PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

A new and unexpected domain–domain interaction in the AraC protein

Stephanie Dirla Cole and Robert Schleif*

Department of Biology, Johns Hopkins University, Baltimore, Maryland 21218

ABSTRACT

An interaction between the dimerization domains and DNA binding domains of the dimeric AraC protein has previously been shown to facilitate repression of the *Escherichia coli araBAD* operon by AraC in the absence of arabinose. A new interaction between the domains of AraC in the presence of arabinose is reported here, the regulatory consequences of which are unknown. Evidence for the interaction is the following: the dissociation rate of arabinose-bound AraC from half-site DNA is considerably faster than that of free DNA binding domain, and the affinity of the dimerization domains for arabinose is increased when half-site DNA is bound. In addition, an increase in the fluorescence intensity of tryptophan residues located in the arabinose-bound dimerization domain is observed upon binding of half-site DNA to the DNA binding domains. Direct physical evidence of the new domain–domain interaction is demonstrated by chemical crosslinking and NMR experiments.

Proteins 2012; 80:1465–1475.

Key words: regulation; interdomain interaction; arabinose; affinity; activation.

INTRODUCTION

Changes in interdomain interactions occur in concert with changes in the activities of a number of bacterial regulatory proteins, including RfaH, FimH, LacI, the two component regulatory proteins such as DrrB, PrrA, and NarL, and presumably many others.[1–6] The *Escherichia coli* AraC protein, the regulatory protein of the arabinose operon, is another example in which interdomain interactions are known to play a central role.[7–10] Extensive physiological, genetic, biochemical, and biophysical studies have long been focused on this protein and much is now known about the regulatory and physical properties and basis for the protein's actions.[8,11–14] Nonetheless, much remains to be learned before the *de novo* design and construction of gene regulatory proteins is possible.

The AraC protein represses and activates transcription of the *araBAD* operon in the absence and presence of arabinose.[11,15,16] Much data support the conclusion that, in the absence of arabinose, interactions between the dimerization domains and DNA binding domains aid binding of AraC to the well-separated DNA half-sites I_1 and O_2 by constraining the DNA binding domains in suitable positions and orientations.[7–10,17] Binding to these well-separated half-sites forms a transcriptionally repressive DNA loop [Fig. 1(A)].[12,18–20] At the same time, these constraints hinder the binding of AraC to the adjacent direct repeat half-sites I_1 and I_2,[20] binding to which is required for induction. Much data also implies that the binding of arabinose to the dimerization domains releases the DNA binding domains, and, as a result, the DNA binding domains are now more free to bind to the adjacent I_1 and I_2 DNA half-sites[10,18,21,22] and induce transcription [Fig. 1(B)].[23,24]

Until recently nothing has suggested that the DNA binding domains and dimerization domains of AraC interact in the presence of arabinose, and, in fact, models without such interactions[17] can be constructed that well replicate the regulatory behavior of AraC. As a result, no such interactions have hitherto been postulated. Several additional lines of reasoning have also led to the deduction of the absence of significant domain–domain interactions in the presence of arabinose. These include the fact that two DNA binding domains connected by a flexible peptide linker can fully induce p_{BAD},[10] suggesting that the DNA binding domains of AraC need not interact with the dimerization domains to induce transcription. Another comes from estimating the DNA binding affinity of the dimeric AraC protein in the presence of arabinose based on the DNA binding affinity of isolated DNA binding domain.[25] The estimation of the affinity rather closely agrees with the experimental data if the DNA binding domains are

Additional Supporting Information may be found in the online version of this article.
Abbreviations: DTNB, 5, 5′dithiobis-2-nitrobenzoic acid; EDTA, ethylenediaminetetraacetic acid; PEAS, *N*-((2-pyridyldithio)ethyl)-4-azidosalicylamide; TCEP, tris(2-carboxyethyl)phosphine.

Grant sponsor: NSF grant; Grant number: MCB-1021031
*Correspondence to: Robert Schleif, Biology Department, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218. E-mail: schleif@jhu.edu.
Received 18 October 2011; Revised 12 January 2012; Accepted 19 January 2012
Published online 6 February 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24044

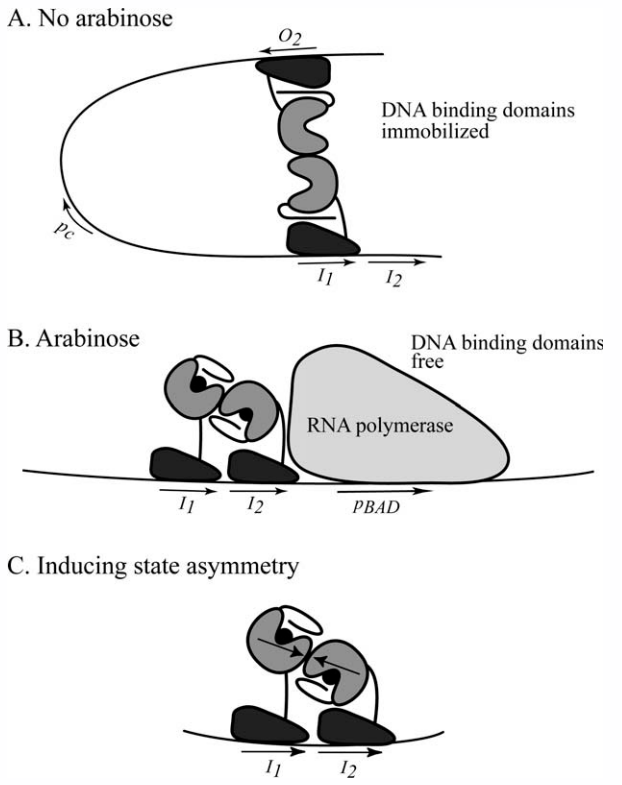

Figure 1

Schematic for the regulation of p_{BAD} by AraC (**A**) in the absence of arabinose and (**B**) in the presence of arabinose. In the absence of arabinose, an interaction between the dimerization domains and DNA binding domains of AraC facilitates DNA looping (A), while the domains have previously been thought to be free of one another in the presence of arabinose (B). Asymmetry between the dimerization domains and DNA binding domains when AraC is bound to $I_1 - I_2$ DNA (**C**) makes is impossible for the dimerization domain and DNA binding domain of each subunit to interact similarly in the inducing state.

assumed not to significantly interact with the dimerization domains. An additional reason to expect that an interaction does not occur between the dimerization domains and DNA binding domains of AraC in the presence of arabinose is that the two-fold rotational symmetry of the dimerization domains of AraC is manifestly different from the translational symmetry of the DNA binding domains when the protein is bound to the direct repeat I_1 and I_2 DNA half-sites at the p_{BAD} promoter.[22,26] Only when the DNA binding domains do not interact similarly with the dimerization domains can such domain–domain interactions occur when the AraC is bound to $I_1 - I_2$ [Fig. 1(C)]. The simplest assumption from these results and arguments is that no significant interaction occurs between the domains of AraC in the presence of arabinose.

In contrast to the assumption that no interaction occurs between the dimerization domains and DNA binding domains of AraC in the presence of arabinose, a newly developed but not fully tested method for the measurement of weak protein–protein interactions detected such an interaction.[27] It is unclear, however, that this interaction is present in the full-length AraC protein because it was detected between free domains of AraC that could have interacted using surfaces that are not normally accessible to the domains due to constraints of their short interdomain linker. As a result, the interactions detected by Frato *et al.* may not be relevant to the full-length protein. Here, we report new experimental data that demonstrate the existence of a significant interaction between the dimerization domains and DNA binding domains of full-length AraC in the presence of arabinose. While the strength of this interaction is sufficient to be detectable by the methods described below, its strength is compatible with the general model described above for the regulation of the DNA binding properties of AraC.

MATERIALS AND METHODS

Plasmids and strains

For overexpression, the wild type *araC* gene, as well as DNA encoding the DNA binding domain of AraC (AraC residues 175–292), was previously cloned between the *NcoI* and *SacI* sites in the multiple cloning region of the pET24d vector (kan^r).[28] DNA encoding the dimerization domain of AraC (residues 1–182) followed by a C-terminal Leu-Glu linker and His_6 tag was previously cloned into the pET21b vector (amp^r).[29]

Mutations were introduced into the *araC* gene using QuikChange Site-directed Mutagenesis (Stratagene). Plasmid DNA was isolated from single colonies of DH5α[30] cells using the Wizard Plus Miniprep DNA purification system (Promega), and mutations were confirmed by DNA sequencing. For the chemical crosslinking experiments, AraC with all the following mutations was used: C66S, C205N, C280S, and either C183I or C268R.

The pET24d vector containing wild type or mutant versions of the *araC* gene was transformed into SH321 ($AraC^-$) cells for the arabinose isomerase assay.[20]

Protein purification and isotopic labeling

For all unlabeled proteins, fresh BL21(DE3) transformants containing the appropriate plasmid were grown overnight in 5 mL YT medium plus 100 μg/mL ampicillin or 40 μg/mL kanamycin. These starter cultures, washed with fresh YT medium when ampicillin resistant plasmids were present, were used to inoculate 2 L baffled flasks containing 500 mL YT medium plus antibiotics. Cells were grown to OD_{600nm} of about 0.8, and protein expression was induced by the addition of 0.4 m*M* IPTG. Cells expressing dimerization domain or DNA binding domain were grown at 37°C for three more hours, while cells expressing AraC were grown for at least 5 h at room

temperature. Cells were then harvested by sedimentation at 6000*g* for 10 min and stored at −80°C.

Preparation of the dimerization domain was performed essentially as described.[29] Cells were lysed in an Avestin EmulsiFlex C3 homogenizer in buffer consisting of 15 m*M* Tris–Cl, pH 8.0, 100 m*M* NaCl, 10 m*M* $MgCl_2$, 10 m*M* arabinose, 5% glycerol, and 10 μg/mL DNase. The cell lysate was centrifuged at 10,000*g* for 15 min and the supernatant was gently rocked in the presence of 4 mL Ni-NTA agarose beads (Qiagen) at 4°C. The bead mixture was then centrifuged for 5 min at 1000*g* and the supernatant was discarded. The beads were then washed with buffer containing 15 m*M* Tris–Cl, pH 8.0, 100 m*M* NaCl, 10 m*M* arabinose, and 10 m*M* imidazole until the absorbance of the flow through at 280 nm was less than 0.05. Dimerization domain was eluted at a concentration of roughly 1 mg/mL in 25 mL of buffer containing 15 m*M* Tris–Cl, pH 8.0, 10 m*M* NaCl, 10 m*M* arabinose, and 2*M* imidazole. The His_6 tag was removed by incubation with 5 μg/mL trypsin overnight at 4°C. Trypsin was inhibited by the addition of 5 μg/mL soybean trypsin inhibitor, and then trypsin was removed by anion exchange chromatography on a 5 mL HiTrap Q HP column (Amersham Biosciences). The sample was loaded in 15 m*M* Tris–Cl, pH 8.0, and 10 m*M* arabinose and eluted with roughly 250 m*M* NaCl in a gradient of 0–1*M* NaCl over 40 mL. Fractions were analyzed using SDS-PAGE and found to be greater than 95% pure. The relevant fractions were concentrated using Amicon Ultra-4 10K devices (Millipore).

DNA binding domain was purified essentially as described.[28] Cells were lysed in an Avestin EmulsiFlex C3 homogenizer in buffer consisting of 50 m*M* phosphate, pH 7.0, 500 m*M* NaCl, 1 m*M* ethylenediaminetetraacetic acid (EDTA), 5% glycerol, and 10 μg/mL DNase. The lysate was centrifuged at 10,000*g* for 15 min, and the supernatant was discarded. The DNA binding domain was refolded from inclusion bodies by dissolving the lysate pellet in 7*M* guanidine-HCl, then diluting in half with 20 m*M* phosphate, pH 7.0, 50 m*M* NaCl. This sample was then dialyzed against 1 L of 20 m*M* phosphate, pH 7.0, 50 m*M* NaCl, and 1 m*M* DTT twice for at least 3 h per dialysis. The protein sample was then centrifuged for 15 min at 10,000*g*, and the supernatant was loaded onto a 5 mL HiTrap Heparin HP column and eluted over a 40 mL gradient from 0 to 1*M* NaCl in 20 m*M* phosphate, pH 7.0 buffer. Fractions were analyzed and concentrated as above, and the resulting protein was found to be greater than 95% pure.

Cells containing overexpressed AraC were lysed in buffer containing 20 m*M* phosphate, pH 6.5, 300 m*M* NaCl, 10 m*M* arabinose, 1 m*M* EDTA, 2 m*M* $MgCl_2$, 5% glycerol, 1 m*M* DTT, and 10 μg/mL DNase using an Avestin EmulsiFlex C3 homogenizer. The lysate was centrifuged at 10,000*g* rpm for 15 min, and the supernatant was diluted in half with 20 m*M* phosphate, pH 6.5 and 10 m*M* arabinose. This supernatant was loaded onto a 5 mL HiTrap Heparin HP column and eluted with a 40 mL gradient of 0 to 1*M* NaCl in 20 m*M* phosphate, pH 6.5, 10 m*M* arabinose. AraC typically eluted in approximately 500 m*M* NaCl, and fractions containing AraC were combined and diluted four-fold in 15 m*M* Tris–Cl, pH 7.5, and 10 m*M* arabinose. This sample was then loaded onto a 5 mL HiTrap Q HP column in 15 m*M* Tris–Cl pH 7.5, and 10 m*M* arabinose, and AraC eluted in the flow through. This flow through was then loaded onto a 5 mL HiTrap Heparin HP column and eluted in a 40 mL gradient from 0 to 1*M* NaCl in 15 m*M* Tris–Cl, pH 7.5, and 10 m*M* arabinose. Fractions were analyzed using SDS-PAGE and AraC was found to be greater than 90% pure. The protein was concentrated as above.

DNA binding domain was ^{15}N-labeled essentially as described,[31] except that cells were grown in minimal medium before induction and labeling. Labeled protein was otherwise expressed and purified as described above.

Comparisons of dissociation kinetics from I_1 half-site DNA

To observe the dissociation rate of AraC or DNA binding domain from I_1 half-site DNA, gel binding assays were performed in which 100 n*M* protein was incubated for 10 min with 10 n*M* Cy5-labeled I_1 DNA in buffer consisting of 10 m*M* Tris-acetate, pH 7.4, 50 m*M* KCl, 10 m*M* arabinose (when present), 1 m*M* EDTA, 5% glycerol, 0.05% NP40, and 1 m*M* tris(2-carboxyethyl)phosphine (TCEP). The sequence for the labeled I_1 DNA is 5′-Cy5-GCTAGCATTTTTATCCATACCTCCCTC-3′, where the underlined bases constitute the I_1 half-site. Before loading the 10% acrylamide gel, the sample was diluted 5-fold in the above buffer in the absence of salt, and unlabeled I_1 competitor DNA was added simultaneously at a final concentration of 500 n*M*. The sequence of unlabeled I_1 DNA is as follows: 5′-CGCCATAGCATTTTTATCCATAAGATC-3′. Samples were then loaded into the gel at various time points and were subjected to electrophoresis for 20 min at 80 V. Fluorescent bands were imaged using a Typhoon 9410 Variable Mode imager and quantified using ImageQuant TL v2003.03.

Equilibrium I_1 DNA binding measurements

For I_1 binding measurements, 10 n*M* 5′ Cy5-labeled I_1 DNA was placed in a 1-cm glass cuvette with buffer consisting of 15 m*M* Tris–Cl, pH 7.5, 100 m*M* KCl, 10 m*M* arabinose (when present), 1 m*M* EDTA, 5% glycerol, and 1 m*M* TCEP. AraC or DNA binding domain protein, which had been dialyzed against 15 m*M* Tris–Cl, pH 7.5, 250 m*M* NaCl, and 10 m*M* arabinose (when present), was titrated into the cuvette and anisotropy was monitored by excitation at 640 nm and emission at 658 nm, collecting 1 point per second for 30 s using a T-configuration Photon International fluorimeter with a 75-W xenon arc lamp

light and a temperature-controlled cuvette holder. The I_1 DNA sequence used in these experiments was the same as described above. Because the DNA concentration was on the same order of magnitude as the K_d, binding was fit to the following equation, which accounts for the reduction in free AraC due to the non-negligible amount of AraC–DNA complex, and where F_b is the fraction bound, x_o is the initial fluorescent signal, Δx is the change in fluorescent signal on binding, P_{total} is the total concentration of protein, and D_{total} is the total concentration of DNA:

$$F_b = x_o + \Delta x * \frac{([P_{total}] + [D_{total}] + K_d) - \sqrt{([P_{total}] + [D_{total}] + K_d)^2 - 4[P_{total}][D_{total}]}}{2[D_{total}]}$$

The activity of the proteins used in I_1 binding measurements was determined by titration in gel binding assays by using 500 n*M* Cy5-labeled I_1 DNA. Protein and DNA were incubated together for at least 10 min at room temperature, and then loaded directly into a gel essentially as described. Fluorescent bands were imaged and analyzed as described above.

Arabinose binding measurements

Arabinose binding was assayed fluorometrically on a T-configuration Photon International fluorimeter with a 75-W xenon arc lamp light and a temperature-controlled cuvette holder.

Arabinose was removed from the AraC protein by dialysis against 15 m*M* Tris–Cl, pH 7.5, and 250 m*M* NaCl. Arabinose binding was measured using 500 n*M* AraC protein in a 1 cm quartz cuvette in 15 m*M* Tris–Cl, pH 7.5, 100 m*M* KCl, 1 m*M* EDTA, 5% glycerol, 1 m*M* TCEP, and 500 n*M* unlabeled I_1 half-site DNA, when present (see above for sequence). Tryptophan fluorescence was monitored as arabinose was titrated into the sample by exciting at 295 nm and integrating the emission from 320 to 370 nm. The emission spectra were corrected for dilution and then the average emission wavelength and integrated fluorescence intensity were calculated. The average emission wavelength versus arabinose concentration was plotted and the data were fit to a simple binding isotherm to determine the K_d. For the measurement of the absolute fluorescence intensity in the presence of I_1-I_1 DNA, the DNA sequence was 5′-GCCA TAGCATTTTTATCCATAAGATTAGCATTTTTATCCATAC CTC-3′.

Chemical crosslinking

The AraC variant containing C183 possessed the mutations C66S, C205N, C268R, C280S, and the variant containing C268 possessed the mutations C66S, C183I, C205N, C280S. The AraC variants at 40 µ*M* concentration and a 140 µL volume were incubated in 96-well plates with 1 m*M* *N*-[(2-pyridyldithio)ethyl]-4-azidosalicylamide (PEAS, Invitrogen) in 20 m*M* phosphate, pH 6.5, 250 m*M* NaCl, and 10 m*M* arabinose in the dark at room temperature overnight. Then, I_1 DNA was added to the reaction mixtures to a final concentration of 40 µ*M* and a final volume of 170 µL. These protein and DNA concentrations are over 1000-fold higher than the equilibrium dissociation constants for AraC binding to I_1 DNA in the presence and absence of arabinose (reported here) and thus all of the protein is bound to the DNA under these conditions. The crosslinking experiments were all done in the presence of DNA because PEAS-AraC in the absence of arabinose and DNA is insoluble. The reaction mixtures were divided in half and dialyzed, using dialysis tubing with a molecular weigh cutoff of 10,000 Da, against 500 mL of 20 m*M* phosphate, pH 6.5, 250 m*M* NaCl, and 10 m*M* arabinose, when present at 4°C for 4 h to remove excess, unreacted PEAS. Finally, the reaction mixtures were exposed to UV light from a Mineralight C51 short wavelength light source (Ultra Violet Prod) by placing them approximately three inches from the light source for 30 min at room temperature. Samples were analyzed by non-reducing SDS-PAGE on 12% acrylamide gels.

The 5,5′dithiobis-2-nitrobenzoic acid (DTNB)[32] reactivity of C183 and C268 was measured in a 1-cm path length plastic cuvette by combining 10 µ*M* of AraC containing C183 or C268 as the only cysteine and 20 µ*M* DTNB to a total volume of 1 mL in buffer containing 20 m*M* phosphate, pH 6.5, 250 m*M* NaCl, and 10 m*M* L-arabinose. The absorbance at 412 nm was measured every second for up to 2 h.

Arabinose isomerase assay

The arabinose isomerase assay was performed as described on cells that had been growing exponentially in M10 minimal medium supplemented with 40 µg/mL kanamycin, 0.4% (v/v) glycerol, 10 µg/mL thiamine, 20 10 µg/mL L-leucine, 1% (w/v) casamino acids, and 0.2% (w/v) L-arabinose, when present.[33]

NMR experiments

Labeled DNA binding domain and unlabeled dimerization domain were dialyzed into buffer consisting of 16.5 m*M* Tris–Cl, pH 7.5, 165 m*M* NaCl, 11 m*M* arabinose, and 1.1 m*M* EDTA. Ten percent D_2O by volume was

added to each sample, along with soybean trypsin inhibitor that had been dissolved in the dialysis buffer to inhibit any trypsin that carried through from the preparation of the dimerization domain. The final experimental buffer contained 15 m*M* Tris–Cl, pH 7.5, 150 m*M* NaCl, 10 m*M* arabinose, 1 m*M* EDTA, 0.1 mg/mL soybean trypsin inhibitor, and 10% D_2O. In each experiment, the ^{15}N-labeled DNA binding domain was kept at a concentration of 60 μ*M*, and the concentration of the dimerization domain varied from 0 to 1.23 m*M*.

^{1}H–^{15}N HSQC spectra were acquired on a Bruker Avance spectrometer at the Biomolecular NMR center at Johns Hopkins University operating at a ^{1}H Larmor frequency of 600 MHz and equipped with a TCI cryoprobe with *z*-axis pulse field gradients. Data were processed using NMR Pipe[34] and analyzed using NMR View.[35]

RESULTS AND DISCUSSION

Arabinose-bound AraC dissociates from half-site DNA faster than free DNA binding domain

As described above, a variety of *in vitro* DNA binding experiments are compatible with the idea that the DNA binding domains of AraC do not significantly interact with the dimerization domains in the presence of arabinose.[17] If this were true, then the dissociation rate of arabinose-bound full-length AraC from a single I_1 half-site, that is, dissociation that involves only a single DNA binding domain, should be the same as the dissociation rate of free DNA binding domain for the same DNA. If, however, a significant difference exists between the I_1 half-site DNA dissociation rate of arabinose-bound AraC and free DNA binding domain, then the DNA binding domains in the two situations differ. This must be the result of an interaction between the dimerization domain and DNA binding domain that occurs in the presence of arabinose.

Experimentally, we found that the dissociation of full-length AraC from I_1 DNA is greater than 150 times faster than the dissociation of free DNA binding domain (Fig. 2). Despite the large differences in the rates of dissociation from I_1 DNA, we find that arabinose-bound AraC and free DNA binding domain bind to such DNA with similar affinity, with equilibrium dissociation constants of 10.5 ± 3.8 n*M* and 10.7 ± 4.3 n*M*, respectively, as measured by fluorescence anisotropy (data not shown). Because $K_d = k_{off}/k_{on}$, the large differences in k_{off} must mean that k_{on} is approximately 150-fold faster for arabinose-bound AraC relative to that of the free DNA binding domain. The surprisingly large apparent differences in the dissociation rates for these proteins from I_1 DNA does not appear to be due to technical problems associated with the gel binding assay because, under similar conditions, this assay reveals that arabinose

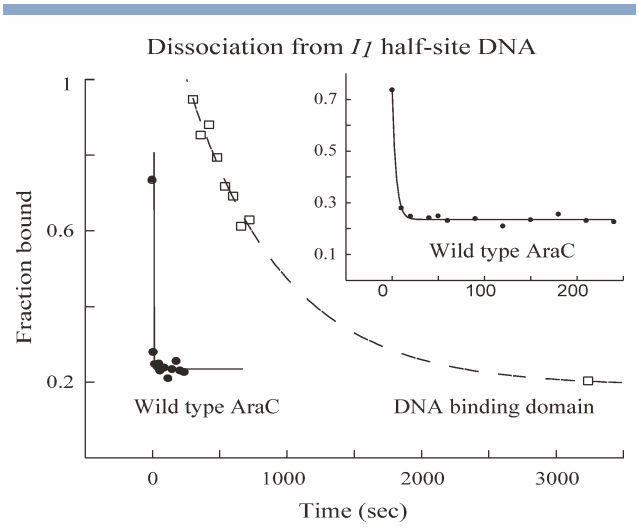

Figure 2

Dissociation of arabinose-bound AraC (inset; filled circles) and free DNA binding domain (unfilled squares) from I_1 half-site DNA. Arabinose-bound AraC dissociates from I_1 half-site DNA approximately 150-fold times faster than the DNA binding domain.

decreases the dissociation rate from $I_1 - I_1$ full-site DNA for full-length AraC (data not shown), as expected and consistent with previous findings.[22]

The large differences in the dissociation rates of full-length AraC in the presence of arabinose and free DNA binding domain appear to be a consequence of interactions in full-length AraC in the presence of arabinose between the DNA binding domains and the dimerization domains. It seems likely that the interactions alter the conformation of the DNA binding domains and this is responsible for the rate differences. While it is conceivable that the mere tethering of the DNA binding domain to the dimerization domain might slightly alter the dissociation rate of the DNA binding domain from half-site DNA, it is hard to imagine how this effect could be as large as 150-fold. It therefore seems most likely that the fast dissociation from half-site DNA is caused by an interaction between the DNA binding domain and dimerization domain of AraC in the presence of arabinose.

The arabinose binding affinity is altered when AraC is bound to I_1 DNA

As described in the introduction, it has seemed probable that the DNA binding domains and dimerization domains of AraC interact in the absence of arabinose but not in its presence. Thus, by the same reasoning, in constitutive mutants of AraC, that is, mutants that do not require the presence of arabinose to be in the inducing state,[11] the DNA binding domains would likely not interact with the dimerization domains. In this section we show that, contrary to expectations, but consistent with the findings of the previous section, an appreciable

Table I
Dissociation Constants for Arabinose Binding to AraC in the Presence and Absence of I_1 Half-site DNA

	K_d *in the absence of* I_1 *half-site DNA (m*M*)*	K_d *in the presence of* I_1 *half-site DNA (m*M*)*
L9H AraC	0.46 ± 0.043	0.22 ± 0.014
WT AraC	0.41 ± 0.075	0.15 ± 0.0093

interaction occurs between the dimerization domains and DNA binding domains of a constitutive AraC mutant. This interaction is revealed by the change in the arabinose binding affinity of the dimerization domain when half-site DNA is bound to the DNA binding domain.

We find that binding I_1 half-site DNA to a constitutive AraC mutant, L9H AraC,[36] increases the affinity by which the dimerization domain binds arabinose by approximately two-fold (Table I). This indicates that instead of the domains being free from one another in a constitutive mutant, as has been assumed, the domains of such a mutant interact. Because constitutive AraC mutants are in the inducing state even in the absence of arabinose, the interactions that occur between the dimerization domains and DNA binding domains of such a mutant occur while the protein is in its inducing state. Similarly, we find that the binding of I_1 half-site DNA to the DNA binding domain of wild type AraC also increases the arabinose binding affinity of its dimerization domain by a comparable amount. Thus, the effect of half-site DNA binding on arabinose binding in wild type AraC protein can be used to suggest that the interaction that occurs in the presence of arabinose in L9H AraC also occurs in wild type AraC and is not a result of the mutation itself. The thermodynamic cycle that describes the four relevant binding events of this situation also requires that arabinose strengthen the binding of AraC to I_1 half-site DNA. This is because the products of the equilibrium constants around both sides of such a thermodynamic cycle must be equal to one another. As such, we find that arabinose strengthens the binding of wild type AraC to I_1 half-site DNA as measured by fluorescence anisotropy (data not shown), with equilibrium dissociation constants of 10.5 ± 3.8 n*M* and 42.8 ± 4.3 n*M* in the presence and absence of arabinose, respectively.

Fluorescence changes supporting a dimerization domain–DNA binding domain interaction in the presence of arabinose

The previous sections utilized binding data to suggest the existence of an interdomain interaction in AraC in the presence of arabinose. In this section, we provide evidence from fluorescence measurements for such an interaction. The rationale is as follows: if, in the presence of arabinose, the DNA binding domains of AraC are not bound to the dimerization domains, then the absolute change in the tryptophan fluorescence on binding to I_1 half-site DNA should be close to the change in the fluorescence of DNA binding domain alone on binding to I_1 DNA. That is, in the absence of an interdomain interaction, the change in fluorescence on DNA binding would be expected to arise primarily from the single native tryptophan of the DNA binding domain and not from the five tryptophans of the dimerization domain.

Experimentally, we find that the tryptophan fluorescence of arabinose-bound AraC increases by roughly 10% upon binding to I_1 half-site DNA, whereas the fluorescence of the DNA binding domain alone decreases by approximately 2% when binding to I_1 DNA (Table II). We verified that the increase in the fluorescence change that occurs with full-length AraC derives from changes in the dimerization domain by removing the only tryptophan residue in the DNA binding domain. In this mutant, W226F, the binding of I_1 half-site DNA in the presence of arabinose still produced a 10% change in fluorescence. Because two of the tryptophan residues of the dimerization domain could be altered without loss of protein activity, W91 and W95, it was possible to further localize the changes responsible for the fluorescence change to the three remaining tryptophan residues, W104, W107, and W110. As these lie near the surface and close to the interface between dimerization domains, their environments could easily be altered by a DNA binding domain-dimerization domain interaction. It should be mentioned that the domain–domain interaction that we are suggesting that occurs in the presence of arabinose is not identical to the interaction that does occur in the absence of arabinose. The reason is that the binding of half-site DNA to AraC in the absence of arabinose produces virtually no change in the absolute tryptophan fluorescence.

Evidence for the plus arabinose domain–domain interaction using chemical crosslinking

Because the previously described experiments are somewhat indirect, we sought direct evidence for the proposed domain–domain interaction. For this, we used the heterobifunctional photoactivatable chemical cross-

Table II
Relative Changes in Tryptophan Fluorescence Intensity Upon I_1 Half-site DNA Binding

	Fluorescence change in the presence of arabinose	Fluorescence change in the absence of arabinose
WT AraC	11% ± 0.12%	−1.1% ± −0.50%
DNA binding domain	—	−2.1% ± −0.16%
W91F AraC	12% ± 1.2 %	−1.9% ± 0.10%
W95F AraC	10% ± 0.82%	1.1% ± 0.047%
W226F AraC	8.4% ± 0.87%	−2.9% ± −0.66%

Table III
Arabinose Isomerase Levels for Wild Type AraC and AraC Mutants Containing C183 or C268 as the Only Cysteine Residue

	Arabinose isomerase levels in units per cell	
	Presence of arabinose	Absence of arabinose
WT AraC	920 ± 70	1.3 ± 0.90
C183 AraC	550 ± 90	2.4 ± 0.71
C268 AraC	1000 ± 78	3.7 ± 1.1

linker PEAS to directly verify the proximity and restricted rotational freedom of the DNA binding domain of AraC in the presence of arabinose. For these experiments, we used a mutant of AraC in which four of the five native cysteine residues, C66, C183, C205, C268, and C280, were replaced, leaving only a single cysteine residue, either at position 183 or at 268 of the DNA binding domain. Both of these AraC variants behave like wild type AraC *in vivo* (Table III), indicating that the substitutions have not altered the basic regulatory properties of the protein. To detect the interdomain interaction using PEAS, its thiol-reactive group was first allowed to form a disulfide bond with the single remaining solvent-exposed cysteine in the protein that is located in the DNA binding domain. This was followed by photoactivation of the aryl azide component of PEAS to form a second bond somewhere within the protein.

Crosslinked species formed with C183 while AraC was bound to I_1 half-site DNA in the presence and absence of arabinose yielded two slowly-migrating bands on nonreducing SDS-PAGE gels (Fig. 3), whereas no similar species were formed with C268. The finding that PEAS-C183 forms crosslinked species and that PEAS-C268 does not suggests that the DNA binding domain is constrained in position and orientation, with PEAS-C183 in a position where it can interact with the dimerization domain and PEAS-C268 is constrained in a position where it cannot interact. Had the DNA binding domains not been constrained, the patterns of reactivity of PEAS-C183 and PEAS-C268 would have been similar. We also expected to observe more than two discrete bands on the gel if the crosslinking were subject to random orientational movement of the DNA binding domains, although a negative control in which no domain–domain interactions occur is necessary to truly rule out this possibility. As such, these crosslinking results are consistent with a plus arabinose domain–domain interaction but do not definitively prove its existence.

It should be noted that the crosslinking results also indicate that the DNA binding domains are constrained in the absence of arabinose, consistent with earlier findings.[7–10,22] The same crosslinked species were also formed with C183 while AraC was bound to I_1 half-site DNA and $I_1 - I_1$ full-site DNA, in both the presence and absence of arabinose. Because the crosslinking experiments address large-scale aspects of the structure of AraC, the similarity between the crosslinking patterns produced in the presence and absence of arabinose is perhaps not surprising, and merely indicates that the DNA binding domains are kept in close proximity to the dimerization domains under both conditions.

To verify that the crosslinker was working as expected, we demonstrated that the crosslinks were from the cysteine in the DNA binding domain to the dimerization domain by showing that breaking the disulfide bond of the crosslinked species produces a new sulfhydryl that is located on the dimerization domain. This was shown by reacting with fluorescein-maleimide and finding that the dimerization domain had become fluorescent (Supporting Information).

The lack of crosslinking observed for C268 is not due to poor solvent exposure of the cysteine, which would prevent PEAS from attaching to it. In the presence of arabinose and the absence of DNA, we find that DTNB requires 4–5 h to fully react with either C183 or C268 at concentrations of protein and DTNB similar to the concentrations of protein and PEAS used in the crosslinking experiments. For the crosslinking experiments, PEAS was allowed to react overnight with both C183 and C268 in the presence of arabinose and the absence of DNA, and only afterward was DNA added and arabinose removed. These results suggest that the lack of crosslinked species observed with C268 was not due to a failure of the crosslinker to form a disulfide bond with the surface exposed cysteine. Also, the attachment of PEAS to C183 or C268 before the photoactivation step does not unfold the DNA binding domain because we find that these AraC variants bind equally as well to I_1 DNA before or after the binding of PEAS in a gel binding assay (data not shown).

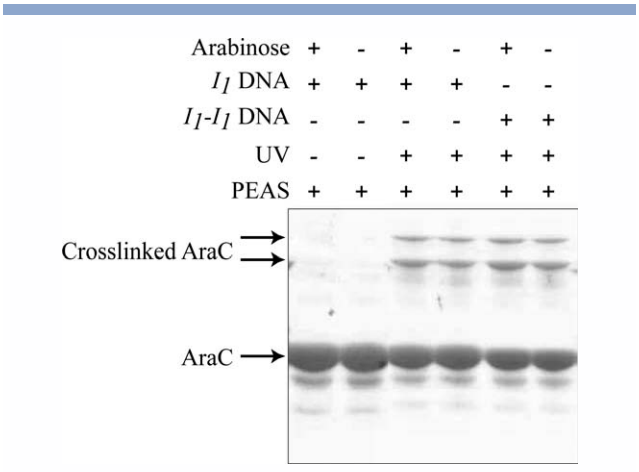

Figure 3
SDS-PAGE gel showing crosslinked species formed by C183-PEAS. The same crosslinked bands appear in the presence and absence of arabinose, and in the presence of I_1 half-site and $I_1 - I_1$ full-site DNA.

The results presented in this section indicate that C183 of at least one of the two DNA binding domains of AraC is in close proximity to the dimerization domain in the presence and absence of arabinose, and presumably it is held there by interactions between the dimerization domain and DNA binding domain. The lack of crosslinking activity of C268 indicates that it does not lie near the interaction interface in the presence or absence of arabinose. Consistent with the crosslinking results presented here is the existence of mutations at C183 and nearby residues, but not at C268, that have been previously identified that cause AraC to repress transcription in both the presence and absence of arabinose.[8] C183 may thus be involved in important domain–domain interactions that regulate the behavior of AraC. An additional finding of the crosslinking experiments is that the same crosslinked species form with C183 when AraC is bound to half-site and full-site DNA, suggesting that a domain–domain interaction that occurs when AraC is bound to arabinose could also occur when AraC is bound to p_{BAD} *in vivo*.

Direct observation of the plus arabinose domain–domain interaction by NMR

The kinetics, affinity, fluorescence, and crosslinking measurements presented above all provide evidence for an interaction between the dimerization domains and DNA binding domains of AraC in the presence of arabinose. NMR ^{1}H–^{15}N-HSQC experiments in the presence of both binding partners have the capability of directly demonstrating the interaction. The resonances observed in such experiments are very sensitive to changes in the chemical environments of their atoms, as occurs for atoms in an interface when two proteins or domains associate or dissociate. An additional virtue of these types of experiments is that if the resonances have been assigned, as is the case for the DNA binding domain of the AraC protein,[28] then the changes can be mapped to specific residues. If the protein complex is large, as unfortunately is also the case for the dimerization domain–DNA binding domain complex, then the intensities of all of the resonances will decrease upon binding due to line broadening. In some cases, however, the strength of the resonances from atoms at the interface decrease more strongly than the others.[37]

The ^{1}H–^{15}N-HSQC NMR technique can provide meaningful signals at the high protein concentrations that are necessary to generate a significant amount of interaction between weakly interacting proteins[38] Such a weak interaction is the case for the dimerization domain and DNA binding domain of AraC. In full-length AraC, the interdomain linker holds the dimerization domains and DNA binding domains at a relative local concentration of roughly 10 m*M*.[27] For the association and dissociation of the domains to be regulated at such high local concentrations, their equilibrium dissociation constant should then be between 0.1 m*M* and 1.0 m*M*.[27]

Consistent with the prediction that the interdomain interaction must be very weak, Frato *et al.* observed an apparent interaction with an equilibrium dissociation constant of roughly 0.3 m*M* in the absence of arabinose, which is close to the strength that is expected in order for the interaction to be regulated, as discussed above.[27] Unexpectedly, a weaker, but measureable interaction was also observed in the presence of arabinose. In the experiments of Frato *et al.*,[27] however, the domains are not connected to one another by their interdomain linker, and therefore it is unclear whether this interaction can occur in normal, full-length AraC. In addition to detecting an interaction between the DNA binding and dimerization domains of AraC, NMR has the capability of precisely locating the interaction interface, whereas the DNA assisted binding method developed by Frato *et al.* only measures the strength of the interaction.[27]

We labeled the DNA binding domain of AraC with ^{15}N and collected HSQC spectra of this domain alone, and in the presence of increasing concentrations of unlabeled dimerization domain. Because the DNA binding domain-dimerization domain complex is large, roughly 50,000 kDa, we expected the bound state to be invisible in the HSQC spectrum as a result of significant line broadening due to increased rotational correlation time. We thus predicted that, on binding, the intensities of most of the resonances would decrease substantially. In the presence of arabinose, we found that the intensities of the resonances of the DNA binding domain decreased with increasing concentration of added dimerization domain [Fig. 4(A)], indicating that the two proteins interact. Because the bound state was invisible in these experiments, we observed no significant changes in the chemical shifts of any of the resonances. A similar decline in intensity on complex formation was observed by Greene *et al.* in their study of the interaction interface of the large complex DnaJ/K, although the DnaJ/K complex was not as large as the AraC DNA binding domain-dimerization complex and as such the authors also observed some changes in the chemical shifts.[39]

At low to moderate concentrations of dimerization domain, some resonances of the AraC DNA binding domain decrease in intensity more than others. At the highest concentrations of dimerization domain, more than half of the DNA binding domain resonances decreased in intensity by 60% to 100% [Fig. 4(A)]. This number of residue changes is typical for large protein–protein complexes, but it makes the comparative analysis of resonance intensities difficult. Thus, we compared the intensity of resonances in the presence of low to moderate concentrations of dimerization domain. We found that the resonances that decrease most in intensity are from most of the residues located on the surface of the DNA binding domain that is opposite the DNA binding surface [Fig. 4(B)]. While we are unable to

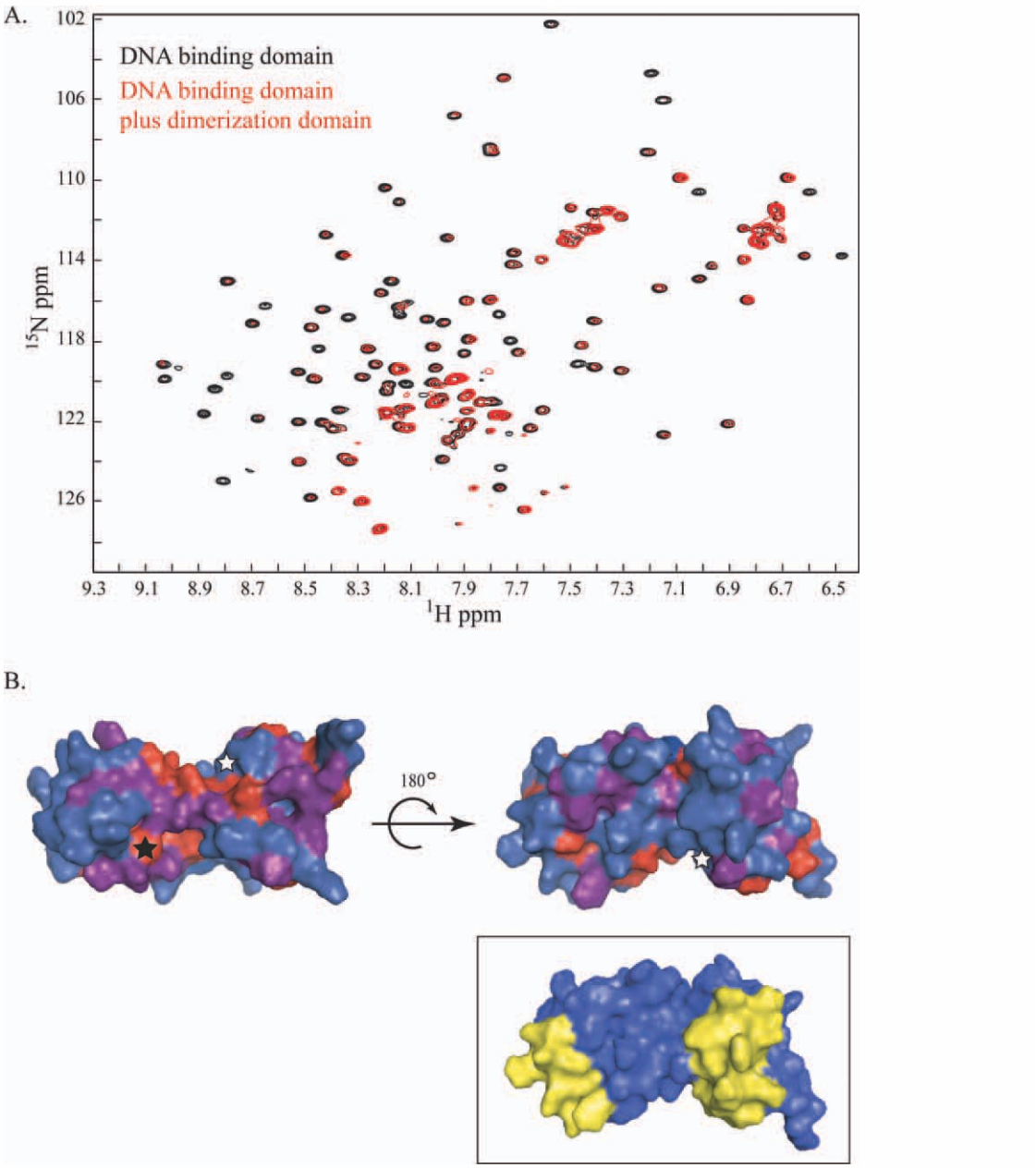

Figure 4

NMR data showing an interaction between the DNA binding domain and dimerization domain of AraC in the presence of arabinose. (**A**) Overlay of HSQC spectra of the ^{15}N-DNA binding domain and ^{15}N-DNA binding domain in the presence of arabinose-bound dimerization domain. (**B**) Map showing areas of the 2K9S DNA binding domain structure where the corresponding HSQC resonances decrease in intensity on binding of the dimerization domain. Red represents residues whose resonances decrease in intensity the most, purple represents residues whose resonances decrease by an intermediate amount, and blue represents residues whose resonances decrease the least. C183, whose resonance decreases by a large amount, is labeled with a black star, and C268, whose resonance does not decrease significantly, is labeled with a white star. The inset shows the DNA contacting helices (yellow; all other residues are blue) of the DNA binding domain for reference.

precisely determine the interaction interface of the DNA binding domain in these experiments, this portion of the AraC DNA binding domain likely includes the interaction interface of this protein.

The potential interface revealed by these NMR experiment may include C183, whose resonance decreases substantially in intensity [Fig. 4(B)], and which the chemical crosslinking experiments identified as also being at or near the interface. The resonance of C268, which the crosslinking experiments indicated is not near the interface, does not decrease significantly in intensity. These NMR experiments directly show that the DNA binding

domain interacts with the dimerization domain in the presence of arabinose and identify a large surface of the DNA binding domain that potentially includes the interaction interface.

FURTHER DISCUSSION

It has been believed for sometime[8,14,17] that an interaction between the dimerization domains and DNA binding domains of AraC in the absence of arabinose constrains the DNA binding domains and makes the formation of a transcriptionally repressive DNA loop energetically favorable.[7–9,12,18–20] Because the evidence also indicates that when arabinose is present, the DNA binding domains are less constrained, and in the interest of simplicity, up to now it has been assumed that the DNA binding domains are free of the dimerization domains in the presence of arabinose.[10,14,22] A further argument in favor of the domains being free in the presence of arabinose is the fact that the two DNA binding domains cannot both interact with the dimerization domains in the same way when bound to the $I_1 - I_2$ DNA site for induction because the dimerization domain possesses C2 rotational symmetry[13,40,41] that is absent in the direct repeat DNA sites.[22,26] Here we show that, in contrast to these extrapolations, a significant interaction between the dimerization domains and DNA binding domains of AraC also occurs in the presence of arabinose.

The following lines of evidence lead to the conclusion that an interaction occurs between the dimerization domains and DNA binding domains of AraC in the presence of arabinose. Full-length AraC in the presence of arabinose dissociates 150 times more rapidly from I_1 half-site DNA than does free DNA binding domain. The arabinose binding affinity of AraC is altered by the binding of I_1 half-site DNA in both a constitutively active mutant in which no domain–domain interaction would be expected, and also in wild type protein. Chemical crosslinking experiments reveal an interaction between the DNA binding domain and dimerization domain of AraC in the presence of arabinose that occurs at or near residue C183 but not near residue C268 of the DNA binding domain. Finally, NMR indicates that an interaction exists between the dimerization domains and DNA binding domains in the presence of arabinose, and, in agreement with the crosslinking experiments, suggests that this interaction occurs on the surface of the DNA binding domain that is on the other side of the domain from the DNA binding surface.

The interaction demonstrated in this work between the DNA binding domains and the dimerization domains of AraC was unexpected in part because, as described above, it is not possible for the dimerization domains and DNA binding domains of each AraC subunit to interact symmetrically and simultaneously bind to the adjacent direct repeat DNA half-sites as is required for induction at p_{BAD}. Thus, when bound to $I_1 - I_2$ DNA, it is impossible for both DNA binding domains of the AraC dimer to be bound to the dimerization domains using similar interfaces. Data presented here, however, show that a DNA binding domain-dimerization domain interaction does occur in the presence of arabinose when AraC is bound to both half-site and full-site DNA. When bound to full-site DNA, that is the direct repeat half-sites $I_1 - I_2$, the interactions between the DNA binding domains and the dimerization domains must either be asymmetrical if both DNA binding domains and dimerization domains are involved in the interaction, or must involve only one DNA binding domain. The results presented here cannot distinguish between these possibilities. We conclude that an interaction occurs between the dimerization domains and DNA binding domains of AraC in the presence of arabinose when AraC is bound to half-site DNA. Some of the methods used here may be useful for the study of interdomain interactions in other systems where structure determination has been problematic.

The data presented in this article do not directly address the question of whether or not the domain–domain interaction that we show in this work to exist in the presence of arabinose utilizes the same interface as the interaction that occurs in the absence of arabinose. On one hand, the fact that the fluorescence change of full-length AraC that occurs in the presence of arabinose on binding to half-site DNA is different from the fluorescence change on binding to half-site DNA in the absence of arabinose suggests that the two interfaces are different. On the other hand, the similarity in the crosslinked products in the experiments with PEAS hints that the interfaces may be the same. These similarities, however, can also be explained by the fact that the linker arm is somewhat long, ~14 Å, and therefore the crosslinking experiments are merely reporting on large-scale aspects of the structure of AraC.

ACKNOWLEDGMENTS

The authors thank Michael Rodgers and Gregory Bowman for assistance, discussion, and helpful comments on this article.

REFERENCES

1. Belogurov GA, Vassylyeva MN, Svetlov V, Klyuyev S, Grishin NV, Vassylyev DG, Artsimovitch I. Structural basis for converting a general transcription factor into an operon-specific virulence regulator. Mol Cell 2007;26:117–129.
2. Aprikian P, Tchesnokova V, Kidd B, Yakovenko O, Yarov-Yarovoy V, Trinchina E, Vogel V, Thomas W, Sokurenko E. Interdomain interaction in the FimH adhesin of *Escherichia coli* regulates the affinity to mannose. J Biol Chem 2007;282:23437–23446.
3. Eldridge AM, Kang HS, Johnson E, Gunsalus R, Dahlquist FW. Effect of phosphorylation on the interdomain interaction of the response regulator, NarL. Biochemistry 2002;41:15173–15180.

4. Nowak E, Panjikar S, Konarev P, Svergun DI, Tucker PA. The structural basis of signal transduction for the response regulator PrrA from *Mycobacterium tuberculosis*. J Biol Chem 2006;281:9659–9666.
5. Robinson VL, Wu T, Stock AM. Structural analysis of the domain interface in DrrB, a response regulator of the OmpR/PhoB subfamily. J Bacteriol 2003;185:4186–4194.
6. Bell CE, Lewis M. A closer view of the conformation of the Lac repressor bound to operator. Nat Struct Biol 2000;7:209–214.
7. Reed WL, Schleif RF. Hemiplegic mutations in AraC protein. J Mol Biol 1999;294:417–425.
8. Saviola B, Seabold R, Schleif RF. Arm-domain interactions in AraC. J Mol Biol 1998;278:539–548.
9. Wu M, Schleif R. Mapping arm-DNA-binding domain interactions in AraC. J Mol Biol 2001;307:1001–1009.
10. Harmer T, Wu M, Schleif R. The role of rigidity in DNA looping-unlooping by AraC. Proc Natl Acad Sci U S A 2001;98:427–431.
11. Englesberg E, Irr J, Power J, Lee N. Positive control of enzyme synthesis by gene C in the L-arabinose system. J Bacteriol 1965;90:946–957.
12. Dunn TM, Hahn S, Ogden S, Schleif RF. 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 U S A 1984;81:5017–5020.
13. Weldon JE, Rodgers ME, Larkin C, Schleif RF. Structure and properties of a truly apo form of AraC dimerization domain. Proteins 2007;66:646–654.
14. Schleif R. AraC protein, regulation of the L-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. FEMS Microbiol Rev 2010;34:779–796.
15. Englesberg E, Squires C, Meronk F, Jr. 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 U S A 1969;62:1100–1107.
16. Sheppard DE, Englesberg E. Further evidence for positive control of the L-arabinose system by gene araC. J Mol Biol 1967;25:443–454.
17. Seabold RR, Schleif RF. Apo-AraC actively seeks to loop. J Mol Biol 1998;278:529–538.
18. Lobell RB, Schleif RF. DNA looping and unlooping by AraC protein. Science 1990;250:528–532.
19. Martin K, Huo L, Schleif RF. 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 U S A 1986;83:3654–3658.
20. Hahn S, Dunn T, Schleif R. Upstream repression and CRP stimulation of the *Escherichia coli* L-arabinose operon. J Mol Biol 1984;180:61–72.
21. Huo L, Martin KJ, Schleif R. Alternative DNA loops regulate the arabinose operon in *Escherichia coli*. Proc Natl Acad Sci U S A 1988;85:5444–5448.
22. Carra JH, Schleif RF. Variation of half-site organization and DNA looping by AraC protein. EMBO J 1993;12:35–44.
23. Zhang X, Schleif R. Obligatory activator-polymerase addition order at promoters. Nucleic Acids Res 1996;24:2575–2577.
24. Lutz R, Lozinski T, Ellinger T, Bujard H. Dissecting the functional program of *Escherichia coli* promoters: the combined mode of action of Lac repressor and AraC activator. Nucleic Acids Res 2001;29:3873–3881.
25. Timmes A, Rodgers M, Schleif R. Biochemical and physiological properties of the DNA binding domain of AraC protein. J Mol Biol 2004;340:731–738.
26. Brunelle A, Schleif R. Determining residue-base interactions between AraC protein and araI DNA. J Mol Biol 1989;209:607–622.
27. Frato KE, Schleif RF. A DNA-assisted binding assay for weak protein-protein interactions. J Mol Biol 2009;394:805–814.
28. Rodgers ME, Schleif R. Solution structure of the DNA binding domain of AraC protein. Proteins 2009;77:202–208.
29. LaRonde-LeBlanc N, Wolberger C. Characterization of the oligomeric states of wild type and mutant AraC. Biochemistry 2000;39:11593–11601.
30. Hanahan D. Techniques for transformation of *E. coli*. In: Glover DM, editor. DNA cloning: a practical approach. Washington D.C.: IRL Press; 1985.
31. Marley J, Lu M, Bracken C. A method for efficient isotopic labeling of recombinant proteins. J Biomol NMR 2001;20:71–75.
32. Ellman GL. Tissue sulfhydryl groups. Arch Biochem Biophys 1959;82:70–77.
33. Schleif, R, Wensink, Practical methods in molecular biology, New York: Springer-Verlag; 1981.
34. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 1995;6:277–293.
35. Johnson BA, Blevins RA. NMR view: a computer program for the visualization and analysis of NMR data. J Biomol NMR 1994;4:603.
36. Ross JJ, Gryczynski U, Schleif R. Mutational analysis of residue roles in AraC function. J Mol Biol 2003;328:85–93.
37. Zuiderweg ER. Mapping protein-protein interactions in solution by NMR spectroscopy. Biochemistry 2002;41:1–7.
38. Vaynberg J, Qin J. Weak protein-protein interactions as probed by NMR spectroscopy. Trends Biotechnol 2006;24:22–27.
39. Greene MK, Maskos K, Landry SJ. Role of the J-domain in the cooperation of Hsp40 with Hsp70. Proc Natl Acad Sci U S A 1998;95:6108–6113.
40. Soisson SM, MacDougall-Shackleton B, Schleif R, Wolberger C. The 1.6 A crystal structure of the AraC sugar-binding and dimerization domain complexed with D-fucose. J Mol Biol 1997;273:226–237.
41. Soisson SM, MacDougall-Shackleton B, Schleif R, Wolberger C. Structural basis for ligand-regulated oligomerization of AraC. Science 1997;276:421–425.