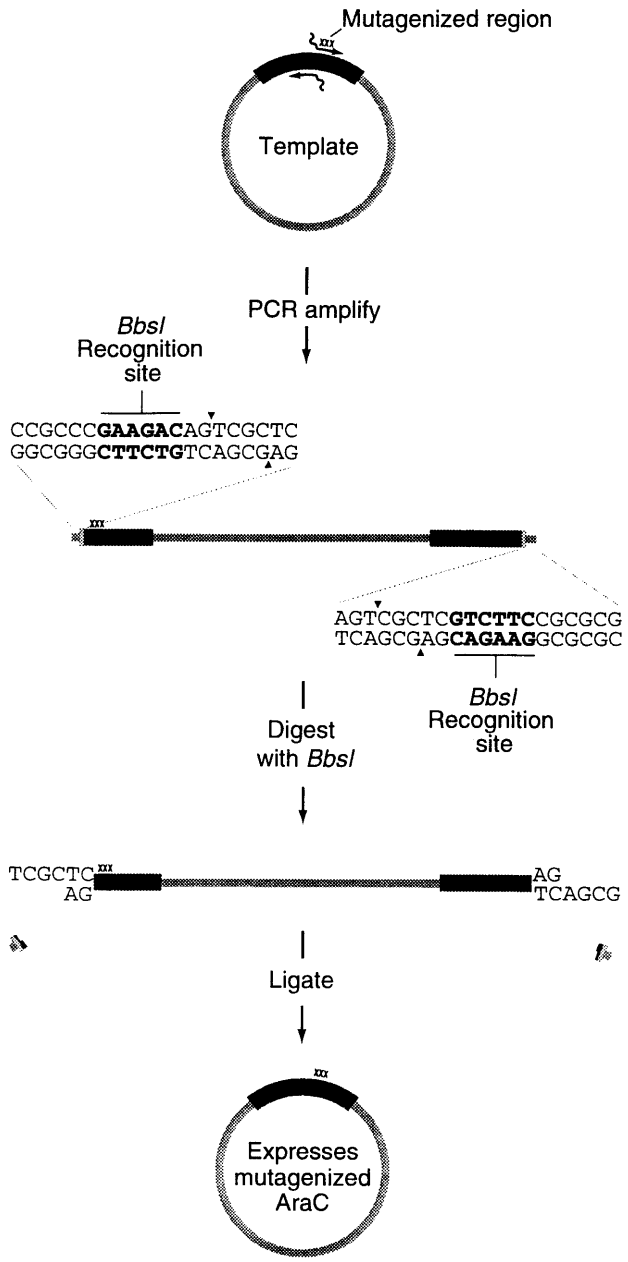

FIG. 1. Mutagenizing the interdomain linker with long-way-around PCR. The first oligonucleotide codes for a *Bbs*I restriction endonuclease site and hybridizes to template DNA coding for AraC amino acids 169 to 177. DNA coding for the linker region, amino acids 171 to 177, was mutagenized. The letter X designates the bases that were mutagenized. The second oligonucleotide codes for a *Bbs*I restriction endonuclease site and hybridizes to the opposite strand of template DNA coding for AraC amino acids 165 to 170. Following a PCR reaction, DNA of the correct size was purified, digested with *Bbs*I, and ligated to reform the plasmid circle.

domized region of the first oligonucleotide in 10 mM Tris-Cl, pH 8, and 10 mM $MgCl_2$ by heating the oligonucleotides to 65°C and allowing them to cool slowly to room temperature. The DNA was made double stranded by adding the Klenow fragment of DNA polymerase I and 0.2 mM deoxynucleotides and heating at 37°C for 1 h. The DNA was then cleaved with *Sph*I and *Pvu*II and cloned into the *Sph*I and *Pvu*II sites in *araC*. Since only a small region of the construct was subjected to in vitro replication in this step, no mutations outside of the region coding for the linker should be generated.

Changing large regions of the linker sequence. To change the sequence of the minimal linker region and part of the maximal linker region, five pairs of oligonucleotides were synthesized. These pairs of oligonucleotides were hybridized together in 10 mM Tris-Cl, pH 8, and 10 mM $MgCl_2$ by heating the oligonucleotides to 65°C and allowing them to cool slowly to room temperature. All oligonucleotides code for an *Sph*I and *Pvu*II restriction site on either end. Following hybridization, the double-stranded DNA was digested with restriction endonucleases and cloned into the *Sph*I and *Pvu*II sites in the plasmid carrying *araC*. The constructs were then sequenced with double-stranded sequencing. As shown in Table 2, the oligonucleotides change part of the sequence of the AraC linker region to TTAGN, TTAGTTAGS, TTAGTTAGN, EFRGSRS, or EFRG SR.

To make the AraC dimerization domain (amino acids 1 to 170)-λ CI repressor linker (amino acids 93 to 131)-AraC DNA binding domain (amino acids 178 to 291) construct, the *lexA* region of a plasmid coding for the fusion protein AraC dimerization domain-λ CI repressor linker-LexA DNA binding domain was replaced by the *araC* region of a plasmid coding for the fusion protein C/EBP leucine zipper-λ CI repressor linker-AraC DNA binding domain (9).

Enzyme assays. Plasmid pAB1003, which expresses wild-type AraC (3), and plasmids coding for proteins with mutant linkers were transformed into RE1 for arabinose isomerase assays. Arabinose isomerase assays, in which the rate of formation of ketose from arabinose is measured by the cysteine-carbazole method, were performed as described by Schleif and Wensink (25), and units reported are the average of three independent assays. For β-galactosidase assays, cells were grown in minimal salts media with glycerol and Casamino Acids (21, 25) and assays were performed as described by Miller (17). Wild-type AraC and proteins with mutant linkers were expressed in strains BS1 and RE5. All values reported are the average of β-galactosidase levels from two independent assays. Because the results from the independent assays rarely differed from one another by as much as 20%, we report the assay averages without error bars.

Immunoblot assay. Cells carrying the plasmids containing the various *araC* mutants were grown in YT medium to about 2×10^8 cells per ml. First, 100 μl of cells was spun down, lysed (4), and loaded on a sodium dodecyl sulfate gel. After electrophoresis, the proteins were transferred to nitrocellulose by electroblotting (BioRad) (22) and incubated with a monoclonal antibody specific for the C-terminal end of AraC protein. Antigen was detected by using alkaline phosphatase-conjugated anti-mouse secondary antibody (Promega) with a Protoblot Western Blot AP system from Promega.

RESULTS

Mutating the linker region. DNA coding for the minimal linker region of AraC, residues 171 to 177 (9), was mutagenized by two PCR methods. The first requires only two oligonucleotides in a long-way-around PCR reaction and is shown in Fig. 1. Both methods are described in Materials and Methods and use oligonucleotides shown in Table 1. Candidate plasmids were transformed into $AraC^-$ bacteria and plated on indicator plates. Both $AraC^+$ and $AraC^-$ candidates were sequenced in the linker region to determine which mutations in the linker are compatible with protein activity and which abolish it. Because the majority of the transformed colonies generated by both methods were $AraC^+$, it is unlikely that a linker mutation making the protein $AraC^-$ and a site-specific second site suppressor elsewhere in the protein together generate an $AraC^+$ phenotype.

Initial studies revealed that both $AraC^+$ and $AraC^-$ phenotypes were found in mutants possessing identical changes in the linker region. The existence of such pairs indicates that at least some of the $AraC^-$ mutants result from alterations elsewhere in the protein, presumably a consequence of PCR mistakes. Therefore, only the $AraC^+$ phenotypes are meaningful in these experiments. Figure 2, which shows all the $AraC^+$ mutants isolated and sequenced, displays the variety of amino acid substitutions found in the linker region. Note that although the nucleotides coding for Arg-177 were mutagenized, no $AraC^+$ colonies were found with alterations at this site. Several constructs contained multiple nonconservative amino acid changes, and in one construct, residue L158 outside of the linker was deleted and residues H171, P172, P173, and D175 were changed to D, A, E, and G, respectively, without abolishing protein activity (Fig. 2). The variety of mutations in the linker that allow protein activity shows that the amino acid sequence of the linker need not be strictly maintained.

We tested all of the linker mutants shown in Fig. 2 for

TABLE 2. Activity of linker variants[a]

Protein	Linker region, amino acids 167-177	Immunoblot assay	I_1-I_2-*lacZ* -ara	I_1-I_2-*lacZ* +ara	I_1-I_1-*lacZ* -ara	I_1-I_1-*lacZ* +ara
No AraC			60	50	10	20
Wild type AraC	N E S L H P P M D N R		50	5200	1900	4000
Linker variants	N E S L H P P I D N R		20	4300	2000	3400
	N E S L **D A E** M **G** N R		50	5800	3200	3300
	N E S L H P P M D N **G**		40	70	30	1400
	N E S L **T T A** G N N R		40	100	120	1700
	T T A G T T A G S △ R		30	40	50	640
	T T A G T T A G N N R		40	160	140	2400
	N E S L **E F R** G S R S		50	90	180	1600
	N E S L **E F R** G S R R		20	70	410	1900
	N E S L **←λ CI repressor linker→**		30	50	900	1550

[a] Two different *ara* promoters fused to *lacZ* were used to analyze AraC linker variants. Wild-type AraC protein and AraC proteins with mutations in the linker region were expressed from high-copy-number plasmids. The letters standing for the mutagenized amino acids in the linker are in boldface. △ indicates that an amino acid was deleted. In the protein in which amino acids 171, 172, 173, and 175 are changed to D, A, E, and G, amino acid 158 is deleted. The immunoblot assay of AraC protein and its derivatives was probed with a monoclonal antibody specific for the C-terminal end of AraC. β-Galactosidase assays were used to measure the level of transcriptional activation in the presence and absence of arabinose. All results are expressed as Miller units and are the average of two independent assays.

transcriptional activation from a p_{BAD} promoter by using either arabinose isomerase assays (25) or β-galactosidase assays (17). For arabinose isomerase assays, transcriptional activation was measured from the wild-type p_{BAD} promoter in the *E. coli* chromosome while β-galactosidase assays measured transcription from a p_{BAD} promoter fused to *lacZ*, also in the chromosome. Mutants were considered to have an AraC$^+$ phenotype only if they activated transcription at least half as well as wild-type AraC. β-Galactosidase assay data for two of these AraC$^+$ mutants and one AraC$^-$ mutant, which were made by mutagenizing the minimal linker region, are shown in Table 2. The AraC$^+$ linker mutants, including a mutant in which four of the seven amino acids in the minimal linker are changed, activate transcription like wild-type AraC. A linker protein in which only amino acid 177 has been changed from R to G is unable to activate transcription from this same promoter.

Changing extensive regions of the AraC linker. The mutation study described above suggests that there are few constraints to the identity of individual amino acids within amino acids 171 to 176 of the linker region. In fact, up to four residues can be changed at the same time without inhibiting the protein's response to arabinose. To examine the possibility that larger regions of the linker cannot be changed without altering protein function, we substituted five new linker sequences for the natural one. The altered linker proteins were assayed for their ability to activate transcription in the absence and presence of arabinose from both the wild-type promoter and a special promoter to which AraC binds very tightly.

Three constructs replace parts of the AraC linker with de-

Mutants	Amino Acids in AraC+ Linker H171	P172	P173	M174	D175	N176	R177
1	L	T					
2	Q				G		
3	Q				G		
4	R				A		
5	L						K
*6	D	A	E		G		
7		R					
8		R	A				
9			A				
10			A				
11			S				
12			A	V			
13				V			
14				V			
15				V			
16				I			
17							D

FIG. 2. AraC linker mutants. Amino acid changes in the AraC interdomain linker that result in an AraC$^+$ phenotype are shown. All changes at position R177 resulted in an AraC$^-$ phenotype. * in front of mutant 6 indicates that amino acid L158 was deleted. Arabinose isomerase assays or β-galactosidase assays were used to determine whether AraC proteins with mutations in the linker could activate transcription from a p_{BAD} promoter. Proteins were considered to have an AraC$^+$ phenotype if their levels of activation were greater than 50% of wild-type levels.

rivatives of a linker from a protein with known structure. In these mutant linker proteins, the number of amino acids in the linker is similar to the number found in wild-type AraC. In the first of these constructs, the number of amino acids in the linker was maintained while the amino acid identities were altered to ones that are flexible and stable when exposed to the solvent (1). The sequence of the linker is derived from an interdomain linker in proteinase A (18); the AraC sequence H171, P172, P173, M174, D175 was changed to TTAGN.

In two more constructs, more of the linker region was altered. AraC amino acids N167, E168, S169, L170, H171, P172, P173, M174, and D175 were changed to TTAGTTAGN and AraC amino acids N167, E168, S169, L170, H171, P172, P173, M174, D175, and N176 were changed to TTAGTTAGS. These two linkers are partial duplications of the linker from proteinase A. In the linker region which replaces AraC amino acids 167 to 176, a serine residue is inserted at position 176 instead of an asparagine residue because position 176 is already an asparagine in wild-type AraC. These mutant linker proteins showed no activation of transcription from the wild-type p_{BAD} promoter fused to *lacZ* (Table 2).

Two more linker mutants were generated which contain amino acids not normally found in interdomain linker regions. Three of the five amino acids included in these constructs, E, F, and R, are found less often in linker regions than they are typically found within proteins in general and could, in fact, disrupt the function of an interdomain linker because of their bulky side chains (1). In one construct, AraC amino acids H171, P172, P173, M174, D175, and N176 are changed to EFRGSR, and in another construct, amino acids H171, P172, P173, M174, D175, N176, and R177 are changed to EFRG SRS. These mutant AraC proteins also were inactive on the wild-type p_{BAD} promoter. The protein did not activate transcription in response to arabinose (Table 2).

A final linker mutant was constructed containing a much longer linker than the one found in wild-type AraC; the 39-amino-acid interdomain linker from the *E. coli* λ CI repressor was substituted for AraC amino acids 171 to 177. Again, changing a large part of the linker region from AraC results in a protein that cannot activate transcription from the wild-type p_{BAD} promoter in response to the sugar arabinose (Table 2).

Linker mutants' activity on promoters with tight DNA binding sites. To explore further the defect in the linker mutants containing multiple amino acid changes, we assayed mutant activities on a derivative of the AraC promoter. The wild-type p_{BAD} promoter consists of a strong binding I_1 half-site adjacent to a weaker binding I_2 half-site. If a second I_1 half-site is substituted for the I_2 site, a much stronger AraC binding site is formed and AraC protein will activate transcription from this even in the absence of arabinose (21). Presumably, on this promoter, the extra binding energies available from the strong binding sites substitute for arabinose-induced changes in the protein and allow wild-type AraC to bind and activate transcription in the absence of arabinose.

Except for the protein containing the λ CI repressor linker, the proteins with large stretches of changes in the AraC linker do not activate transcription from a promoter with an I_1-I_1 binding site fused to *lacZ* in the absence of arabinose but do so in the presence of arabinose (Table 2). Replacing the interdomain linker of AraC protein with a linker region from a different protein, therefore, does not abolish the conformational change in the response to arabinose, suggesting that the linker region is not required for this response.

Unlike the shorter peptides derived from the proteinase A interdomain linker, the longer linker from the λ CI repressor protein enables AraC to activate transcription at 50% of the wild-type level even in the absence of arabinose. In the presence of arabinose, transcriptional activation with this mutant increases to the same level as with wild-type AraC (Table 2).

Amount of mutant linker proteins in cells. It is possible that replacing the interdomain linker in AraC protein with a foreign linker makes the protein susceptible to proteolytic attack. If this were true, then lack of transcriptional activation could be low because of a lack of protein rather than an altered arabinose response of the protein. To measure the levels of intact AraC protein in cells, we performed immunoblot assays using antibodies specific for AraC protein.

As shown in Table 2, all of the mutant linker proteins, including those that respond to arabinose only in the presence of a tight DNA binding site, are present in amounts similar to those of wild-type protein. The mutant linker proteins as well as wild-type protein were expressed in the same strain and from the same expression vector. Except for the protein with the linker from the λ CI repressor, all of the linker mutants electrophorese to the same position on the gel as wild-type protein. The protein with the λ CI repressor has 32 more amino acids than wild-type AraC and migrates more slowly on the gel.

In some of the lanes containing linker variants (Table 2), a faster migrating band is seen which is also recognized by the antibody specific for the C-terminal end of AraC. This fragment could represent the DNA binding domain of AraC which has been released because of proteolysis of the linker. Because the intensity of the bands representing the intact mutant proteins does not decrease relative to the band for wild-type protein when this faster migrating protein is present, the majority of the protein from the linker variants must remain intact.

DISCUSSION

The experimental evidence we have presented in this paper shows that individual amino acids in the linker region between the dimerization and DNA binding domains of AraC, amino acids 166 to 176, may be altered without substantial effect on the arabinose response of the protein. When four amino acids in the region were altered, we found the same result, but when larger segments of the linker were changed, the protein was partially defective. Two main conclusions can be drawn from the data: one that is relatively firm and one that is less definitive and more in the nature of a conjecture.

The fact that the individual amino acids and even a group of four could be altered without apparent effect on the functioning of AraC indicates that the amino acids of the linker region do not play crucial structural and functional roles in the arabinose response of the protein. These amino acids thus appear largely to connect the dimerization and DNA binding domains of the protein.

The fact that no active AraC mutants were found in which residue 177 was altered implies that either this one residue does play an important role in linking the domains of the protein or that this amino acid is within the DNA binding domain and plays an important role there.

The proteins in which linkers containing five or more amino acids of the linker were altered or substituted with linkers from other proteins were not normal. These proteins did not induce the wild-type *ara* p_{BAD} promoter in the presence of arabinose. Also, they did not activate transcription in the absence of arabinose from the promoter containing an I_1-I_1 binding site in place of the weaker-binding wild-type I_1-I_1 site whereas wild-type AraC does fully activate p_{BAD} in the absence of arabinose from such a promoter. When arabinose was present and the I_1-I_1 p_{BAD} promoter was used, these proteins did activate tran-

scription. Thus, both the better binding site and the presence of arabinose are necessary for them to be active. Clearly, the linker region is not crucial to generating an arabinose response by the protein. We can picture these proteins as requiring both the added binding energy provided by the I_1-I_1 site and the presence of arabinose to drive them into an inducing state. We might conjecture that the inactive linkers were insufficiently flexible and AraC containing such linkers requires extra energy to be driven into the inducing conformation. This result is similar to that found for LexA, in which individually changing five different amino acids in the LexA protein interdomain linker region does not affect the ability of the repressor protein to bind DNA but changing all five amino acids at the same time prevents the protein from binding DNA (19).

ACKNOWLEDGMENTS

We are grateful to the members of this lab for critical discussion and comments on the manuscript, to Beatrice Saviola for strain BS1, and to Beth MacDougall Shackleton for technical assistance.

This material is based upon work supported by the National Institutes of Health grant GM18277 to Robert Schleif.

REFERENCES

1. **Argos, P.** 1990. An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. J. Mol. Biol. **211:**943–958.
2. **Astromoff, A., and M. Ptashne.** 1995. A variant of λ repressor with an altered pattern of cooperative binding to DNA sites. Proc. Natl. Acad. Sci. USA **92:**8110–8114.
3. **Brunelle, A., and R. Schleif.** 1989. Determining residue-base interactions between AraC protein and *araI* DNA. J. Mol. Biol. **209:**607–622.
4. **Bustos, S. A., and R. Schleif.** 1993. Functional domains of the AraC protein. Proc. Natl. Acad. Sci. USA **90:**5638–5642.
5. **Carra, J. H., and R. Schleif.** 1993. Variation of half-site organization and DNA looping by AraC protein. EMBO J. **12:**35–44.
6. **Dunn, T., S. Hahn, S. Ogden, and R. Schleif.** 1984. An operator at −380 base pairs that is required for repression of *araBAD* operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. Proc. Natl. Acad. Sci. USA **81:**5017–5020.
7. **Englesberg, E., J. Irr, J. Power, and N. Lee.** 1965. Positive control of enzyme synthesis by gene C in the L-arabinose system. J. Bacteriol. **90:**946–957.
8. **Englesberg, E., C. Squires, and F. Meronk.** 1969. The L-arabinose operon in *Escherichia coli* B/r: a genetic demonstration of two functional states of the product of a regulator gene. Proc. Natl. Acad. Sci. USA **62:**1100–1107.
9. **Eustance, R., S. Bustos, and R. Schleif.** 1994. Reaching out: locating and lengthening the interdomain linker in AraC protein. J. Mol. Biol. **242:**330–338.
10. **Gubler, M., and T. Bickle.** 1991. Increased protein flexibility leads to promiscuous protein-DNA interactions in type IC restriction-modification systems. EMBO J. **10:**951–957.
11. **Hahn, S.** 1984. Ph.D. thesis. Brandeis University, Waltham, Mass.
12. **Hendrickson, W., C. Stoner, and R. Schleif.** 1990. Characterization of the *araFGH* and *araJ* promoters, J. Mol. Biol. **215:**497–510.
13. **Horii, T., T. Ogawa, T. Nakatani, T. Hase, H. Matsubara, and H. Ogawa.** 1981. Regulation of SOS functions: purification of *E. coli* LexA protein and determination of its specific site cleaved by the RecA protein. Cell **27:**515–522.
14. **Kraft, R., J. Tardiff, K. Krauter, and L. Leinwand.** 1988. Using mini-prep plasmid DNA for sequencing double stranded templates with Sequenase. BioTechniques **6:**544–546.
15. **Lobell, R., and R. Schleif.** 1990. DNA looping and unlooping by AraC protein. Science **250:**528–532.
16. **Martin, K., L. Huo, and R. Schleif.** 1986. The DNA loop model for *ara* repression: AraC occupies the proposed loop sites *in vivo* and repression-negative mutations lie in these same sites. Proc. Natl. Acad. Sci. USA **83:**3654–3658.
17. **Miller, J.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Moult, J., F. Sussman, and M. James.** 1985. Electron density calculations as an extension of protein structure refinement. *Streptomyces griseus* protease A at 1.5 Å resolution. J. Mol. Biol. **182:**555–566.
19. **Oertel-Buchheit, P., T. Schmidt-Dorr, M. Granger-Schnarr, and M. Schnarr.** 1993. Spacing requirements between LexA operator half-sites can be relaxed by fusing the LexA DNA binding domain with some alternative dimerization domains. J. Mol. Biol. **229:**1–7.
20. **Reece, R., and M. Ptashne.** 1993. Determinants of binding site-specificity among yeast C-6 zinc cluster proteins. Science **261:**909–911.
21. **Reeder, T., and R. Schleif.** 1993. AraC protein can activate transcription from only one position and when pointed in only one direction. J. Mol. Biol. **231:**205–218.
22. **Sambrook, J., E. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Sander, C., and R. Schneider.** 1991. Database of homology-derived protein structures and the structural meaning of sequence alignment. Proteins **9:**56–68.
24. **Sauer, R., M. Ross, and M. Ptashne.** 1982. Cleavage of the lambda and P22 repressors by recA protein. J. Biol. Chem. **257:**4458–4462.
25. **Schleif, R., and P. Wensink.** 1981. Practical methods in molecular biology. Springer-Verlag, New York.
26. **Soisson, S., B. B. MacDougall-Shackleton, R. Schleif, and C. Wolberger.** Submitted for publication.
27. **Stemmer, W., and S. Morris.** 1991. Enzymatic inverse PCR: a restriction site independent, single-fragment method for high-efficiency, site-directed mutagenesis. BioTechniques **13:**214–220.
28. **Stoner, C., and R. Schleif.** 1983. The *araE* low affinity L-arabinose transport promoter. J. Mol. Biol. **171:**369–381.
29. **Tobin, J., and R. Schleif.** 1987. Positive regulation of the *Escherichia coli* L-rhamnose operon is mediated by products of tandemly repeated regulatory genes. J. Mol. Biol. **196:**789–799.
30. **Vershon, A., and A. Johnson.** 1993. A short, disordered protein region mediates interactions between the homeodomain of the yeast α2 protein and the MCM1 protein. Cell **72:**105–112.