doi:10.1016/S0022-2836(03)00262-6 J. Mol. Biol. (2003) 328, 85–93

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Mutational Analysis of Residue Roles in AraC Function

Jennifer J. Ross, Urszula Gryczynski and Robert Schleif*

Department of Biology, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA

The previously isolated hemiplegic, induction-negative, repression-positive mutants, H80R and Y82C, were found to be defective in the binding of arabinose. Randomization of other residues close to arabinose in the three-dimensional structure of AraC or that make strong interactions with arabinose yielded induction-negative, repression-positive mutants. The induction and repression properties of mutants obtained by randomizing individual residues of the N-terminal arm of AraC allowed identification of the domain with which that residue very likely makes its predominant interactions. Residues 8–14 of the arm appear to make their predominant interaction with the DNA-binding domain. Although the side-chain of residue 15 interacts directly with arabinose bound to the N-terminal dimerization domain, the properties of mutant F15L indicate that this mutation increases the affinity of the arm for the DNA-binding domain.

© 2003 Elsevier Science Ltd. All rights reserved

Keywords: AraC protein; mutation analysis; arm–domain interactions; fluorescence; computation

*Corresponding author

Introduction

The regulatory protein of the L-arabinose operon in *Escherichia coli*, AraC, senses the presence of arabinose,[1] and when the sugar is present, transmits this information from the arabinose-binding domains to the protein's DNA-binding domains.[2] The activities that are seen in AraC of sensing an intracellular condition and then transmitting this information to another part of the protein are common to many allosteric proteins. While much is known about the phenomenon of allostery in general, our understanding of the detailed atomic motions, energetics, and their consequences in such systems is limited. The experiments described here were designed to improve our understanding of these processes in AraC.

In the absence of arabinose, the two DNA-binding domains of the dimeric AraC protein are held in position with respect to each other such that the protein prefers to bind to the well separated half-sites I_1 and O_2 and form a DNA loop that blocks transcription from the p_c and p_{BAD} promoters (Figure 1).[3–8] When arabinose is present, the DNA-binding domains are no longer as constrained and it is energetically more favorable for AraC to bind to the adjacent I_1 and I_2 half-sites, a position from which AraC activates transcription of the adjacent p_{BAD} promoter.[9–11] It is the N-terminal arms of AraC that control the positioning of the DNA-binding domains.[6–12] In the absence of arabinose, the arms, which extend from the dimerization domains, bind to the DNA-binding domains, and hold them such that DNA looping is facilitated. When arabinose is present, it binds to the dimerization domains, and the N-terminal arms bind over the arabinose. The repositioned N-terminal arms can no longer bind to the DNA-binding domains, thus freeing them to bind to I_1 and I_2. This is termed the light-switch mechanism, because the system is transcriptionally "on" when the arms are in one position and "off" when they are in the other position. We have addressed two basic questions in the work described here. First, can mutations that prevent arabinose from binding be identified? Second, since the N-terminal arms at times interact with the dimerization domains, and at other times with the DNA-binding domains, some residues may have a stronger interaction with one or the other of the domains of AraC. If so, can we identify the domain of stronger interaction for each such residue?

Present address: J. J. Ross, HFM-481, 1401 Rockville Pike, Rockville, MD 20852-1448, USA.

Abbreviations used: I^-, induction-negative; R^+, repression-positive.

E-mail address of the corresponding author: rschleif@jhu.edu

0022-2836/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved

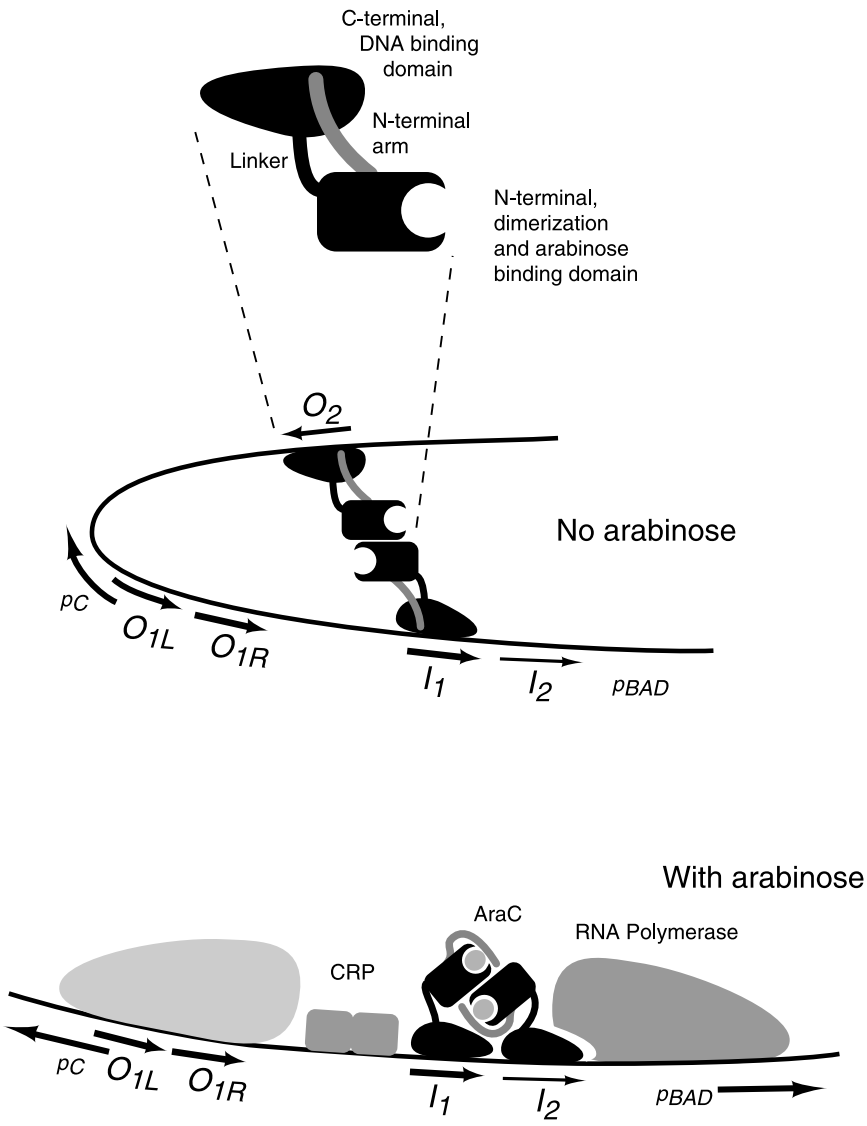

Figure 1. Light switch mechanism of AraC plus the domain and arm structure showing DNA looping in the absence of arabinose and *cis* binding in the presence of arabinose that induces transcription from the p_{BAD} promoter.

Results

Measurement of arabinose binding

Hemiplegic mutants of AraC protein have been identified in the dimerization/arabinose-binding domain that repress transcription from p_{BAD} normally, but are unable to activate transcription in response to arabinose.[13] The two strongest such mutations are H80R and Y82C. The tyrosine residue lies within 2.5 Å and the histidine residue lies within 4 Å of arabinose.[14] These mutant proteins are presumed to fold correctly because they are able to repress transcription. They could therefore be defective in binding arabinose, or defective in communicating to the DNA-binding domain the

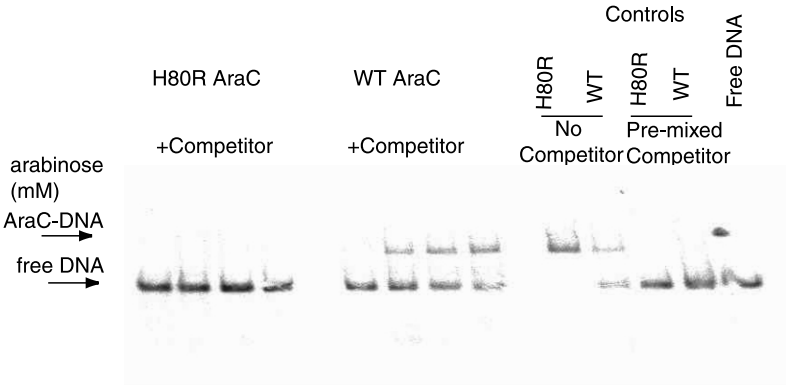

Figure 2. *In vitro* DNA binding response of wild-type and H80R AraC. Cell extracts containing AraC were incubated with the indicated concentrations of arabinose and 10,000 cpm/lane of 64 bp ^{32}P end-labeled dsDNA (I_1–I_2). A 100-fold molar excess of unlabeled 64 bp DNA (I1–I2) was added as competitor before or after addition of arabinose. The samples were incubated and subjected to electrophoresis in a 6% polyacrylamide gel.

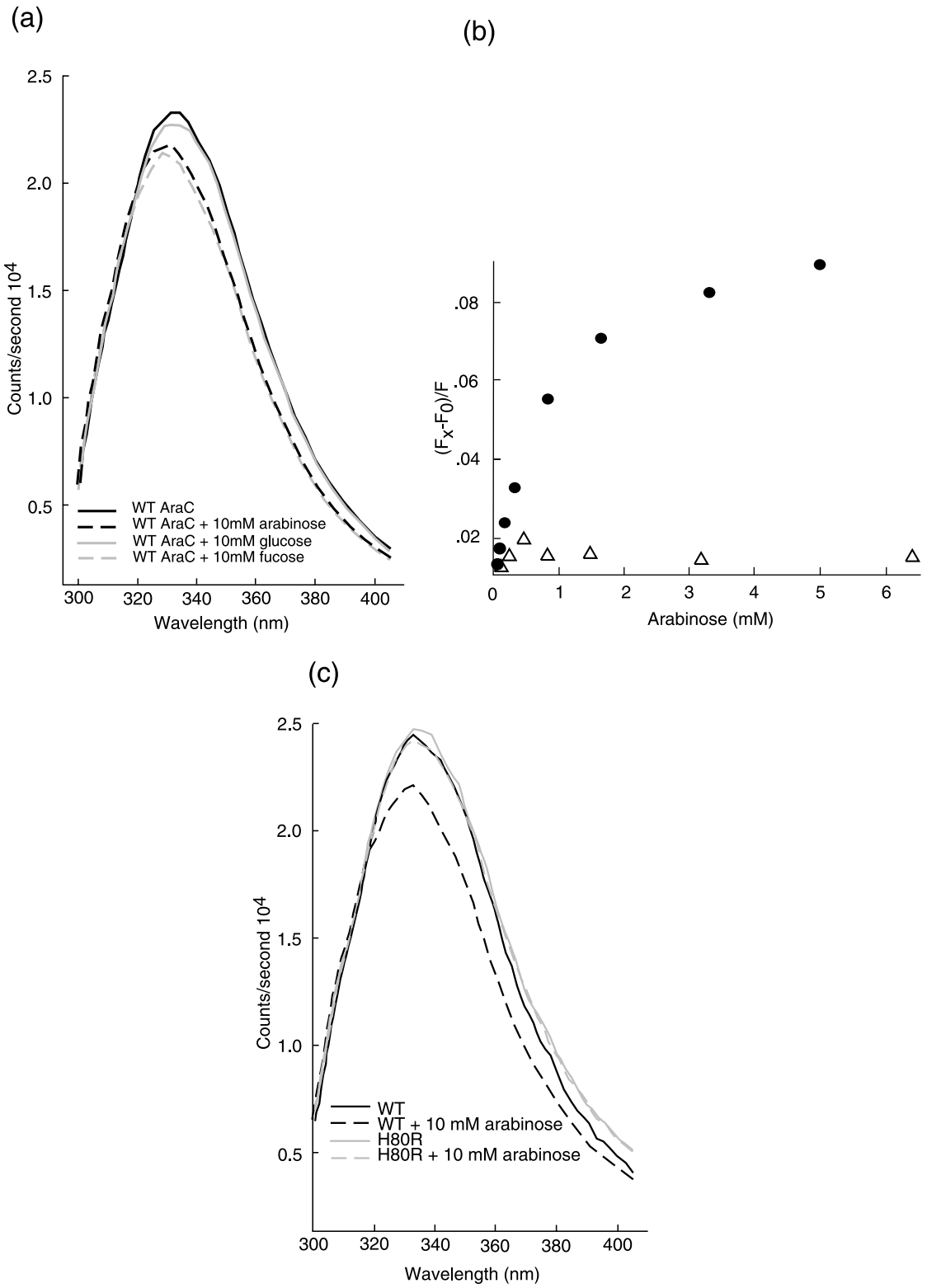

Figure 3. (a) The fluorescence emission spectra of wild-type AraC dimerization domain with no sugar, glucose, arabinose, or fucose addition. Purified, C-terminal His_6-tagged, AraC dimerization domain at 10 μg/ml was incubated with each sugar (10 mM) separately and excited at 280 nm. (b) Measurement of arabinose affinity of wild-type, circles, and Y82C, triangles. The area under each fluorescence emission spectrum of AraC in the presence of the indicated concentration of arabinose, *Fx*, was subtracted from the area under the fluorescence emission spectrum of AraC when buffer lacking sugar was added, *F*0. The resulting difference was expressed relative to the peak fluorescence of AraC before any additions, *F*, and plotted. (c) The fluorescence emission spectra of wild-type and H80R purified, C-terminal His_6-tagged, AraC dimerization domains with and without addition of arabinose. AraC dimerization domain at 10 μg/ml was incubated with the indicated concentrations of arabinose and excited at 280 nm.

Table 1. Properties of residues and mutants close to arabinose

Position	Closest approach to arabinose (Å)	Interaction energy with arabinose, (kcal/mol)	No. I^-R^+ mutants[a]	No. I^-R^- mutants	No. I^+ mutants
I24	2.5	−2.25	2/2	0/2	0/2
I26	4.5	−0.09	2/48	19/48	27/48
R38	2.5	−31	4/4	0/4	0/4
I46	2.5	−0.6	2/48	45/48	1/48
H80	4.0	−0.5	9/9	0/9	0/9
Y82C	2.5	−0.05	1/1	0/1	0/1
H93	5.5	−4.97	1/2	1/2	0/2
W95	3.0	−5.54	0/5	5/5	0/5
Y97	4.0	−0.09	16/48	31/48	1/48

The hemiplegic mutants, I^-R^+, were: I24, E, P; I26, R; R38, A, Q, D, V, L; I46, R; H80, A, R, H, F, C, W, S, Y, Q; H93, D; W95, H; Y97, not sequenced.

[a] I is induction and R is repression.

status of arabinose binding. We have determined which step is defective in these two mutants.

Before examining directly whether arabinose binds to the mutants H80R or Y82C, or whether they are defective for some other reason, it seemed advisable to confirm *in vitro* that at least one lacks an arabinose response. To do this, we used the DNA migration retardation assay to check for the arabinose-dependent stabilization of AraC binding to the *araI* site.[7] Figure 2 shows that the binding of wild-type AraC, but not of H80R AraC, is stabilized by arabinose because the wild-type protein, but not the mutant protein, remained bound to the labeled DNA when cold competitor DNA was added. Hence, H80R AraC displays a mutant phenotype *in vitro* with no additional macromolecule present other than DNA.

The intrinsic tryptophan fluorescence of AraC protein provides a convenient assay for the binding of arabinose to the protein.[15,16] For C-terminal His_6-tagged, wild-type AraC dimerization domain, residues 1–178, we found a 5–10% reduction in the total fluorescence emission and an approximately 2 nm blue shift in the wavelength of the maximum emission in response to the binding of arabinose (Figure 3(a)). Additionally, the L-arabinose analog, D-fucose, which also binds in the pocket, produced a similar fluorescence change. No fluorescence change was observed in response to the addition of D-glucose, sucrose, glycerol, or lactose, which is expected because these compounds are not known to bind to AraC, as they do not activate transcription by AraC. The half-maximal change in fluorescence of wild-type dimerization domain occurs at about 0.4 mM arabinose (Figure 3(b)). Mutants Y82C and H80R both display no significant fluorescence change upon arabinose addition (Figure 3(b) and (c)), and these proteins most likely did not bind arabinose in the experiment.

Additional arabinose pocket mutations

Can additional hemiplegic induction-negative (I^-), repression-positive (R^+) mutations be isolated? The mutant searches that yielded H80R and Y82C found these mutations multiple times,[13] suggesting that not many additional mutations of this type can be found. On the other hand, directing mutagenesis to critical residues may reveal additional, mutations of the same phenotype. Therefore, we mutagenized residues whose side-chains either closely approach arabinose or that make a strong interaction with arabinose and screened for the I^-R^+ phenotype (Table 1). We found hemiplegic, I^-R^+ mutants at each of the positions tested, except W95.

Randomization to identify arm residues involved in induction and repression

The N-terminal arm of AraC appears to interact primarily with the dimerization domain of the protein in the presence of arabinose, and with the DNA-binding domain in the absence of arabinose.[6,7] Any particular residue of the arm might make significant contacts with either or both domains. Our objective here was to identify whether a residue predominantly interacts with a domain, and if so, to identify which domain.

Since the 20 amino acids are of different sizes and charges, it is likely that the majority of single amino acid substitutions at a position in a protein will eliminate an interaction made by the wild-type residue and only a few, if any, will create an interaction as strong as or stronger than those made by the wild-type residue (see Discussion). Applying this reasoning to the N-terminal arm of AraC, it seems likely that if a residue in the arm is changed to a variety of other residues, then the majority of the introduced amino acids will result in weaker interactions than those made by the wild-type residue. Weakening the interactions between the arm and the DNA-binding domain should reduce the ability of AraC to loop the DNA and hence would increase the activation of p_{BAD} by AraC, especially in the absence of arabinose. On the other hand, weakening the interactions between the arm and the dimerization domain or arabinose would increase the arm's

Table 2. Properties of mutations in the N-terminal arm

	N6	D7	P8	L9	L10	P11	G12	Y13	S14	F15	N16	A17	H18
Trp													U
Tyr	Ind	Ind				C		* * *	C	U	Ind		U[a]
Phe		U	C			C				* * *	Ind		
Met				Ind	C[b]					U			
Ile		Ind	C		Ind								
Leu	Ind		Ind[b]	* * *	* * *		C		Ind[c]	U	C	Ind	C
Val	Ind	U		C					Ind			Ind	
Thr	Ind				C	C	C	C	Ind		C	U	
Pro			* * *	C[b]	C	* * *		C	C[d]			U	C[d]
Cys	Ind	Ind	Ind	C	Ind	C	Ind	U	C				Ind
Ser	C	Ind	Ind	C	Ind	C	C	C	* * *		Ind	Ind	
Ala		Ind[e]	Ind	C	C		Ind		C	U	Ind	* * *	Ind
Gly			C	C			* * *	Ind		U			
Arg		Ind	C		C	C	C	C	Ind				Ind
Lys	Ind			C				C					
His		Ind	C	C				C				Ind	* * *
Glu					C			C	C	U			
Gln					C					U			
Asn	* * *				C	C			Ind		* * *		
Asp	Ind	* * *		C	C	C	C			U	C[f]	Ind	Ind
Ind	8	7	4	1	3	0	3	1	5	0	4	7	4
U	0	2	0	0	0	0	0	1	0	8	0	2	2
C	0	0	5	9	9	8	5	7	5	0	3	0	2
Total	8	9	9	10	12	8	8	9	10	8	7	9	8

Mutants were classified as inducible (Ind) if they induce greater than 20% of wild-type, constitutive (C) if they express at least 20% of the wild-type induced level in the absence of arabinose, and uninducible (U), if induction is less than 20% of wild-type: * * * indicates the wild-type residue. The actual uninduced and induced expression levels as a percentage of fully induced wild-type were: for N6, Y <5, 130; L <5, 160; V 100, 140; T <5, 100; C <5, 200; S 100, 150; K <5, 400; D 100, 100. For D7, Y <5, 30; F <5, 20; I <5, 50; V <5, 20, C <5, 130; S <5, 100; A <5, 25; R <5, 50; H <5, 50. For P8, F 50, 200; I 30, 200; L <5, 100; C 10, 250; S 20, 200, A <5, 100; G 100, 200; R 30, 300, H 30, 200. For L9, M 10, 150; V 100, 100, P 100, 200; C 200, 350; S 200, 200; A 90, 90; G 250, 150; K 100, X; H 300, 250; D 150, 150. For L10, M 60, 150; I 10, 100; T 100, 200; P 150, 200; A 100, 150; R 150, 200; E 100, 200; Q 100, 300; N 200, 200; D 200, 200. For P11, Y 30, 100; F 100, 150; T 60, 200; C 50, 200; S 60, 200; R 200, 180; N 60, 200; D 50, 300. For G12: L 40, 200; T 50, 400; C 20, 200; S 40, 300; A 10, 200; R 100, 200; D 30, 300. For Y18, T 10, 100; P 200, 100; C 3, 10; S 100, 100; G <5, 50; R 100, 100; K 120, 120; H 50, 150; E 100, 80. For S14, Y 60, 60; L <5, 75; V <5, 100; T 10, 100; P 100, 100; C 100, 200; A 100, 150; R <5, 100, E 100, 100, N 20, 100. For F15, Y <5, <5; M <5, 10; L <5, <5; A <5, 10; G <5, <5; E <5, 10, Q <5, 10; D <5, <5. For N16, Y 20, 200; F 20, 200, L 30, 200; T 100, 200; S 20, 200; A 5, 200; D 60, 180. For A17, L <1, 60; V <1, 30; T <1, 18; P <1, 10; S <1, 100; H 5, 40; D 9, 140. For H18, W <1, 24; Y <5, <5; L 100, 100; P 100, 150; C 20, 150; A 20, 125; R 5, 110; D 10, 130.

[a] Reed & Schleif.[13]
[b] Saviola *et al.*[7]
[c] M. Ghosh, personal communication.
[d] Wallace.[26]
[e] M. Wu, personal communication.
[f] Wu & Schleif.[12]

binding to the DNA binding domain and hence would result in reduced expression from p_{BAD}in the presence and absence of arabinose. Thus, if the majority of the alterations at a position elevate p_{BAD} expression, we conclude that the wild-type residue most likely makes its more important contacts to the DNA-binding domain, and if p_{BAD} expression is reduced, that the wild type residue most likely makes its more important interactions with the dimerization domain or arabinose.

We used oligonucleotides containing mixed bases at the codons of interest to randomize individually residues 6–18 of the arm of AraC. Resultant mutants were screened for loss of repression ability and for induction ability by growing cells in the absence and in the presence of arabinose and assaying arabinose isomerase. Mutants unable to induce transcription from p_{BAD} even in the presence of arabinose were tested for their ability to repress transcription from p_c so as to identify mutants that have lost both induction and repression abilities. Table 2 shows the properties of the arm mutants. All substitutions at residue 6 leave the protein able to induce transcription to at least half-maximal level, indicating that this is not a critical residue for induction or repression. This is in agreement with the finding that deleting the first six residues of AraC does not inactivate inducibility.[7] Residues 9–14 seem most important for repression, as many substitutions at these positions yielded protein that constitutively activates transcription, and residue Phe15 is apparently important for induction of transcription, as all substitutions there left the protein unable to induce transcription.

An exceptional mutant, F15L

All changes at arm position 15 left AraC unresponsive to arabinose. This suggests that the side-chain of Phe15 makes critical interactions with the dimerization domain or with arabinose. Not surprisingly, Phe15 is the only residue of the N-terminal arm whose side-chain contacts

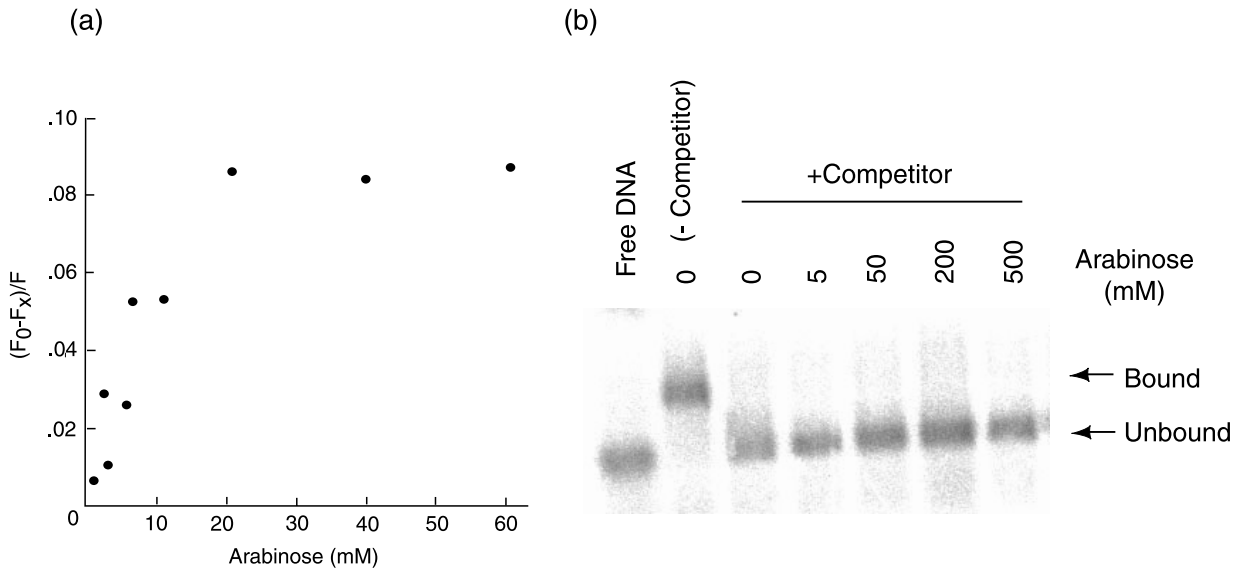

Figure 4. (a) Measurement of affinity of arabinose binding to F15L AraC dimerization domain. Purified, C-terminal His_6-tagged, AraC dimerization domain at 10 μg/ml was incubated with different concentrations of sugar and excited at 280 nm. The area under each fluorescence emission spectrum of AraC, *Fx*, was subtracted from the area under the fluorescence emission spectrum of AraC when buffer lacking sugar was added, *F*0. The resulting difference was expressed relative to the peak fluorescence of AraC before any additions, *F*, and plotted. (b) Lack of an *in vitro* DNA binding response when arabinose is added to full-length F15L AraC. Cell extracts containing AraC were incubated with the indicated concentrations of arabinose and 10,000 cpm/lane of 64 bp ^{32}P end-labeled dsDNA (I_1–I_2). A 100-fold molar excess of unlabeled 64 bp DNA (I1–I2) was added as competitor before or after addition of arabinose. The samples were incubated and subjected to electrophoresis in a 6% polyacrylamide gel.

arabinose directly in the pocket. Because mutant F15L exhibits a particularly strong R^+, I^- phenotype, we expected that the strength of the Leu15–arabinose interaction would be substantially weaker than the Phe15–arabinose interaction. The calculated interaction energy between Leu15 and arabinose was not, however, dramatically weaker, −0.6 kcal/mol (1 cal = 4.184 J) compared to −1.2 kcal/mol for Phe15. Because the phenotype of F15L was stronger than what seemed to be explained by the energetics of interaction with arabinose, we examined the biochemical properties of the F15L mutant more carefully.

In the fluorescence assay, the mutant F15L dimerization domain shows a half-maximal response to arabinose at about 10 mM (Figure 4(a)) as compared to the wild-type dimerization domain, whose half-maximal response is at about 0.4 mM. In the DNA-binding stabilization assay described earlier that utilizes full-length AraC rather than dimerization domain, we found that mutant F15L was not stabilized on DNA even by the very high concentration of 500 mM arabinose (Figure 4(b)), showing that the presence of the DNA-binding domain dramatically weakens the strength of arabinose binding to the dimerization domain in the mutant.

Discussion

AraC protein utilizes one of the simplest possible mechanisms for allosteric regulation. The N-terminal arm of the protein communicates to the DNA-binding domain the status of arabinose binding by the dimerization domain. Even in this simple system, our understanding of ligand binding and arm binding is far from allowing us to build new systems that operate by the same principles. In the work described here, we sought to take a step in this direction with AraC by assigning roles in this signaling system to several residues near the arabinose-binding pocket and to residues of the N-terminal arm.

Mutant AraC proteins H80R and Y82C repress the p_c promoter in the absence of arabinose, but do not induce the p_{BAD} promoter in the presence of arabinose. Due to the proximity of these residues to the arabinose-binding pocket in the protein, we thought it is likely that the structural changes generated by these mutations interfere with arabinose binding. To test this, we used the change in intrinsic tryptophan fluorescence of the protein as a monitor of arabinose binding because this signal is more likely to reflect arabinose binding than a signal like DNA-binding affinity, which measures both arabinose binding and transmission of the binding signal to the DNA-binding domain. Unfortunately, full-length AraC that contains both the dimerization domains and DNA-binding domains is difficult to purify and difficult to use in biochemical experiments, mainly due to the presence of its unstable DNA-binding domains. Therefore, we measured the fluorescence of wild-type and mutant dimerization domain proteins.

While wild-type dimerization domain of AraC shows a 5–10% reduction in intrinsic tryptophan fluorescence and a 2 nm emission maximum wavelength shift in response to the binding of arabinose, neither of the dimerization domain mutant proteins showed any response to arabinose. It is conceivable that the absence of a fluorescence response in the two proteins results from the fortuitous cancellation of both intensity and wavelength effects in the two proteins. We think it much more likely, however, that the absence of the responses in the two mutant proteins results from their failure to bind arabinose.

One of the two previous reports on fluorescence measurements of AraC protein observed fluorescence changes of full-length AraC protein and of the dimerization domain very similar to those we observed for the dimerization domain.[16] This is not an unexpected signal size, as five of the six tryptophan residues in AraC lie in the dimerization domain, but only one of them contacts arabinose.[14] Additionally, there is almost no change in the structure of the dimerization domain upon the binding of arabinose, and with all the tryptophan residues well separated from one another,[14] sizeable fluorescence changes are not likely. We do not understand the origin of the large (45%) reduction in tryptophan fluorescence, that was reported by Wilcox.[15]

We then tested the effects of randomizing additional residues that lie close to arabinose. We reasoned that the desired mutations could either generate steric constraints that interfere with arabinose binding or they could eliminate interactions that normally hold arabinose in place. At six of these locations, we found mutants whose induction abilities were reduced to less than 20% of wild-type, but whose repression abilities remained greater than 50% of wild-type. In one of the sites tested, Trp95, all mutants were I^-,R^-, probably indicating that all these mutants were unable to fold properly. Arg38 is calculated to make a particularly strong interaction with arabinose, and therefore, it is not surprising that all mutations in this position yielded strong I^- mutants. Such mutants were probably not seen by Reed & Schleif[13] in the original mutant screen because that involved passing a plasmid carrying the AraC gene through a mutator strain. This method limited the type of substitutions that could result, since the most frequently observed base-pair change was from AT to GC.[13] The codon for Arg38 is CGA. Therefore, if the AT base-pair were changed to GC, the codon would become CGG and would still specify arginine.

As explained earlier, in the absence of arabinose, the N-terminal arm of AraC contacts the DNA-binding domain, and in the presence of arabinose, contacts the dimerization domain. Any particular residue of the arm might therefore make significant contact with one, the other, or with both domains. We reasoned that if a residue of the N-terminal arm makes important contacts, then most amino acid substitutions of this residue will make weaker interactions rather than stronger interactions. There is precedence for such reasoning. Out of ~730 single amino acid mutations made over the 61 amino acid residue DNA-binding domain of *lac* repressor, only 21 mutations at six different positions increased the operator binding affinity.[17] Similarly, only two out of 53 single residue mutations in the gp120 binding site of CD4 resulted in increased interaction affinity between gp120 of HIV and CD4.[18]

Weaker interactions of the N-terminal arm of AraC with the DNA-binding domain would result in higher expression from p_{BAD} in the absence and in the presence of arabinose, whereas weaker interactions with the dimerization domain would result in lower expression. Therefore, we randomized, individually, residues 6–16 of the N-terminal arm of AraC and examined expression. Most changes in residues 7–14 of the arm of AraC yielded AraC that failed to repress normally so that the uninduced expression level of the p_{BAD} promoter was elevated substantially. We conclude that these residues make their more important interactions with the DNA-binding domain. Most alterations at position 15 left AraC uninducible by arabinose. This result was not surprising, because the side-chain of Phe15 contacts arabinose directly.

Calculation of the strengths of the F15 and L15 interactions with arabinose suggested that a weakening of the interaction of residue 15 with arabinose was insufficient to account for the phenotype of the mutant. Indeed, comparison of the arabinose-binding properties of the L15 dimerization domain and of the arabinose response in DNA binding of full-length mutant L15 AraC showed the binding of arabinose to the full protein to be much weaker than to the dimerization domain. A simple hypothesis that is consistent with the data is that the mutant arm is held more tightly to the DNA-binding domain than the wild-type arm. In such a case, the mutant protein might fail to show an arabinose response in the DNA-binding assay because the arm cannot be released from the DNA-binding domain and allow the DNA-binding domains to orient themselves for proper DNA binding. Alternatively, it is possible that arabinose binding to the dimerization domain requires that the arm close over the arabinose. In such a case, the coupling of binding energies would necessitate that an increase in the binding energy of the arm to the DNA-binding domain would result in a comparable decrease in the binding energy of arabinose to the dimerization domain. Our experiments do not resolve the two possibilities, but we can conclude that in mutant F15L, the strength of the arm-DNA binding domain interaction is increased substantially.

In our analysis of the roles of specific residues in the action of the allosteric protein AraC, two complementary approaches were utilized that should be applicable in many other systems. Within the dimerization domain, whose structure

in the presence and in the absence of arabinose is known with high resolution, computation proved to be an effective tool. Computation predicted that the N-terminal arm of AraC in the F15L mutant protein would bind to the dimerization domain stronger than the wild-type arm and not weaker, as expected from visual inspection. We then confirmed biochemically that the uninducible phenotype of the mutant agreed with the computational result. The second approach involved the DNA-binding domain, whose precise structure is not known. Because the N-terminal arm is bifunctional and makes interactions with either domain of the protein, we randomized individual residues of the arm to determine with which domain a residue made its more important interactions. This approach should be applicable to situations in which a residue has multiple functions.

If we used the criterion of being able to predict function from structure, it would seem that we have much to learn about protein structure and function. On the other hand, we know a lot already. One very important generalization is that most mutations alter structure locally. Categorization of the defects in 4000 single amino acid substitutions in *lac* repressor showed that, for the most part, substitutions affected functions of adjacent portions of the protein rather than of distal regions.[17] The work reported here on AraC protein similarly found that the effects of mutations were local. Changing residues near arabinose affected arabinose binding, and changing residues in the N-terminal arm affected its behavior. Because arm–domain interactions are likely to be found in many systems,[19] the approach we used in determining the domain with which a particular arm residue interacts more strongly is likely to prove of use. Finally, the addition of tools like those used here for the computation of interaction energies and atomic motions are likely to find increasing use as molecular biologists probe more deeply into the functioning of individual proteins.

Materials and Methods

General procedures, plasmids, and strains

Mutations in the AraC gene were made with the Quickchange Site-Directed Mutagenesis kit from Stratagene (La Jolla, CA) starting with 10 ng of plasmid. Oligonucleotides were obtained from IDT (Coralville, IA) with standard desalting purification. Oligonucleotides longer then 39 bases were purified by polyacrylamide gel electrophoresis. For randomization of a residue, the oligonucleotides contained three degenerate bases corresponding to the codon of interest. Often the third base was designated C or T to reduce the frequency of stop codons. All oligonucleotides had at least 50% GC content on both sides of the codon to be mutagenized and, where possible, extended 12 bases on either side of the codon to be mutagenized.

Sequencing was done at the Johns Hopkins University DNA analysis facility on 2.4 μg of mutagenized plasmid DNA. This was isolated with a Qiagen (Valencia, CA) Plasmid Mini kit or 0.3 μg of PCR product from the mutagenized plasmid DNA after purification using the Qiagen (Valencia, CA) PCR purification kit.

Plasmid pWR03 is derived from pSE380 (amp^r) with the AraC gene inserted between the *Nco*I and *Sac*I sites.[13] Mutagenized plasmids were transformed into strain SH321 (F^-, Δ*araC-1022* Δ*lac74*, *galK*$^-$, *str*^{r}) to test for transcription from p_{BAD} by colony color on arabinose/tetrazolium indicating plates or by measuring the levels of arabinose isomerase. Candidate plasmids that were to be tested for repression of p_c activity were then transformed into strain SH10 (F^+, *araC:lacZ fusion*, *leu*$^-$ *str*^{r}, Δ*lac araD*$^-$). Colonies were scored for p_c activity by colony color on X-gal plates or by measurement of the level of β-galactosidase expressed from p_c.

Arabinose isomerase and β-galactosidase assays were performed on cells grown on M10 minimal salts growth medium supplemented with 0.4% (v/v) glycerol, 10 μg/ml of thiamine, and 0.02% (w/v) L-leucine.[13,20,21]

Plasmid AraCTF[22] encodes for amino acid residues 1–182 of AraC, a leucine-glutamic acid linker, and a His_6 tag cloned into the *Nde*I and *Ava*I sites, downstream of a phage T7 promoter, in pET-21b (amp^r) from Novagen, Madison, WI. AraCTF constructs were transformed into BL21 (DE3) cells (Stratagene, La Jolla, CA), which express phage T7 RNA polymerase, for overexpression of AraC dimerization domain. The His_6 tag was not removed, since it was not found to affect the arabinose affinity.

Radioactive DNA migration retardation assay

Cell lysates from a 2 ml culture were made from SH321 cells transformed with pWR03 plasmid. A 64 bp dsDNA template derived from pES51[23] contained the I_1–I_2 AraC binding site (underlined) was formed by annealing complementary oligonucleotides and end labeling with ^{32}P using phage T4 kinase. (p_{BAD} site-derived oligonucleotides: Oligonucleotide 1: 5′CTTTGCTAGCCCATAGCATTTTTATCCATAAGATTAGCGGATCCTACCTGACGCTTTTTATCGC

Oligonucleotide 2: 5′GCGATAAAAAGCGTCAGGTAGGATCCGCTA

Purification of dimerization domain

Ni-NTA agarose beads (Qiagen, Valencia, CA) were used to purify C-terminal His_6-tagged AraC dimerization domain over-expressed from the AraCTF plasmid essentially as described,[24] except that precipitation in ammonium sulfate was not required. RNase was present at 10 μg/ml and DNase at 5 μg/ml during the purification. To remove eluent imidazole, the 4–5 ml of supernatant was applied to a Sephadex G25 column in buffer of 15 mM Tris–HCl (pH 8.0), 100 mM NaCl, 5% glycerol, 1 mM β-mercaptoethanol. Samples were checked for size and purity after being subjected to electrophoresis on an SDS/14% polyacrylamide gel.

Fluorescence assay of arabinose binding

Fluorescence measurements were made at room temperature on a T-format PTI fluorimeter (Monmouth Junction, NJ) with 75W xenon arc lamp. For measurement of arabinose binding to the dimerization domain by intrinsic tryptophan fluorescence, excitation was set at 280 nm, emission was integrated over 300–400 nm, with slit-widths set to give 5 nm band-widths, and data

were collected at 1 nm intervals at one second/point. A sample (3 ml at a concentration of 10 μg/ml) of protein was incubated with varying amounts of arabinose in 15 mM Tris–HCl (pH 8.0), 100 mM NaCl, 5% glycerol, 1 mM β-mercaptoethanol.

Calculations on AraC

All calculations were performed with the CHARMM program[25] using the Protein Data Bank coordinates of AraC, 2ARC as input and the par_all22_prot and par_all22_sugar parameters and non-bonded interactions. Residues 7–166 of a monomer were immersed in a 12 Å shell of TIP3 water molecules and energy-minimized before the calculation of distances or surface exposure. Residues 7 through 120 of the dimerization domain, which includes the N-terminal arm and the β-sheet jelly roll that binds arabinose, were also immersed in a 12 Å shell of TIP3 water molecules, energy minimized, brought to 300 K and equilibrated at 300 K before simulation totaling 1 ns. Coordinates were saved each 10^{-12} seconds and were used in the calculation of average interaction energies.

Acknowledgements

We thank members of the laboratory for discussions. This work was supported by NIH grant GM18277.

References

1. Englesberg, E., Irr, J., Power, J. & Lee, N. (1965). Positive control of enzyme synthesis by gene c in the L-arabinose system. *J. Bacteriol.* **90**, 946–957.
2. Schleif, R. (2000). Regulation of the L-arabinose operon of *Escherichia coli*. *Trends Genet.* **16**, 559–565.
3. Dunn, T., Hahn, S., Ogden, S. & Schleif, R. (1984). An operator at −280 base-pairs that is required for repression of ara_{BAD} operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. *Proc. Natl Acad. Sci. USA*, **81**, 5017–5020.
4. Martin, K., Huo, L. & Schleif, R. (1986). The DNA loop model for *ara* repression: AraC protein occupies the proposed loop sites *in vivo* and repression-negative mutations lie in these same sites. *Proc. Natl Acad. Sci. USA*, **83**, 3654–3658.
5. Lobell, R. & Schleif, R. (1990). DNA looping and unlooping by AraC protein. *Science*, **250**, 528–532.
6. Seabold, R. & Schleif, R. (1998). Apo–AraC actively seeks to loop. *J. Mol. Biol.* **278**, 529–538.
7. Saviola, B., Seabold, R. & Schleif, R. (1998). Arm–domain interactions in AraC. *J. Mol. Biol.* **278**, 539–548.
8. Hahn, S., Dunn, T. & Schleif, R. (1984). Upstream repression and CRP stimulation of the *Escherichia coli* L-arabinose operon. *Proc. Natl Acad. Sci. USA*, **81**, 5017–5020.
9. Greenblatt, J. & Schleif, R. (1971). Arabinose C protein: regulation of the arabinose operon *in vitro*. *Nature New Biol.* **233**, 166–170.
10. Hendrickson, W. & Schleif, R. (1984). Regulation of the *Escherichia coli* L-arabinose operon studied by gel electrophoresis DNA binding assay. *J. Mol. Biol.* **178**, 611–628.
11. Zhang, X., Reeder, T. & Schleif, R. (1996). Transcription activation parameters at *ara* p_{BAD}. *J. Mol. Biol.* **258**, 14–24.
12. Wu, M. & Schleif, R. (2001). Mapping arm-DNA-binding domain interactions in AraC. *J. Mol. Biol.* **307**, 1001–1009.
13. Reed, W. & Schleif, R. (1999). Hemiplegic mutations in AraC protein. *J. Mol. Biol.* **294**, 417–425.
14. Soisson, S., MacDougall-Shackleton, B., Schleif, R. & Wolberger, C. (1997). Structural basis for ligand-regulated oligomerization of AraC. *Science*, **276**, 421–425.
15. Wilcox, G. (1974). The interaction of L-arabinose and D-fucose with AraC protein. *J. Biol. Chem.* **249**, 6892–6894.
16. Lauble, H., Georgalis, Y. & Heinemann, U. (1989). Studies on the domain structure of the *Salmonella typhimurium* AraC protein. *Eur. J. Biochem.* **185**, 319–325.
17. Suckow, J., Markiewicz, P., Kleina, K., Miller, J., Kisters-Woike, B. & Müller-Hill, B. (1996). Genetic studies of the lac repressor XV: 4000 single amino acid substitutions and analysis of the resulting phenotypes on the basis of the protein structure. *J. Mol. Biol.* **261**, 509–523.
18. Tsui, P., Sweet, R., Sathe, G. & Rosenberg, M. (1992). An efficient phage plaque screen for the random mutational analysis of the interaction of HIV-1 gp120 with human CD4. *J. Biol. Chem.* **267**, 9361–9367.
19. Schleif, R. (1999). Arm–domain interactions in proteins: a review. *Proteins: Struct. Funct. Genet.* **34**, 1–3.
20. Schleif, R. & Wensink, P. (1981). *Practical Methods in Molecular Biology*, Springer-Verlag, New York.
21. Miller, J. (1972). *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
22. LaRonde-LeBlanc, N. & Wolberger, C. (2000). Characterization of the oligomeric states of wild-type and mutant AraC. *Biochemistry*, **39**, 11593–11601.
23. Huo, L., Martin, K. & Schleif, R. (1988). Alternative DNA loops regulate the arabinose operon in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **85**, 5444–5448.
24. Ghosh, M. & Schleif, R. (2001). Biophysical evidence of arm–domain interactions in AraC. *Anal. Biochem.* **295**, 107–112.
25. Brooks, B. R., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 178–217.
26. Wallace, R. W. (1985). Ph.D. Thesis, Biological Sciences, University of California, Santa Barbara.

Edited by M. Gottesman

(Received 7 January 2003; received in revised form 13 February 2003; accepted 13 February 2003)