PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Solution structure of the DNA binding domain of AraC protein

Michael E. Rodgers and Robert Schleif*

Biology Department, Johns Hopkins University, Baltimore, Maryland

ABSTRACT

We report the solution structure of the DNA binding domain of the *Escherichia coli* regulatory protein AraC determined in the absence of DNA. The 20 lowest energy structures, determined on the basis of 1507 unambiguous nuclear Overhauser restraints and 180 angle restraints, are well resolved with a pair wise backbone root mean square deviation of 0.7 Å. The protein, free of DNA, is well folded in solution and contains seven helices arranged in two semi-independent sub domains, each containing one helix-turn-helix DNA binding motif, joined by a 19 residue central helix. This solution structure is discussed in the context of extensive biochemical and physiological data on AraC and with respect to the DNA-bound structures of the MarA and Rob homologs.

Proteins 2009; 77:202–208.

Key words: NMR; transcription factor; gene regulation.

INTRODUCTION

Members of the very large AraC family of prokaryotic regulatory proteins typically share ~20% sequence identity over their ~110 residue DNA binding domains (DBDs).[1,2] Most members, including AraC, also contain a domain involved in controlling the protein's activity. In AraC, the additional domain dimerizes the protein,[3,4] contains a binding site for arabinose, and possesses an N-terminal arm that plays a key role in controlling the arabinose-dependent DNA binding properties of the protein.[5,6] Considerable genetic data indicates that the N-terminal arms on the dimerization domains interact with the protein's C-terminal DNA binding domains in the absence of arabinose,[6,7] but not in the sugar's presence.[4,8–11] Thus, the arms control the orientational freedom of the DNA binding domains,[5,8,12–14] (see Fig. 1).

Elucidation of the structural details of many of the important, functional, intramolecular interactions in AraC requires high-resolution structural information. To date, however, despite intense effort, X-ray structure determinations of either AraC or its DNA binding domain have been unsuccessful. Inability to crystallize these proteins, either alone or bound to DNA, may be a consequence of intra- or inter-domain structural flexibility and may also be a consequence of the poor solubility properties of AraC. Other members of the AraC family share the property of low solubility.[15–18] Structure determination has been possible, however, for DNA-bound forms of two AraC family members, MarA[19] and Rob.[20] These findings plus the fact that many DNA binding domains undergo substantial structural rearrangements upon binding to DNA[21] raise the possibility that DNA binding domains in the AraC family may be partially or even largely unfolded when free in solution. Consequently, they would substantially restructure or refold upon binding to DNA. Therefore, it is of considerable interest to determine the structure of the DNA binding domain of an AraC family member in the absence of DNA.

METHODS

Expression vectors

DNA encoding the AraC DNA binding domain region, amino acid residues 175-292, was inserted between the *Nco*1 and *Sac*1 sites in the multiple cloning region in the pET24d vector. The C-terminal truncation was generated using quick-change mutagenesis (Stratagene) to introduce two tandem stop codons (TAATAA) after amino acid position E281. All constructs were verified by DNA sequencing.

Additional Supporting Information may be found in the online version of this article.
Grant sponsor: NIH; Grant number: GM18277.
*Correspondence to: Robert Schleif, Biology Department, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218. E-mail: schleif@jhu.edu
Received 8 January 2009; Revised 28 February 2009; Accepted 15 March 2009
Published online 24 March 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.22431

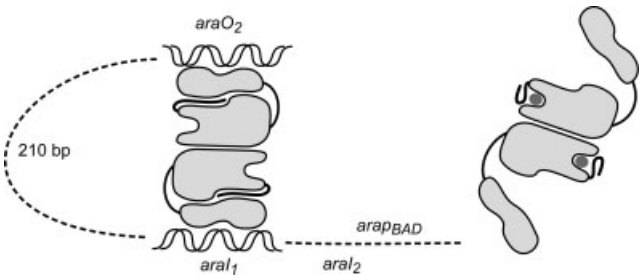

Figure 1

Schematic diagram of the AraC protein and the Ara operon. In the minus arabinose repressing state, left, the AraC DNA binding domains are held in a position that favors looping by binding to the two widely separated DNA half-sites $araO_2$ and $araI_1$. Upon binding arabinose, right, the arms reposition on the dimerization domains over the arabinose. This frees the DNA binding domains to reorient and allows binding to the adjacent direct repeat DNA half-sites $araI_1$ and $araI_2$ partially overlapping the RNA polymerase binding site p_{BAD}.

Purification and isotopic labeling

An isolated colony of cells, BL21(DE3), freshly transformed with the appropriate pET24d expression construct was grown for 8-16 h in 5 mL yeast-tryptone (YT) medium[22] containing 40 μg/mL kanamycin. Unlabeled protein was expressed for 3 h in cultures grown in YT and induced with 0.4 mM isopropyl β-D-thiogalactopyranoside (IPTG) once the A_{550} reached 0.8. ^{15}N-labeled proteins were overexpressed using a modification of the method of Marley *et al.*[23] A starter cell culture grown overnight in minimal glucose medium was used to inoculate new cultures also in minimal glucose medium, which were grown with vigorous aeration at 37°C to an A_{550} of ~0.8. Cells were harvested, washed with M9 salts[22] lacking NH_4Cl, and immediately resuspended in 1/4 of the previous culture volume of minimal medium containing 1 g/L ^{15}N NH_4Cl and either glucose (1% w/v) or ^{13}C-glucose (0.4% w/v) as the carbon source. After 5 min in labeling medium, overexpression was induced by the addition of IPTG to 0.4 mM and cultures were grown for an additional 3 h at 37°C, harvested, and either frozen at −80°C or processed immediately.

Frozen cells, ~5 g, were thawed on ice, resuspended in three volumes of lysis buffer, 20 mM sodium phosphate, pH 6.0, 300 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol, 5% glycerol, and lysed by two passages through a French press. A volume of 1/100, 0.1*M* phenylmethanesulfonyl fluoride, freshly dissolved in ethanol, was added immediately before and again after lysis. The lysate was diluted with an equal volume of lysis buffer and centrifuged for 10 min at 12,000 rpm in a Sorvall SS-34 rotor. The pellet was resuspended in three volumes of 7*M* guanidium-HCl and 10 mM dithiothreitol and allowed to stand for 15 min at room temperature. The sample was diluted in half with 50 mM NaCl, 20 mM Na phosphate, pH 6, and dialyzed twice into 1 L of the same buffer. The dialysate was centrifuged for 10 min at 12,000 rpm in a Sorvall SS-34 rotor and the supernatant loaded onto a 5 mL Heparin Hi-Trap column, Pharmacia, pre-equilibrated with 20 mM Na phosphate, pH 6. Protein was eluted from the column with a 90 mL, 0-1*M* NaCl gradient. The DNA binding domain elutes at about 0.45*M* NaCl, is the only significant elution peak, and is >95% pure based on sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis. The pooled peak fractions from the Heparin HiTrap column were further chromatographed on a Mono-S column using the same buffers and protocol. Heavily overloaded samples run on 15% SDS polyacrylamide gels showed no trace of contaminating proteins. Typically, about 50-100 mg of purified DBD are obtained from 5 g of cells. Purified protein concentrations were calculated from OD_{280} using an extinction coefficient of 8480 $M^{-1}cm^{-1}$ as calculated from the amino acid content by the method of Pace *et al.*[24] Because the protein was refolded out of guanidinium-HCl, the fraction of active protein was assessed using a fluorimetric DNA binding assay.[25] The protein was fully active within experimental error.

NMR spectroscopy

Samples were prepared by dialyzing labeled protein into 444.4 mM NaCl, 55.5 mM Na phosphate, pH 6. D_2O was then added to 10% (v/v) giving a final composition of 400 mM NaCl, 50 mM Na phosphate. If necessary, samples were concentrated using a Millipore Centricon YM-10 concentrator. Typical final protein concentrations were between 200 and 600 μM. Samples in pure D_2O were prepared using two passes through a buffer exchange column equilibrated with 400 mM NaCl, 50 mM Na phosphate, 100% D_2O, pD 6 (apparent pH = 5.6[26]). All NMR experiments were carried out at 25°C on various spectrometers operated by the Biomolecular NMR Center at the Johns Hopkins University. Most data used in the assignments was collected on a Varian INOVA spectrometer operating at a ^{1}H Larmor frequency of 500 MHz. The 3D ^{15}N and ^{13}C-edited NOESY spectra were obtained on a Varian INOVA spectrometer operating at a ^{1}H Larmor frequency of 800 MHz and equipped with a cryogenic probe with *z*-axis pulsed field gradients. All aromatic HSQC and NOESY data were collected on a Bruker Avance spectrometer operating at a ^{1}H Larmor frequency of 600 MHz and equipped with a TCI cryoprobe with *z*-axis pulsed field gradients. Data were processed using NMRPipe[27] and analyzed using NMRView.[28]

Structure determination

Chemical shift assignment of HN, N, CA, and CB resonances was performed manually using four triple resonance experiments, HNCA, HNCACB, HN(CO)CA, and CBCA(CO)NH, data not shown. The fully assigned

^{1}H-^{15}N-HSQC spectrum of the DBD containing residues 175-281 of AraC is shown in Supporting Information Figure 1. The remainder of the chemical shift assignments were also obtained manually using a combination of several spectra including an aliphatic ^{13}C-CTHSQC, aromatic sensitivity enhanced CH-HSQC, aromatic sensitivity enhanced CH-CTHSQC, ^{13}C-edited aromatic 3D NOESY, HNCO, HCC(CO)NH-TOCSY, H_CC(CO)NH-TOCSY, and HCCH-COSY. Overall completeness of assignments is 92.7%.

NOE spectra were converted to XEASY format using the program CARA.[29] Automated NOE peak picking, identification, and assignment were performed with the stand-alone ATNOS/CANDID program.[30,31] The NMR solution structure was calculated using the program CYANA.[32,33] Conversion of NOE peak intensities to distance restraints was done using automatic calibration. Initial dihedral angle restraints were generated using TALOS[34] and were input during the CYANA calculations. The final structure was refined using CNS with explicit water as the solvent using scripts obtained from the RECOORD project web site.[35] Structures were analyzed using PROCHECK-NMR.[36] Graphical representations were generated using either MOLMOL[37] or VMD.[38]

The structure and NMR restraints have been deposited with the RCSB PDB and the BMRB. The accession numbers are RCSB ID code ***rcsb100855***, PDB ID code ***2k9s***, and BMRB accession number ***16001***.

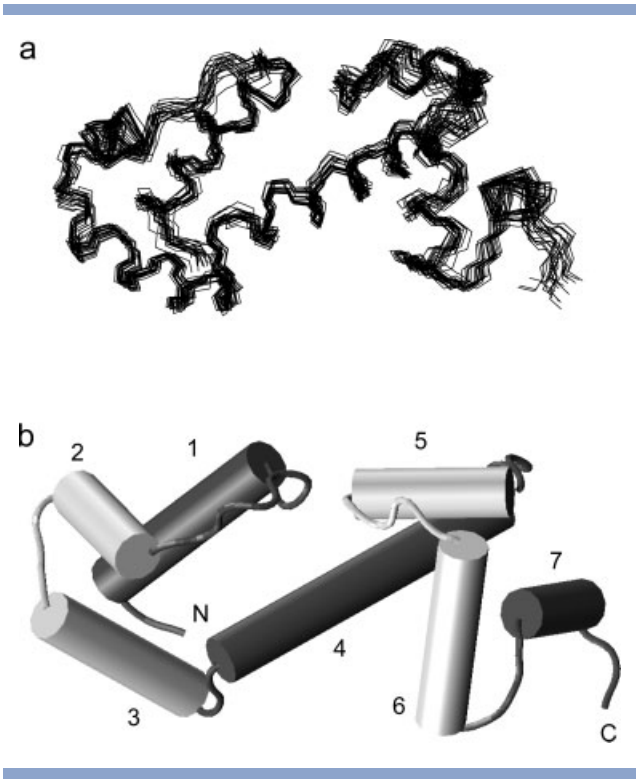

Figure 2

Structure of AraC DBD. (**a**) Overlay of the CA traces of the ensemble of 20 structures with the lowest final target function values. The RMSD to the mean was 0.68 Å. (**b**) Representation of the lowest energy structure refined in explicit water. The orientation is such that the two DNA contacting helices, 3 and 6, are foremost. The two helix-turn-helix (H-T-H) motifs are light colored. The minimal contact region between the two H-T-H subdomains is at the top center of the diagram.

RESULTS AND DISCUSSION

Structure determination of the AraC DNA binding domain

Initial NMR studies of the DNA binding domain of AraC suggested that the C-terminal portion of the domain is unstructured. The ^{1}H-^{15}N HSQC spectrum of the AraC DBD contained about 10 extremely intense, sharp resonances relative to the others, suggesting that they derived from an unstructured segment. Chemical shift redundancy of these crosspeaks with several of the less intense peaks interfered with initial resonance assignment. As an 11-residue C-terminal tail of AraC can be deleted without affecting the protein's inducing or repressing properties[39] and is absent from some natural homologs of the protein, we removed these residues. As expected, the ^{1}H-^{15}N-HSQC spectrum from this construct lacked the very sharp resonances, and all other backbone crosspeaks were of comparable intensity (Supp. Info. Fig. 1). We therefore used the C-truncated DBD for all subsequent work reported here.

The resonance assignments obtained as outlined in the Methods section, and the three-dimensional NOE spectroscopy experiments, were used as inputs for the structure calculation. Initial dihedral restraints were generated by chemical shift analysis and the final structure ensemble was generated using these restraints and NOE distance restraints. The 20 structures with the lowest final target function values, [Fig. 2(a)], possessed an average backbone root mean square deviation (RMSD) to mean of 0.7 Å. The NMR structure statistics are summarized in Table 1. The final lowest energy structure from 20 structures, refined in explicit water, is shown [Fig. 2(b)].

Structure description

The AraC DNA binding domain consists of seven alpha helices in two weakly interconnected subdomains. Each subdomain contains one helix-turn-helix (H-T-H) motif, helices 2 and 3 and helices 5 and 6 in the N- and C-terminal subdomains, respectively. The two subdomains are joined by a 19-residue central helix, helix 4. Only eight long-range NOE restraints are observed between the two subdomains, those involving residues L16, A17, V79, and G80. The absence of other contacts between the subdomains allows significant rotation of one H-T-H subdomain with respect to the other. This freedom is seen in the NMR structure ensemble in which either the N-terminal or the C-terminal subdomains can

Table I
Refinement Statistics

	Ensemble
NOE upper distance limits	
Total	1507
Short range $\|i-j\| \leq 1$	766
Medium range $1 < \|i-j\| < 5$	396
Long range $\|i-j\| \geq 5$	345
Dihedral angle constraints	180
Residual target function value (Å^2)	0.84 ± 0.12
Residual distance restraint violations	
Number > 0.1 Å	3 ± 1
Maximum, Å	0.16 ± 0.04
Residual dihedral angle restraint violations	
Number > 5°	1 ± 0
Maximum (°)	9.09 ± 0.64
RMSD from average structure (residues 4–103)	
Backbone (Å)	0.68 ± 0.14
Heavy atoms (Å)	1.10 ± 0.12
PROCHECK parameters	
Most favored (%)	98.4
Additionally allowed (%)	1.6
Generously allowed (%)	0.0
Disallowed (%)	0.0

individually be overlaid well, but the relative positioning of the two subdomains shows considerably greater dispersion (see Fig. 3). That dispersion leads to the relatively large RMSD values reported in Table 1. The backbone RMSD of the individual subdomains is considerably lower, 0.48 and 0.49 Å for the N- and C-terminal subdomains, respectively. Also consistent with the idea that the two subdomains may move somewhat independently of one another is the fact that calculating the structures of each subdomain separately, residues 1-57 and residues 57-107, using the same NOE data sets as above gives much better structural statistics with an average backbone RMSD to mean of 0.47 Å for both. Our data provide no evidence for the existence of any significantly populated alternate structures or of a substantial amount of unstructured protein.

Our results do not support the possibility mentioned above that the poor solubility of AraC results from unfolded DNA binding domains. Other sources of the protein's low solubility and aggregation include oxidation of cysteine residues or domain swapping of helix-turn-helix regions. Oxidation seems unlikely, however, because fully reduced DBD at high concentration aggregates slowly even when stored under argon. The possibility of intermolecular domain swapping[40] of the helix-turn-helix subdodmains of AraC DBD appears compatible with the very weak interdomain interactions between the two helix-turn-helix regions as well as the likelihood that in the process of binding to DNA, one of the helix-turn-helix regions in the AraC DBD undergoes considerable rotation with respect to the other helix-turn-helix region. The increased molecular weight associated with such swapping could be detected with pulsed-field gradient NMR techniques[41,42] or other physical techniques sensitive to the hydrodynamic radius.

Mechanistic considerations

The sequence identity between AraC and the homologs MarA and Rob is only 21 and 25%, respectively, and similarity is ~45% for both. Nevertheless, the solution structure of the DNA binding domain of AraC is similar to the DNA bound structures of MarA and Rob with pair wise backbone RMSDs of 3.2 and 5 Å, respectively. The seven helical regions in the three proteins begin and end at about the same positions in the homology aligned sequences, (see Fig. 4), and they have roughly similar positions and orientations in space with respect to each other as shown in the overlay of AraC and MarA (see Fig. 5).

What are the implications of the AraC DBD structure for binding to DNA? If the two H-T-H regions of AraC were to interact homologously with B-form DNA sites one turn apart, the regions would be oriented similarly, and would be separated by about 34 Å. In fact, in the structure, the second helix of the two H-T-H regions differ in orientation by nearly 50°, and they are separated by only about 26 Å. Thus, either the interactions are not

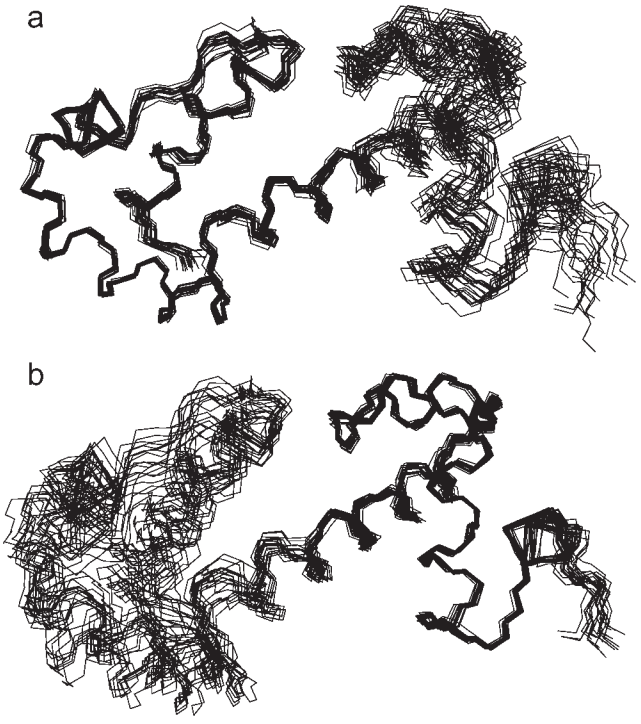

Figure 3
Overlays of subdomains of AraC DBD. (**a**) Ensemble of the 20 structures from Fig. 2b in which residues 4-57 were used for RMS positioning. The RMSD to the mean was 0.48 Å. (**b**) Like (a) but residues 57-103 were used. The RMSD to the mean was 0.49 Å.

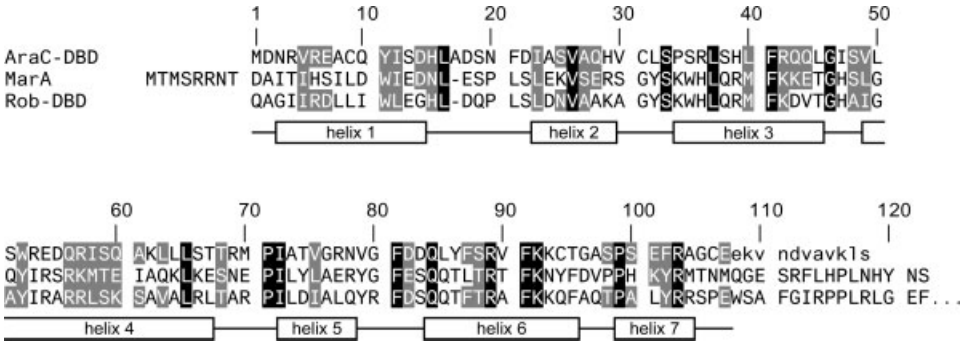

Figure 4

Alignment of AraC with homologs MarA and Rob. The upper line shows the sequence number corresponding to the AraC DBD used in this article. Residue 1 corresponds to residue 175 in the native AraC. Alignment between the AraC-DBD sequence, the complete MarA sequence, and the Rob DNA binding domain is presented. Identical residues in all three are highlighted in black. Gray highlighted residues are similar or identical between AraC-DBD and either or both MarA and Rob. The boxes at the bottom correspond to alpha helical regions in AraC-DBD. The helix ends in the MarA and Rob structures are at identical positions ± 1 residue of the indicated positions. The N-terminal and C-terminal H-T-H motifs comprise helices 2-3 and 5-6, respectively. The lowercase residues in AraC-DBD are the residues that were truncated prior to structure determination.

homologous and one turn apart or significant distortions in the protein and/or DNA occur upon binding. Analysis of DNA contacts by AraC show that the major energetic contributions from bases in the $araI_1$ DNA half-site derive from two regions of five or six base pairs separated by 10 or 11 base pairs.[43] Taken together these facts indicate that both AraC DBD and the DNA binding site are significantly distorted in the protein-DNA complex. This deduction is compatible with the observation that the binding of AraC to the full *araI* site generates a substantial bend, about 90°.[44]

The two H-T-H regions of MarA clearly do not interact homologously with DNA. In the crystal structure of MarA bound to DNA, the DNA-contacting helix of the second H-T-H region sits within the major groove of the DNA, whereas the DNA-contacting helix of the first H-T-H region points more end-on into the major groove of the significantly distorted DNA binding site.[19] In the Rob-DNA structure, the DNA is unbent and the second helix does not extend into the major grove at all, it rests on the side of the DNA binding site.[20] This raises the question of whether the crystal structure of Rob and the *micF* promoter accurately represents the solution structure or the biologically functional structure of Rob-DNA complexes because in the complexes of Rob and the *zwf* and *fumC* promoters, Rob does bend the DNA.[45] In the DNA-bound structures of both MarA and Rob, the distance between the centers of the DNA-contacting helices is about 26 Å, about the same as we find in the DNA-free solution structure of AraC. Thus, as expected, the DNA to which MarA has bound is substantially bent, ∼35°.

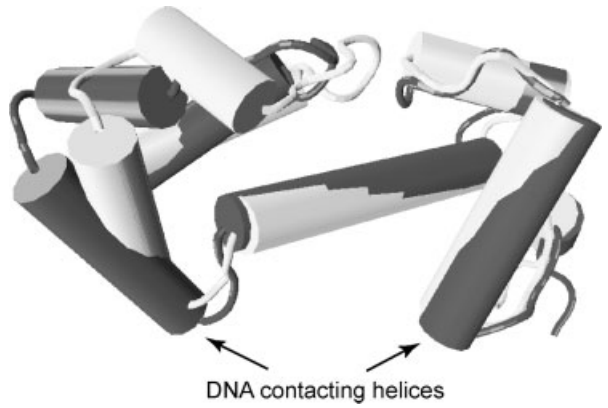

Figure 5

Overlay of AraC and MarA. The solution structure of AraC-DBD, dark, overlaid with the DNA-bound structure of MarA, light. Only backbone atoms of the C-terminal H-T-H regions were used to calculate the overlay, which had an RMS difference of 1.3 Å.

As mentioned earlier, the final 11 residues of the AraC DBD, unstructured in the 0.4*M* NaCl used in the structure work, are not essential to the regulatory activities of AraC and were deleted for the structure determination. The poorly homologous C-terminal sequence in MarA is structured,[19] at least when the protein is bound to DNA and in the lower salt concentration crystallization buffer. Although nonessential, it is possible that the C-terminal tail of AraC is also structured when bound to DNA and in low salt concentration.[46] Although truncation of unstructured peptides often facilitates crystal formation, attempts to crystallize the C-truncated DBD were unsuccessful and similar truncation in "full length" AraC renders the protein considerably less soluble than normal.

In summary, the solution structure of the DNA binding domain of AraC contains two weakly interacting

helix-turn-helix regions connected by an α-helix. It is similar, but not identical to the DNA-bound structures of two family members with 20% sequence identity. Either AraC or the DNA or both distort by on the order of 5 Å upon DNA binding. No evidence is seen that the DBD of AraC or any significant part of it is unstructured for a significant fraction of the time when the protein is free in solution. This structure and the dimerization domain structure will permit docking calculations to address the question of information transfer between the two domains that is required for normal AraC function.

ACKNOWLEDGMENTS

The authors thank Dr. Ananya Majumdar and the Johns Hopkins Biomolecular NMR Center for extensive assistance with all NMR experiments; Drs. Tingting Ju and Joel Tolman for use of their computers and assistance in the structure calculation; Dr. Blake Hill for additional assistance; and Katie Frato, Stephanie Dirla, and Jennifer Seedorff for comments on the manuscript.

REFERENCES

1. Gallegos M, Schleif R, Bairoch A, Hofmann K, Ramos J. AraC/XylS family of transcriptional regulators. Microbiol Mol Biol Rev 1997;61:393–410.
2. Egan SM. Growing repertoire of AraC/XylS activators. J Bacteriol 2002;184:5529–5532.
3. Soisson S, MacDougall-Shackleton B, Schleif R, Wolberger C. Structural basis for ligand-regulated oligomerization of AraC. Science 1997;276:421–425.
4. Bustos S, Schleif R. Functional domains of AraC protein. Proc Natl Acad Sci USA 1993;90:5638–5642.
5. Rodgers ME, Holder N, Dirla S, Schleif R. Functional modes of the regulatory arm of AraC. Proteins 2009;74:81–90.
6. Saviola B, Seabold R, Schleif R. Arm-domain interactions in AraC. J Mol Biol 1998;278:539–548.
7. Wu M, Schleif R. Mapping arm-DNA-binding domain interactions in AraC. J Mol Biol 2001;307:1001–1009.
8. Lobel R, Schleif R. DNA looping and unlooping by AraC protein. Science 1990;250:528–532.
9. Carra J, Schleif R. Variation of half-site organization and DNA looping by AraC protein. EMBO J 1993;12:35–44.
10. Seabold R, Schleif R. Apo-AraC actively seeks to loop. J Mol Biol 1998;278:529–538.
11. Harmer T, Wu M, Schleif R. The role of rigidity in DNA looping-unlooping by AraC. Proc Natl Acad Sci USA 2001;98:427–431.
12. Schleif R. Regulation of the L-arabinose operon of *Escherichia coli*. Trends Genet 2000;16:559–565.
13. Schleif R. The AraC protein: a love-hate relationship. Bio Essays 2003;25:274–282.
14. Martin K, Huo L, Schleif R. 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 Nat Acad Sci USA 1986;83:3654–3658.
15. Grainger DC, Belyaeva TA, Lee DJ, Hyde EI, Busby SJW. Binding of the *Escherichia coli* MelR protein to the melAB promoter: orientation of MelR subunits and investigation of MelR-DNA contacts. Mol Microbiol 2003;48:335–348.
16. Egan S, Schleif R. DNA-dependent renaturation of an insoluble DNA binding protein. J Mol Biol 1994;243:821–829.
17. Domínguez-Cuevas P, Marín P, Busby S, Ramos JL, Marqués S. Roles of effectors in XylS-dependent transcription activation: intramolecular domain derepression and DNA binding. J Bacteriol 2008;190:3118–3128.
18. Yamazaki H, Tomono A, Ohnishi Y, Horinouchi S. DNA-binding specificity of AdpA, a transcriptional activator in the A-factor regulatory cascade in *Streptomyces griseus*. Mol Microbiol 2004;53:555–572.
19. Rhee S, Martin RG, Rosner JL, Davies DR. A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. Proc Natl Acad Sci USA 1998;95:10413–10418.
20. Kwon HJ, Bennik MH, Demple B, Ellenberger T. Crystal structure of the *Escherichia coli* Rob transcription factor in complex with DNA. Nat Struct Biol 2000;7:424–430.
21. Spolar RS, Record MT Jr. Coupling of local folding to site-specific binding of proteins to DNA. Science 1994;263:777–784.
22. Schleif RF, Wensink PC. Practical methods in molecular biology. New York: Springer-Verlag; 1981. pp 196–198.
23. Marley J, Lu M, Bracken C. A method for efficient isotopic labeling of recombinant proteins. J Biomol NMR 2001;20:71–75.
24. Pace CN, Vajdos F, Lee L, Grimsley G, Gray T. How to measure and predict the molar absorption coefficient of a protein. Protein Sci 1995;4:2411–2423.
25. Timmes A, Rodgers ME, Schleif R. Biochemical and physiological properties of the DNA binding domain of AraC protein. J Mol Biol 2004;340:731–738.
26. Mikkelsen K, Nielsen SO. Acidity measurements with the glass electrode in H_2O-D_2O mixtures. J Phys Chem 1960;64:632–637.
27. 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.
28. Johnson BA, Blevins RA. NMRView: a computer program for the visualization and analysis of NMR data. J Biomol NMR 1994;4:603–614.
29. Keller R. The computer aided resonance assignment tutorial, 1st ed. Switzerland: Cantina-Verlag; 2004. ISBN 3-85600-112-3.
30. Herrmann T, Güntert P, Wüthrich K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. J Biomol NMR 2002;24:171–189.
31. Herrmann T, Güntert P, Wüthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Mol Biol 2002;319:209–227.
32. Lopez-Mendez B, Jee J-G, Pantoja-Uceda D, Güntert P. Improved automated NMR structure calculation with CYANA 20 NMR. Toronkai Koen Yoshishu 2003;42:282–283.
33. Güntert P. Automated NMR structure calculation with CYANA. Meth Mol Biol 2004;278:353–378.
34. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 1999;13:289–302.
35. Nederveen AJ, Doreleijers JF, Vranken W, Miller Z, Spronk CAEM, Nabuurs SB, Güntert P, Livny M, Markley JL, Nilges M, Ulrich EL, Kaptein R, Bonvin AMJJ. RECOORD: a REcalculated COORdinates Database of 500+ proteins from the PDB using restraints from the BioMagResBank. Proteins 2005;59:662–672.
36. Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 1996;8:477–486.
37. Koradi R, Billeter M, Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graphics 1996;14:51–55.
38. Humphrey W, Dalke A, Schulten K. VMD—visual molecular dynamics. J Mol Graphics 1996;14:33–38.

39. Eustance R, Bustos S, Schleif R. Reaching out: locating and lengthening the interdomain linker in AraC protein. J Mol Biol 1994;242: 330–338.
40. Liu Y, Eisenberg D. 3D domain swapping: as domains continue to swap. Protein Sci 2002;11:1285–1298.
41. Stejskal EO, Tanner JE. Spin diffusion measurements: spin echos in the presence of a time-dependent field gradient. J Chem Phys 42 1965;42:288–291.
42. Gibbs SJ, Johnson CSJ. A PFG experiment for accurate diffusion and flow studies in the presence of eddy currents. J Magn Reson 1991;93:395–402.
43. Niland P, Hühne R, Müller-Hill B. How AraC interacts specifically with its target DNAs. J Mol Biol 1996;264:667–674.
44. Saviola B, Seabold R, Schleif R. DNA bending by AraC: a negative mutant. J Bacteriol 1998;180:4227–4232.
45. Jair K, Yu X, Skarstad K, Thöny B, Fujita N, Ishihama A, Wolf R Jr. Transcriptional activation of promoters of the superoxide and multiple antibiotic resistance regulons by Rob, a binding protein of the *Escherichia coli* origin of chromosomal replication. J Bacteriol 1996;178:2507–2513.
46. Harmer T, Schleif R. The C-terminal end of AraC tightly binds to the rest of its domain. J Biol Chem 2001;276:4886–4888.