doi:10.1006/jmbi.2001.4531 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2001) **307**, 1001–1009

JMB

Mapping Arm-DNA-binding Domain Interactions in AraC

Martin Wu and Robert Schleif*

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

**Corresponding author*

AraC protein, the regulator of the L-arabinose operon in *Escherichia coli* has been postulated to function by a light switch mechanism. According to this mechanism, it should be possible to find mutations in the DNA-binding domain of AraC that result in weaker arm-DNA-binding domain interactions and which make the protein constitutive, that is, it no longer requires arabinose to activate transcription. We isolated such mutations by randomizing three contiguous leucine residues in the DNA-binding domain, and then by systematically scanning surface residues of the DNA-binding domain with alanine and glutamic acid. As a result, a total of 20 constitutive mutations were found at ten different positions. They form a contiguous trail on the DNA-distal face of the DNA-binding domain, and likely define the region where the N-terminal arm that extends from the N-terminal dimerization domain contacts the C-terminal DNA-binding domain.

Keywords: arm interactions; domain; AraC; arm-domain; alanine-glutamic scan

Introduction

In the absence of arabinose, the dimeric AraC protein loops between the *ara* p_{BAD} I_1 and O_2 half-sites that are separated by 210 bp (Figure 1).[1-3] This loop interferes with access of RNA polymerase to the p_{BAD} and p_C promoters, and is the source of repression by AraC.[4] Upon the addition of arabinose, AraC ceases looping and instead binds to the two adjacent I_1 and I_2 half-sites, where it acts positively to activate transcription from the p_{BAD} promoter.[3] The change in the DNA-binding properties of AraC appears to be regulated by the N-terminal arm consisting of 18 amino acid residues that extend from the dimerization domain.[5-8] In the absence of arabinose, the arms are thought to bind to the DNA-binding domains and hold them such that their relative orientation favors formation of the DNA loop. When arabinose is present, the arms are thought to prefer binding to the dimerization domains, thus freeing the DNA-binding domains. Hence, the activity of AraC is regulated by switching the binding of its arms between its dimerization domains and its DNA-binding domains. Consequently, mutations in either the arm or the two domains should change the behavior of AraC. For example, AraC should become constitutive, that is, not require arabinose to shift it to an inducing state, as a result of mutations in either the arm or the DNA-binding domain that sufficiently weaken the arm-DNA-binding domain interactions.

Although mutations have been reported in the N-terminal arm of AraC that generate constitutivity, presumably by weakening interactions between the arm and the DNA-binding domain,[5,8] mutations with a similar effect, but lying in the DNA-binding domain have not previously been identified. We sought such mutations, because first, their existence would provide additional support for the light switch mechanism and, second, such mutations could locate the residues in the DNA-binding domain that contact the arm.

Alanine scanning is widely used to probe protein-protein interactions.[9-12] Substitution of a residue with alanine residue eliminates all side-chain atoms beyond the β-carbon atom and, therefore, would appear capable of identifying an amino acid whose side-chain atoms participate in an interaction. Unfortunately, however, entropy-enthalpy compensation nearly neutralizes the effect of the loss of the interactions and minimizes the magnitude of effects from alanine residue substitutions.[13] Hence, the effects from alanine residue substitutions often are not large, and frequently amino acid residues are missed that participate in the interactions being assayed. In addition, the resi-

E-mail address of the corresponding author: bob@gene.bio.jhu.edu

0022-2836/01/041001–9 $35.00/0

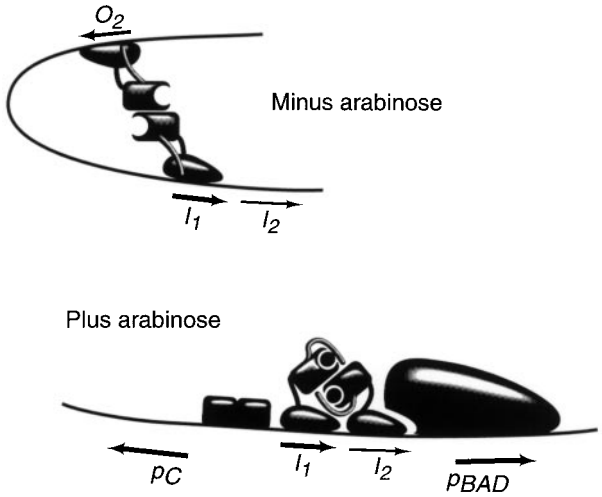


Figure 1. Binding sites in the *ara* p_{BAD} regulatory region and the mechanism of DNA looping and unlooping. In the absence of arabinose, AraC loops between O_2 and I_1, and in the presence of arabinose the N-terminal arms are bound over arabinose to the dimerization domains, and AraC is free to bind to I_1-I_2, where it activates transcription from p_{BAD}. The two boxes to the left of AraC bound at I_1-I_2 represents cyclic AMP receptor protein that also stimulates transcription at p_{BAD}.

dues in a protein-protein interface usually do not contribute equally to the overall binding energy, and often a small number of residues account for most of the binding energy. Therefore, residues other than these few can produce no effects in alanine scanning.[11,12] Rather than removing interactions by introduction of alanine residues, replacing a surface amino acid by another residue with a larger side-chain is more likely to generate an observable effect, albeit, with a somewhat lower ability to localize an interaction site. A combination of both alanine and glutamic acid scans should then provide a robust method for elucidating the binding sites of the arm on the DNA-binding domain of AraC. This approach is now possible because the recently available structures of the AraC homologs MarA[14] and Rob[15] allows a rather accurate model of AraC DNA-binding domain to be constructed and surface residues may be targeted for alteration.

Results

Identifying a probable arm interaction site

Before investing in the work of scanning many residues on the surface of the DNA-binding domain for effects resulting from weaker arm-domain interactions, we sought first to determine whether any such interactions exist at all. Mutagenesis of the entire AraC gene followed by selection of constitutive mutations in the past has yielded no mutations in the DNA-binding domain. Therefore, we sought clues implicating any small portion of the domain in interacting with the arm. The following line of reasoning suggested that the three leucine residues (L237, L238, and L239) might interact with L9 or L10 of the arm.

Several constitutive mutations (L9V, L9M, L9P, L9K, L9R, L10M, and D7P8L9L10 → V) have been isolated in the N-terminal arm of AraC.[5,8] Most likely they weaken the arm-DNA-binding domain interactions for the following reasons: first, substitution of L9 with a number of different amino acid residues all made AraC constitutive. This suggests that the mutations weaken or disrupt an interaction rather than create a new one. Second, residues at those positions interact with dimerization domain exclusively through hydrogen bonds formed between their backbones, the bound sugar arabinose and water molecules.[16] Therefore, changing side-chains at those positions is unlikely to strengthen or create new interactions between the arm and the dimerization domain. We conclude that the side-chains of L9 and L10 probably interact with the DNA-binding domain, and that changing them weakens the interaction.

A potential target for L9-L10 binding was identified by the following reasoning. The mutation S240I appears to strengthen the interaction between the arm and the DNA-binding domain.[6] As this mutation introduces the hydrophobic isoleucine residue adjacent to the three leucine residues L237, L238, and L239, it seems possible that L9 or L10 interacts *via* hydrophobic interactions with L237, L238, or L239, and that the introduction of the isoleucine residue at position 240 allows the hydrophobic interaction to be even stronger. In support of this hypothesis, we found that the mutation D7P8L9L10 → V in the arm rescued the activating ability of S240I. If these ideas are correct, then changing the three leucine residues in the DNA-binding domain should also generate constitutive mutations.

Mutagenesis of the LLL site

Leucine residues 237, 238, and 239 were, therefore, randomized through site-directed mutagenesis. We used randomized nucleotides consisting of 62.5 % of wild-type, and 12.5 % of each of the other three non-wild-type bases for the nine nucleotides coding for the leucine residues. On average, this procedure should produce 3.4 bp changes with the frequencies of single, double and triple base changes being 7.8 %, 18.9 % and 26.4 %, respectively.[17] After mutagenesis and transformation into AraC $^-$ cells, plasmids containing constitutive AraC mutations were selected by their abilities to support growth on minimal arabinose-fucose plates.[18-21] Approximately 2 % of the transformed cells were able to grow on the plates in three days. Candidates were grown in liquid medium, and arabinose isomerase was assayed to identify inducible and constitutive mutants. The entire AraC gene from each candidate was

sequenced and every candidate proved to be mutant. Table 1 shows the sequences found and the induction levels. Notably, all mutants are induced very well by fucose, a natural anti-inducer. At the amino acid level, three were single mutations in the first or second amino acid and the remainder were altered in both the first and second amino acid residues. The third leucine residue remained unchanged. This was not a failure of the randomization process as a number of the sequenced mutants carried silent mutations in the codon for the third amino acid. The most constitutive mutant has the sequence of VQL and it activates the p_{BAD} promoter about eightfold over the basal level in the absence of arabinose. Table 1 shows that the level of constitutivity induced by the mutants was well below the normal full induction level induced by wild-type AraC protein in the presence of arabinose. This suggests that the mutations weakened, but did not abolish the interaction with the N-terminal arm. Upon the addition of arabinose, the constitutive mutants were induced to wild-type or higher than wild-type levels.

Leucine residues are not frequently found on the surface of proteins, and the likelihood that three consecutive leucine residues lie on the surface seems very low. To obtain experimental data bearing on this, we separately randomized each leucine residue on the expectation that surface residues would be highly tolerant to change. Screening on tetrazolium arabinose-indicating plates, we found that only about 5 % of randomized residue 237 or 238 transformants had lost the ability to activate p_{BAD}, but that 30 % of the randomized residue 239 transformants did. As the efficiency of mutagenesis in these experiments was about 50 %, our result implies that L239 is considerably more important, if not essential, in maintaining the stability or inducing activity of AraC protein. This is also consistent with the finding that, amongst those three leucine residues, only L239 is highly conserved in the AraC/XylS family.[22] Since leucine residues 237 and 238 can tolerate many mutations, they are more likely to be located on the surface of the protein.

The MarA and Rob proteins possess sufficient sequence similarity to the DNA-binding domain of AraC that they all should possess the same tertiary structure. Indeed, the recently published MarA and Rob protein structures[14,15] show them to be similar and enabled us to build an AraC DNA-binding domain structural model. Since the last 13 residues of the C terminus are not conserved and are not required for normal AraC function (data not shown), we excluded them from the model. In this model, the side-chains of leucine residues 237 and 238 are solvent exposed and thus accessible to the N-terminal arm from the dimerization domain. Leucine 239, however, is deeply buried inside the protein and contributes to a hydrophobic core of the domain. Overall, the genetic and structural homology-modeling data fit well with each other.

DNA-binding affinity of the LLL mutant

Leucine residues 237, 238 are on the back face of the DNA-binding domain, opposite to its DNA-binding sites, which is an ideal location for the arm to contact. Nonetheless, it is conceivable that mutations at L237 and L238 could affect the structure of the DNA-binding motif and somehow increase the DNA-binding domains' intrinsic affinities to the I_1-I_2 half-sites, thereby generating the constitutive behavior of the mutants. To reduce the likelihood of this possibility, we compared the tightness of the DNA-binding of the mutant VQL, the most constitutive candidate, with that of wild-type AraC. To measure the intrinsic DNA-binding affinities of the domains, the experiments were carried out in the presence of arabinose. In this situation, the N-terminal arm binds over the sugar-binding pocket in the dimerization domain, and hence the affinity measurement is free of the complexity generated by the interactions between the arm and DNA-binding domain. Figure 2 shows that the VQL mutant protein actually dissociates a little bit faster than wild-type AraC from the I_1-I_2 DNA fragment, thus showing the mutation does not increase the affinity of the DNA-binding domain for I_1-I_2 DNA.

Table 1. Abilities of the AraC proteins to activate p_{BAD}

AraC protein	Arabinose isomerase activity in units per cell		
	Minus arabinose	Plus arabinose	Plus fucose
W.T. (LLL)	80	2600	90
VAL	290	2660	3800
VEL	430	2810	2130
VIL	290	3100	1630
VQL	690	3770	3750
VLL	420	2950	2080
VVL	80	4270	570
IEL	330	3170	3660
LEL	180	4100	1390
DLL	200	2800	1720

W.T., wild-type.

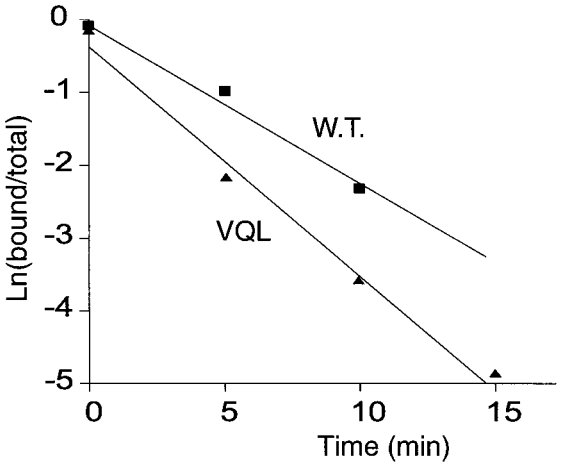


Figure 2. Dissociation kinetics of wild-type AraC and AraC L237V, L238Q, L239L AraC in the presence of arabinose from DNA containing I_1-I_2 half-sites at 50 mM KCl and 2.5 mM $MgCl_2$.

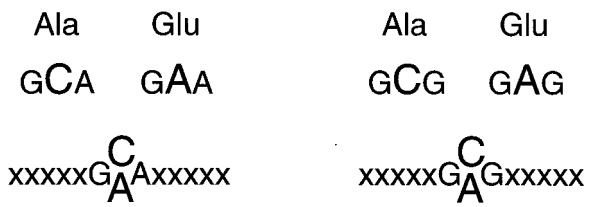


Figure 3. Two possible ways that mixed oligonucleotides may be used to introduce only alanine or glutamic acid at the same position.

Alanine and glutamic acid scanning the AraC DNA-binding domain

As discussed above, alanine substitutions normally generate weak phenotypes. Therefore, we simultaneously carried out both alanine and glutamic acid scans of the AraC DNA-binding domain and examined the effects of these substitutions on transcription from the p_{BAD} promoter. Based on the structural model, a total of 28 and 24 surface residues were selected for alanine and glutamic acid scanning, respectively (see Materials and Methods). Since the codons for alanine and glutamic acid differ by only one base (Figure 3), either of the two mixed-base primers shown can be used to introduce both codons in a single QuickChange™ (Stratagene, La Jolla, CA) reaction that does not introduce codons for any of the other amino acid residues. Screening for constitutive mutations was further facilitated by use of a plasmid encoding both AraC and a reporter green fluorescence protein fused to the p_{BAD} promoter. This allowed a rapid and sensitive visual assay of constitutivity. Candidates with interesting phenotypes were sequenced and their promoter activities were quantified.

The alanine and glutamic acid scanning results are presented in Figure 4. Out of all the residues scanned, only residues L237, L238 and R251, when replaced with alanine residues, showed significant effects on the transcription from p_{BAD} in the absence of arabinose. Glutamic acid substitution of residues 230, 233, 240, 241, 244, 248 and 252, in addition to residues 237, 238 and 251, all generated constitutivity, indicating that these residues are either close to or inside the arm-domain interface. The relative activities of those mutants in the absence of arabinose are shown in Figure 5. Similar to the behavior of the constitutive mutations at L237, L238 and L239, the alanine and glutamic residue-substituted AraC protein in the presence of arabinose activated transcription to levels equal to or above those activated by wild-type AraC induced by arabinose. It is remarkable that the residues form a contiguous path along the back surface of the DNA-binding domain, an ideal region for arm binding.

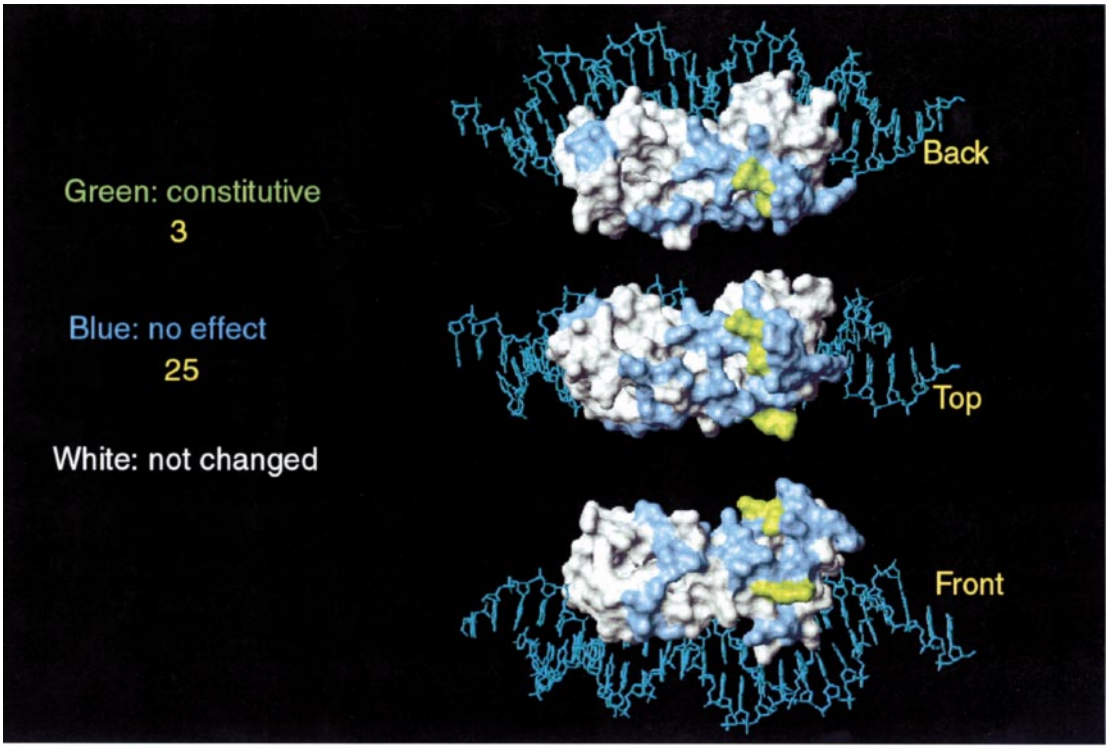
The effects of the scanning on AraC-RNA polymerase interaction

AraC activates transcription at p_{BAD} *via* direct contacts between RNA polymerase and the DNA-binding domain of AraC.[23–25] The actual activating sites on the DNA-binding domain remain elusive, however. Possibly then, the constitutive behavior of the mutants in the DNA-binding domain, especially those from glutamic acid scanning, could result from tighter binding between the AraC DNA-binding domain and RNA polymerase rather than being the result of weaker binding between the AraC DNA-binding domain and the N-terminal arm of AraC. To eliminate this possibility, we compared the interaction between the mutant AraC T241E and RNA polymerase and the interaction between wild-type AraC and RNA polymerase. We did this by measuring the kinetics of open complex formation as a function of RNA polymerase concentration so that the K_d for RNA polymerase binding and the rate of open complex formation, k_2 could be determined. The results (Figure 6 and Table 2), show that both parameters are nearly unchanged for the T241E mutation, and thus that the mutation does not significantly strengthen the interaction between AraC and RNA

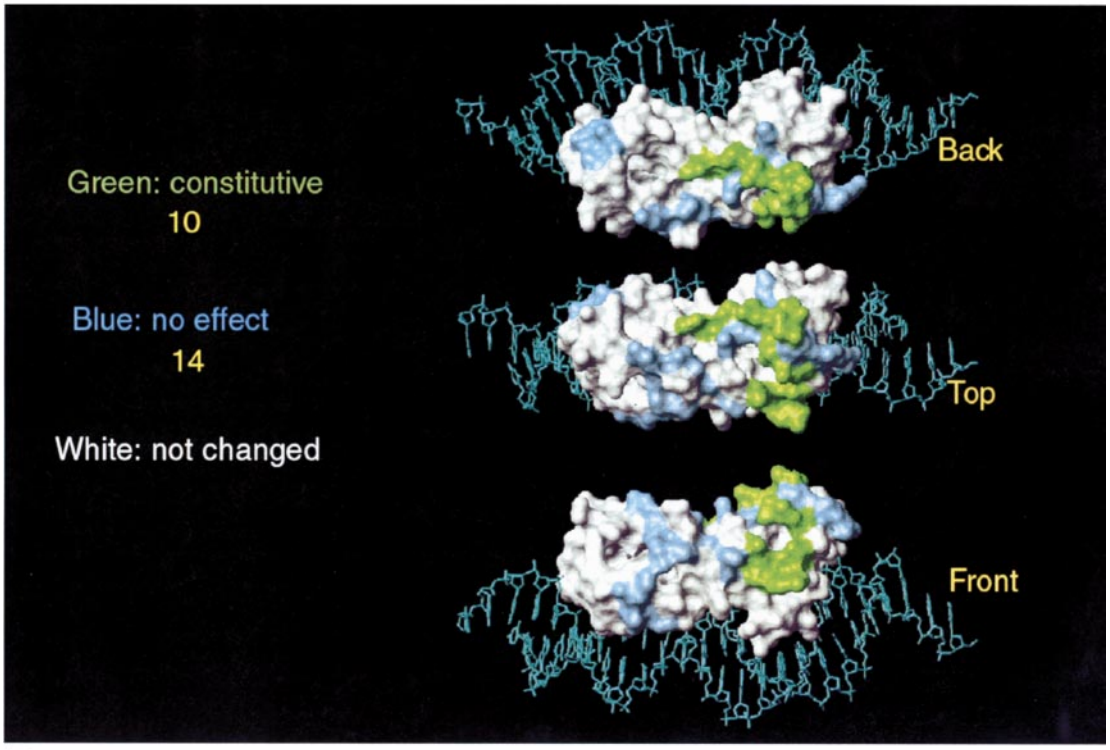
Table 2. Kinetic parameters of open complex formation on the *ara P3* promoter

AraC protein	K_d (nM)	$1/k_2$ (seconds)
WT	0.30	55
T241E	0.39	59

WT, wild-type.



A



B

Figure 4. The amino acid residues that, when changed to (a) alanine or (b) glutamic acid, make AraC become constitutive are shown in green. Those residues unresponsive to the scanning are colored blue, and the residues in white have not been tested. Back, top, and front views of the AraC-DNA complex are shown.

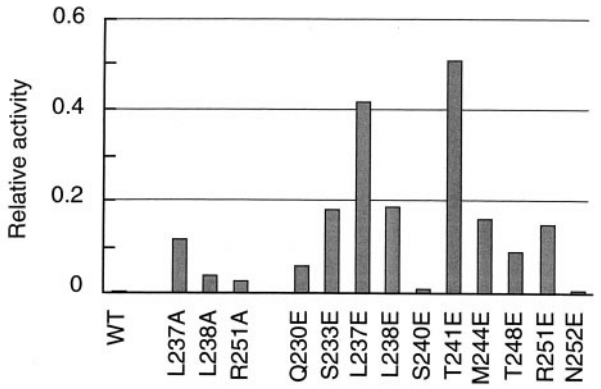


Figure 5. The relative activities of the p_{BAD} promoter in response to AraC in the absence of arabinose and carrying the indicated mutation. The activity was measured by the level of green fluorescence protein in the cell. In the presence of arabinose, wild-type AraC protein has a value of 1.0.

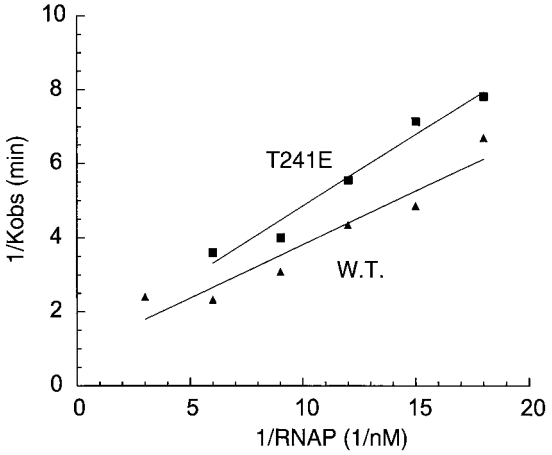


Figure 6. A plot of the inverse of the rate of open complex formation on the *ara P*3 promoter as a function of the inverse of the RNA polymerase concentration. The intercept with the $1/k_{obs}$ axis gives the value of k_2 and the ratio of the slope to the value of k_2 gives K_d.

polymerase. These values are nearly the same as those determined by Zhang *et al*.[24]

Discussion

In the work presented here, we searched for and found constitutive mutations in AraC protein that result from mutations in its DNA-binding domain. This was done by rationally altering the three consecutive leucine residues in the DNA-binding domain that are located at positions 237-239, and then by systematic scanning much of the surface of the domain with alanine and glutamic acid. A total of 20 constitutive mutations at ten different locations were found. The constitutive mutations, including those from randomizing the three leucine residues, those from scanning and the repression negative mutation M244T isolated by Reed,[6] form a contiguous path on the DNA-distal face of the DNA-binding domain. This is as expected if these residues are contacted by the N-terminal arm of AraC.

Two hydrophobic residues (L237 and L238) and one positively charged residue (R251), when substituted with alanine, made AraC constitutive. Possibly these three residues interact directly with the two critical hydrophobic residues L9, and L10 in the arm and with either of the only charged residues in the arm, the negatively charged E3 or D7. Such interactions would also be consistent with the finding that introduction of either hydrophobic or positively charged amino acid residues on the surface of the DNA-binding domain can strengthen its interactions with the arm.[5,6] The N-terminal arm appears to interact with the DNA-binding domain in a way reminiscent of a1/α2 complex in yeast, where the tail of α2 binds to a1 on its back face, distal to the DNA-binding site, stabilized primarily by hydrophobic interactions between three leucines in the tail and a solvent-exposed hydrophobic patch on a1.[26]

Above we concluded that the residues of the DNA-binding domain that when altered can generate constitutive behavior of AraC are the sites of an interaction between the DNA-binding domain and the N-terminal arm of AraC. In principle, mutations generating constitutivity could also result from other effects. One of the more probable alternatives is that the mutations increase the affinity of an interaction between the DNA-binding domain of AraC and RNA polymerase. We already know that such an interaction occurs,[24] but the residues of the AraC DNA-binding domain required for this interaction are unknown. Possibly then, our new constitutive mutations merely strengthen this interaction. Therefore, it was necessary to determine whether any of the mutations alter the apparent binding, K_d, or isomerization rate, k_2, of AraC stimulated transcription initiation at p_{BAD}. We tested two of the constitutive mutants and found their binding and isomerization rates to be similar to those of wild-type AraC. We infer that the rest of the constitutive mutations behave similarly.

We did not scan the entire surface of AraC for several reasons. Because we were substituting alanine and glutamic acid residues, it made no sense to scan residues that in wild-type AraC are already alanine or glutamic acid residues. Also, we choose not to scan residues that lie far from those that do yield constitutive mutations. We consider it unlikely, but not impossible, that their alteration would yield constitutive mutations. Hence, our

conclusion that the constitutive mutations form a localized patch on the surface of the DNA-binding domain has not been exhaustively tested. Also, it was not possible for us to model the final 13 amino acid residues of the DNA-binding domain of AraC because they are not highly homologous to MarA or Rob. We think it is unlikely that the N-terminal arm interacts with these residues, however, because deleting the final 13 amino acid residues from AraC yields a protein that regulates p_{BAD} normally (data not shown).

Throughout we have used the term constitutive to describe the mutations being studied here. These mutations resulted in a substantial increase in the basal activity of p_{BAD}, that is, its activity in the absence of arabinose. Few of the constitutive mutations stimulated p_{BAD} to levels near those induced by AraC protein in the presence of arabinose however, suggesting that few of the mutations entirely abolished the interaction between the N-terminal arm and the DNA-binding domain. All the constitutive mutants could be fully induced by the addition of arabinose though. This is to be expected by the light switch mechanism, since in the mutants, the strength of the dimerization domain-arabinose interaction with the N-terminal arm should be unaltered, and hence arabinose should transfer the arm from the DNA-binding domain to the dimerization domain to the same or greater extent as in wild-type AraC.

As mentioned in the Introduction, alanine and glutamic acid scans are complementary in their sensitivities and selectivities. Fortunately, due to the structure of the genetic code, separate oligonucleotides do not have to be synthesized for the introduction of either of these and only these amino acids at a position. As a consequence, after *in vitro* mutagenesis at a position, the absence of a mutant phenotype means that neither alanine nor glutamic acid residue produces the mutant effect, and sequencing needs to be done only on candidates that do show a mutant phenotype, and then only to determine whether the mutant phenotype is the result of valine or glutamic acid residue. Thus, alanine-glutamic acid scanning provides a powerful yet convenient tool in probing the protein-protein interactions. Glutamic acid scanning was especially suitable in identifying binding sites for the N-terminal arm of AraC as the glutamic acid substitution, in addition to creating a steric hindrance, provides extra electrostatic repulsion to the two negatively charged amino acid residues in the N-terminal arm of AraC.

In summary, we have combined genetic, biochemical and homology-modeling analyses to identify sites on the AraC protein DNA-binding domain that when altered generate constitutivity in AraC. These are likely to be residues that, in the absence of arabinose, are contacted by the N-terminal arm of AraC protein when the protein represses p_{BAD} by forming a DNA loop.

Materials and Methods

General methods, plasmids and strains

Arabinose isomerase was assayed as described.[27] Reproducibility of this assay was approximately 20%. The DNA migration retardation assay was performed as described.[28] Cells were grown in M10 minimal salts medium plus 0.4% (v/v) glycerol, 10 μg/ml thiamine, 0.02% (w/v) L-leucine, and 0.2% (w/v) casamino acids.[27]

Plasmid pWR03 encoding AraC protein[6] was used to randomize leucine residues 237-239. Alanine and glutamic acid scanning were carried out on AraC gene in plasmid GFPuv-p_{BAD}, which also encodes green fluorescence protein under control of the *araBAD* promoter.[29]

RNA polymerase was purified from *Escherichia coli* K12 strain RL916 (F$^-$ λ$^-$ *araD139* Δ*(argF-lac)U169 rpsL150 (Str) relA flbB5301 ptsF25 deoC1 rpoC3531(His6) zja::kan*; derived from MC4100[30]), a generous gift from Dr Robert Landick. In this strain, hexahistidine is attached to the C terminus of the RNA polymerase β-subunit, thus providing a convenient source of ''His-tagged'' RNA polymerase. For randomization of the three leucine residues, strain SH321 (F$^-$ Δ*araC-leu1022* Δ*lac74 galK*$^-$ *Str*^{r} *thi1*[4]) was used. For alanine and glutamic acid scanning, the XL1-Blue cells (*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′ proAB lacI*^{q}*ZΔM15 Tn10(Tet*^{r}*)]*; Strategene, La Jolla, CA) and strain RE5 (Δ*araC-leu1022, araB*$^+$*A*$^+$*D*$^+$*,* Δ*lac74, galK*$^-$ *str*^{r}*,* (λ *ara* I_1-I_2-*lacZ)*[31]) were used.

Site-directed mutagenesis

Site-directed mutagenesis was carried out using Stratagene QuickChange™ Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA). In the experiments of randomizing the three leucine residues, the mutated plasmids were electroporated directly into SH321 cells.

Alanine and glutamic acid scanning

The scanning was performed using site-directed mutagenesis. The residues scanned with alanine were Q184, Y185, S187, H189, D192, S193, S199, Q202, H203, Q217, Q218, D229, Q230, S233, Q234, K236, L237, L238, S240, T241, T242, R243, M244, T248, R251, N252, D256, and D257. Except for D192, D229, D256, and D257, the same residues were scanned with glutamic acid. The oligonucleotide primers were purchased from Integrated DNA Technologies. After transformation, the candidate colonies grown on YT medium lacking arabinose were screened under UV light. Since alanine and glutamic acid share two bases in their genetic codes, both scannings were carried out simultaneously using equal amounts of oligonucleiotides coding for the two amino acid residues at the mutation site. When two residues to be scanned were neighbors, the number of oligo primers was further reduced by making the top primer mutagenic for one of the residues, and the bottom primer mutagenic for the other residue. Usually the pair of primers was then offset from each other by three bases.

Green fluorescent protein assay

Cells grown in minimal medium to A_{600} of ~0.8 were washed once and then diluted to final A_{600} of ~0.02 with

double-distilled water. The fluorescence per cell was measured by exciting the cell suspension at 395 nm and the emission was measured at 508 nm and scattering at 395 nm. Before a day's measurements, the sensitivities of the two detectors are adjusted at 395 nm so that the detector that will later be used for the fluorescence emission at 508 nm is 1.5 times as sensitive as the other. The activity of the p_{BAD} promoter was then taken as the specific activity of green fluorescence protein, Signal at 508 nm/Signal at 395 nm. The reproducibility of these measurements was within 5%.

Kinetics measurement of AraC-stimulated open-complex formation

The kinetics of the AraC-RNA polymerase-promoter open complex formation was measured by DNA migration retardation assay as described.[24] The p_{BAD} promoter used was based on $P3$-I_2-I_1.[23] The DNA fragment containing I_2-I_1 and the RNA polymerase binding sites was generated and radio-labeled by the polymerase chain reaction from the template (AGCGGATCCTTCC TGAAGATTAGCAATTTTATCCATAGATCCTGGTACC GAATTCATGGATCCTACCTGACGCTTTTTATCGGAG CT, with I_2 and I_1 underlined) using two primers: ACTTTGCTCCAGATTAGCGGATCCTTCCTG and TAG AAACAGTAGAGAGCTCCGATAAAAAGC. AraC protein was purified as described.[32] The RNA polymerase was purified as described.[33]

Acknowledgments

We thank Ula Gryczynski for assistance, and NIH Grant 18277 for support.

References

1. Dunn, T. M., Hahn, S., Ogden, S. & Schleif, R. F. (1984). An operator at −280 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.
2. Martin, K., Huo, L. & Schleif, R. F. (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.
3. Lobell, R. B. & Schleif, R. F. (1990). DNA looping and unlooping by AraC protein. *Science,* **250**, 528-532.
4. Hahn, S., Dunn, T. & Schleif, R. F. (1984). Upstream repression and CRP stimulation of the *Escherichia coli* L-arabinose operon. *J. Mol. Biol.* **180**, 61-72.
5. Saviola, B., Seabold, R. & Schleif, R. F. (1998). Arm-domain interactions in AraC. *J. Mol. Biol.* **278**, 539-548.
6. Reed, W. L. & Schleif, R. F. (1999). Hemiplegic mutations in AraC protein. *J. Mol. Biol.* **294**, 417-425.
7. Harmer, T., Wu, M. & Schleif, R. (2001). The role of rigidity in DNA looping-unlooping by AraC. *Proc. Natl Acad. Sci. USA,* **98**, 427-431.
8. Wallace, R. (1982). AvaC and its constitutive mutations in *Escherchia coli* BIR: nucleotide sequences and comparison of altered activities on transcription *in vitro*. *PhD thesis*, University of California, Santa Barbara, USA, CA.
9. Cunningham, B. C. & Wells, J. A. (1989). High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science,* **244**, 1081-1085.
10. Wells, J. A. (1991). Systematic mutational analysis of protein-protein interfaces. *Methods Enzymol.* **202**, 390-411.
11. Clackson, T. & Wells, J. A. (1995). A hot spot of binding energy in a hormone-receptor interface. *Science,* **267**, 383-386.
12. Bogan, A. A. & Thorn, K. S. (1998). Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **280**, 1-9.
13. Dunitz, J. D. (1995). Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chem. Biol.* **2**, 709-712.
14. Rhee, S., Martin, R. G., Rosner, J. L. & Davies, D. R. (1998). A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. *Proc. Natl Acad. Sci. USA,* **95**, 10413-10418.
15. Kwon, H. J., Bennik, M. H., Demple, B. & Ellenberger, T. (2000). Crystal structure of the *Escherichia coli* Rob transcription factor in complex with DNA. *Nature Struct. Biol.* **7**, 424-430.
16. Soisson, S. M., MacDougall-Shackleton, B., Schleif, R. F. & Wolberger, C. (1997). Structural basis for ligand-regulated oligomerization of AraC. *Science,* **276**, 421-425.
17. McNeil, J. B. & Smith, M. (1985). *Saccharomyces cerevisiae* CYC1 mRNA 5′-end positioning: analysis by *in vitro* mutagenesis, using synthetic duplexes with random mismatch base-pairs. *Mol. Cell. Biol.* **5**, 3545-3551.
18. 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.
19. Beverin, S., Sheppard, D. E. & Park, S. S. (1971). D-Fucose as a gratuitous inducer of the L-arabinose operon in strains of *Escherichia coli* B-r mutant in gene araC. *J. Bacteriol.* **107**, 79-86.
20. Greenblatt, J. & Schleif, R. F. (1971). Arabinose C protein: regulation of the arabinose operon *in vitro*. *Nature New Biol.* **233**, 166-170.
21. Nathanson, N. M. & Schleif, R. F. (1973). Novel mutation to dominant fucose resistance in the L-arabinose operon of *Escherichia coli*. *J. Bacteriol.* **115**, 711-713.
22. Gallegos, M. T., Schleif, R. F., Bairoch, A., Hofmann, K. & Ramos, J. L. (1997). Arac/XylS family of transcriptional regulators. *Microbiol. Mol. Biol. Rev.* **61**, 393-410.
23. Reeder, T. & Schleif, R. F. (1993). AraC protein can activate transcription from only one position and when pointed in only one direction. *J. Mol. Biol.* **231**, 205-218.
24. Zhang, X., Reeder, T. & Schleif, R. F. (1996). Transcription activation parameters at *ara* p_{BAD}. *J. Mol. Biol.* **258**, 14-24.
25. Zhang, X. & Schleif, R. F. (1998). CAP mutations affecting activity of the *araBAD* promoter. *J. Bacteriol.* **180**, 195-200.
26. Li, T., Stark, M. R., Johnson, A. D. & Wolberger, C. (1995). Crystal structure of the MATa1/MATα2 homeodomain heterodimer bound to DNA. *Science,* **270**, 262-269.

27. Schleif, R. F. & Wensink, P. C. (1981). *Practical Methods in Molecular Biology*, Springer-Verlag, New York.
28. Hendrickson, W. & Schleif, R. F. (1984). Regulation of the *Escherichia coli* L-arabinose operon studied by gel electrophoresis DNA-binding assay. *J. Mol. Biol.* **178**, 611-628.
29. Cariello, N. F., Narayanan, S., Kwanyuen, P., Muth, H. & Casey, W. M. (1998). A novel bacterial reversion and forward mutation assay based on green fluorescent protein. *Mutat. Res.* **414**, 95-105.
30. Casadaban, M. J. (1976). Transposition and fusion of the lac genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. *J. Mol. Biol.* **104**, 541-555.
31. Eustance, R. J. & Schleif, R. F. (1996). The linker region of AraC protein. *J. Bacteriol.* **178**, 7025-7030.
32. Schleif, R. F. & Favreau, M. A. (1982). Hyperproduction of araC protein from *Escherichia coli*. *Biochemistry,* **21**, 778-782.
33. Kashlev, M., Martin, E., Polyakov, A., Severinov, K., Nikiforov, V. & Goldfarb, A. (1993). Histidine-tagged RNA polymerase: dissection of the transcription cycle using immobilized enzyme. *Gene,* **130**, 9-14.

Edited by J. A. Wells

(Received 15 November 2000; received in revised form 2 February 2001; accepted 2 February 2001)