Structural Basis for Ligand-Regulated Oligomerization of AraC

Stephen M. Soisson, Beth MacDougall-Shackleton, Robert Schleif,* Cynthia Wolberger*

The crystal structure of the arabinose-binding and dimerization domain of the *Escherchia coli* gene regulatory protein AraC was determined in the presence and absence of L-arabinose. The 1.5 angstrom structure of the arabinose-bound molecule shows that the protein adopts an unusual fold, binding sugar within a β barrel and completely burying the arabinose with the amino-terminal arm of the protein. Dimer contacts in the presence of arabinose are mediated by an antiparallel coiled-coil. In the 2.8 angstrom structure of the uncomplexed protein, the amino-terminal arm is disordered, uncovering the sugar-binding pocket and allowing it to serve as an oligomerization interface. The ligand-gated oligomerization as seen in AraC provides the basis of a plausible mechanism for modulating the protein's DNA-looping properties.

AraC is a regulator of transcription that changes the way in which it binds DNA when the protein forms a complex with its monosaccharide ligand, L-arabinose. In *E. coli*, the AraC protein controls expression of genes necessary for uptake and catabolism of arabinose (*1*). In the absence of arabinose, a single AraC dimer contacts the two widely separated I_1 and O_2 half-sites, forming a 210–base pair DNA loop and repressing transcription from the p_{BAD} and p_C promoters (*2–5*) (Fig. 1). The binding of arabinose to AraC causes the protein to cease DNA looping and favor binding to the adjacent I_1 and I_2 half-sites (*4*, *6*), resulting in activation of transcription from the p_{BAD} promoter (*7*). DNA looping, which was discovered in the arabinose operon (*2*), plays an important role in both eukaryotic and prokaryotic transcription and has been shown to be a mechanism for allowing control elements distally located from the start site of transcription to participate in gene regulation (*8*).

The 292-residue AraC protein consists of an NH_2-terminal domain that binds arabinose and mediates dimerization (residues 1 to 170) joined to a COOH-terminal DNA-binding domain (residues 178 to 292) (*9*, *10*) by a linker of at least five residues (*10–12*). The two domains of AraC appear to be functionally independent, as they retain their respective activities when fused to other DNA-binding or dimerization domains (*9*). At protein concentrations near those found in the cell, AraC is a dimer in solution both in the presence and absence of arabinose (*13*). It is also a single AraC dimer that binds to the I_1-I_2 site in the presence of arabinose and to the widely separated I_1 and O_2 sites in the absence of arabinose (*4*, *5*). DNA-binding and footprinting experiments indicate that the binding of arabinose restricts the freedom of the DNA-binding domains of AraC to contact nonadjacent DNA half-sites, probably by reducing the distance that the DNA-binding domains can be separated (*14*). The binding of arabinose to AraC must therefore alter the protein dimer in a way that changes its preference for binding site spacing.

To learn the basis for the pronounced effect of arabinose on the dimerization domain of AraC, we have determined its structure in the presence and absence of arabinose. Crystals of the dimerization domain of AraC complexed with L-arabinose were prepared from a tryptic fragment of the protein consisting of the NH_2-terminal domain of AraC (residues 2 to 178) (*15*). The crystals form in space group $P2_1$ with unit cell dimensions a = 39.75 Å, b = 93.84 Å, c = 50.33 Å, β = 95.62° and contain two monomers in the asymmetric unit. The structure of the AraC-arabinose complex was determined at 1.5 Å resolution (Table 1) and contains residues 7 to 167 of monomer A and residues 7 to 170 of monomer B, two molecules of L-arabinose, and 412 water molecules. The structure of the unliganded AraC dimerization domain was determined at 2.8 Å resolution (Table 1) from crystals that form in space group $P3_121$ with unit cell dimensions $a = b$ = 57.37 Å, c = 83.45 Å (*16*) and contain one monomer in the asymmetric unit. The model of the unliganded dimerization domain contains residues 19 to 167 of AraC.

The structure of AraC complexed with arabinose shows that the fold of the sugar-binding and dimerization domain of AraC contains an eight-stranded antiparallel β barrel (β1 to β8) with jelly-roll topology (Fig. 2A). The β barrel is followed by a long linker that contains two turns of 3_{10} helix, followed by a ninth β strand (β9) that forms part of one sheet of the β barrel. The last β strand is followed by two α helices (α1 and α2), each approximately 20 amino acids in length, that pack against the outer surface of the barrel. One molecule of L-arabinose binds in a pocket within the β barrel (Fig. 2, A and B). Residues 7 to 18 of AraC constitute an NH_2-terminal arm that lies across the sugar-binding pocket, fully enclosing the arabinose molecule within the β-barrel domain (Fig. 2A). The overall fold of the NH_2-terminal domain of AraC and the manner in which it binds ligand has not, to our knowledge, been previously observed. Residues 2 to 6 are not visible at the NH_2-terminus of either monomer in the asymmetric unit and are presumably disordered, as are the COOH-terminal residues 168 to 178 of monomer A and residues 171 to 178 of monomer B. The disorder at the COOH-terminus is in agreement with previous experiments indicating that these residues are part of a flexible and mutatable linker that joins the sugar-binding and dimerization domain to the DNA-binding domain (*9*, *10*, *12*).

Each monomer of AraC binds one molecule of α-L-arabinose (*17*) in the full-chair conformation within the open end of the β barrel, occupying a small pocket lined with both polar and nonpolar side chains (Fig. 2,

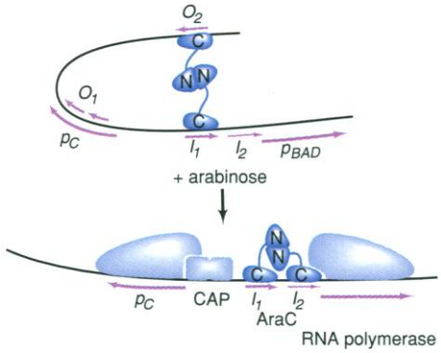

Fig. 1. Regulation of the *araBAD* operon by the AraC protein. Relative to the start site of transcription, the I_1 and I_2 half-sites are positioned at −35, the O_1 site at −125, and the O_2 site at −285. Binding of the CAP protein is not essential, but is required to achieve the maximum level of transcriptional activation from the *araBAD* promoter. N, NH_2-terminal arabinose-binding and dimerization domain of AraC; C, COOH-terminal DNA-binding domain.

S. M. Soisson, Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.
B. MacDougall-Shackleton and R. Schleif, Biology Department, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA.
C. Wolberger, Department of Biophysics and Biophysical Chemistry and the Howard Hughes Medical Institute, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.

*To whom correspondence should be addressed. E-mail: bob@gene.bio.jhu.edu (R.S) or cynthia@groucho.med.jhmi.edu (C.W.)

A to C). The sugar stacks against the indole ring of Trp^{95} and is stabilized by hydrogen bonds formed by side chains and water molecules within the binding pocket with the sugar hydroxyls and ring oxygen (Fig. 2C). The binding site is completed by the NH_2-terminal arm of the protein (residues 7 to 18), which loops around to close off the end of the β barrel in which arabinose is bound. The NH_2-terminal arm forms both direct and indirect contacts with the sugar, resulting in complete burial of arabinose within the protein (*18*). The position of the arm is stabilized by the main chain carbonyl of Pro^8, which makes a direct hydrogen bond with the anomeric hydroxyl (OH-1) of the bound sugar. An extensive network of bound water molecules in the sugar-binding pocket mediates hydrogen bonds between the sugar and protein main chain atoms of residues Leu^9, Leu^{10}, Gly^{12}, Tyr^{13}, and Phe^{15} in the NH_2-terminal arm (Fig. 2, B and C). Because the NH_2-terminal arm blocks access to the arabinose-binding pocket, the arm must move aside when arabinose enters or leaves.

The structure of the unliganded AraC monomer is very similar to the structure of the liganded monomer, with residues 19 to 167 of the two structures superimposable with a root-mean-square difference (rmsd) of 0.6 Å for main chain atoms (Fig. 2D). There are, however, several notable differences between the liganded and unliganded molecules. Most prominent of these is that, in the absence of sugar, the NH_2-terminal arm (residues 7 to 18) that formerly closed off the opening to the ligand-binding site is disordered. Evidence for the disorder comes both from the absence of electron density corresponding to the NH_2-terminal arm and from a packing arrangement in the crystal (described below) that is made possible by the absence of the NH_2-terminal arm from the opening of the sugar-binding pocket. In the presence of arabinose, both direct and water-mediated hydrogen bonds to the sugar help stabilize the NH_2-terminal arm. It is therefore not surprising that the arm is unstructured in the absence of arabinose.

At intracellular concentrations, the AraC protein exists as a dimer in both the presence and absence of arabinose (*13*). In the presence of arabinose, the NH_2-terminal domain of AraC crystallizes with a dimer in the asymmetric unit (Fig. 3A).

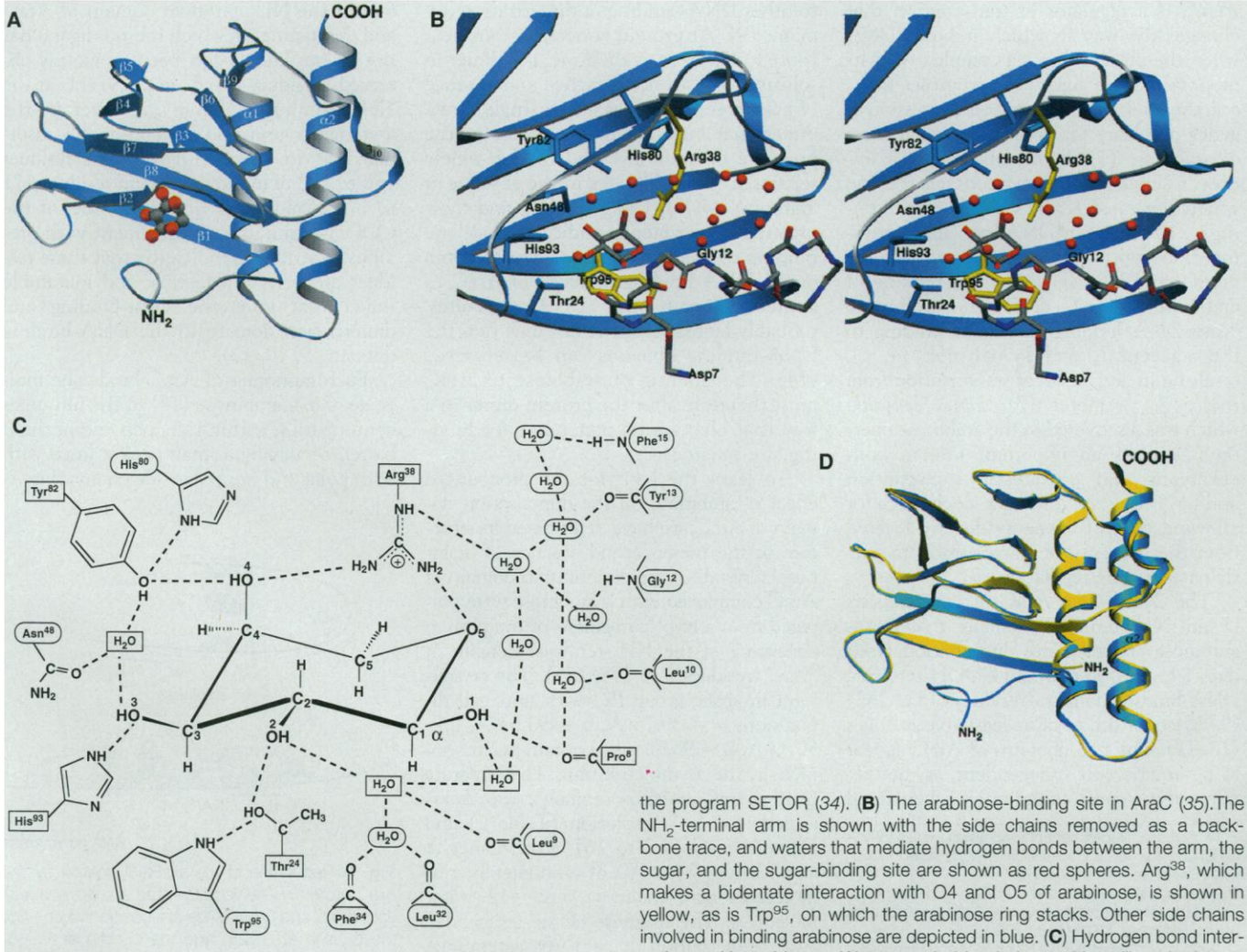

Fig. 2. The AraC sugar-binding and dimerization domain monomer. (**A**) Monomer in the presence of arabinose. The protein is shown as a blue ribbon trace, and the sugar as a ball-and-stick model. The NH_2- and COOH-termini, as well as secondary-structure elements referred to in the text, are indicated. This figure, as well as Figs. 2, B and D, and 3, A and B, were prepared with the program SETOR (*34*). (**B**) The arabinose-binding site in AraC (*35*).The NH_2-terminal arm is shown with the side chains removed as a backbone trace, and waters that mediate hydrogen bonds between the arm, the sugar, and the sugar-binding site are shown as red spheres. Arg^{38}, which makes a bidentate interaction with O4 and O5 of arabinose, is shown in yellow, as is Trp^{95}, on which the arabinose ring stacks. Other side chains involved in binding arabinose are depicted in blue. (**C**) Hydrogen bond interactions with arabinose. Waters and side chains making direct hydrogen bonds to the sugar are in boxes, whereas those making indirect hydrogen bonds to the sugar are in ovals. Several waters have been omitted from the network of water-mediated hydrogen bonds along the right side of the figure for clarity. (**D**) Superposition of the AraC structures in the presence (blue) and absence (yellow) of arabinose.

This arrangement presumably reflects the solution dimer form of the protein, as the crystals contain no other twofold axes relating AraC monomers. The two monomers in the crystal associate by an antiparallel coiled-coil formed between the terminal α helix, helix 2, of each monomer (Fig. 3A). Although a few additional contacts are made between the short 3_{10} helices, most of the 1200 Å^2 dimer interface consists of the coiled-coil interaction. Each end of the coiled-coil is anchored by a triad of leucine residues, Leu^{150} and Leu^{151} from one monomer, and Leu^{161} from the second, that pack together in a knobs-into-holes manner. This packing scheme persists along the length of the helices, consistent with the model proposed by Crick for a coiled-coil (*19*).

The two nonpolar leucine triads flank a hydrophilic central core of the coiled-coil consisting of polar side chains that form a network of hydrogen bonds with each other and with a molecule presumed to be a water (*20*) that is buried at the interface between the two helices. The buried water stabilizes the central portion of the coiled-coil interface by forming four hydrogen bonds, one each to Asn^{154} and Gln^{158} from both monomers. Glu^{157} from each helix further stabilizes Asn^{154} through a hydrogen bond. Although two-helix antiparallel coiled-coils between two monomers are relatively rare, the coiled-coil seen in the AraC arabinose dimer is remarkably similar to the coiled-coil seen in the MetJ-DNA complex (*21*). In that case, MetJ dimers bound to adjacent sites on the DNA interact by way of an antiparallel coiled-coil that contains hydrophobic interactions at its ends flanking a hydrophilic core that buries a water molecule.

In the absence of arabinose, the dimerization domain of AraC crystallizes with a monomer in the crystallographic asymmetric unit. The crystal-packing interactions must therefore be examined to identify the solution dimer interface. Each unliganded AraC monomer in the crystal has two neighboring monomers related to it by independent crystallographic twofold axes, giving rise to two potential choices of dimerization interface. One of the two dimer interfaces in the crystals of the unliganded protein reproduces nearly the same coiled-coil interaction seen in the crystals of arabinose-bound AraC. In the unliganded protein, however, a distortion in the top of helix 2 slightly reorients the leucine triad at the ends of the coiled-coil, resulting in an approximate 13° rotation of one monomer relative to the other as compared with the coiled-coil dimer formed by AraC with bound arabinose. The rotation is a hinge-like motion about the coiled-coil axis, giving rise to a slight rearrangement of the packing interactions at the coiled-coil interface.

The second dimer interface in the uncomplexed AraC crystals (Fig. 3B) brings together two β barrels, in a face-to-face manner, and buries 75% more area than the coiled-coil interface, giving a total of 2100 Å^2 of buried surface area. This seems likely to be the dimerization interface in the absence of arabinose, as in crystals of virtually all multimeric proteins, the biologically relevant interface is that which buries the most surface area (*22*). It is the absence of arabinose and the disordering of the NH_2-terminal arm that exposes the β-barrel surface, allowing it to function as an oligomerization interface. A striking feature of the β-barrel interface is the insertion of Tyr^{31} from each monomer into the sugar-binding site of the opposing monomer where the tyrosine side chain takes the place of arab-

Table 1. A native data set from the arabinose-bound crystals was collected to a resolution of 1.8 Å on an RAXIS-IIC detector with CuKα radiation. All data were processed and reduced with DENZO and SCALEPACK (*26*). A mercury derivative with two atoms per asymmetric unit was prepared by soaking crystals in stabilizing solution containing 0.9 M ethylmercurithiosalicylate for 3 days. These crystals were transferred to the standard cryosolution for 20 min and flash-frozen (*15*). A complete derivative data set was collected at beamline X4A of the National Synchrotron Light Source at Brookhaven National Laboratories (NSLS, Upton, New York) with the use of a Fuji BAS-2000 image plate detection system. Data were collected at a wavelength of 1.006 Å, just above the mercury L_{III} absorption edge, in order to maximize the anomalous scattering from the mercury atoms in the crystal. These data were scaled to the native data set with SCALEIT (*27*), and single isomorphous replacement plus anomalous scattering (SIRAS) phases were calculated with MLPHARE (*27*). The initial phases were improved further by applying solvent flattening and histogram matching with DM (*27*), resulting in a high-quality 1.8 Å electron density map. The model was built with the interactive graphics program O (*28*), and refinement proceeded with X-PLOR (*29*) through use of a 1.5 Å native data set recorded at the NSLS and standard refinement protocols (*30*). The structure of the unliganded AraC sugar-binding and dimerization domain was determined by molecular replacement with AMoRe with data collected at beamline X4A (*31*). One monomer of the plus arabinose AraC structure with the sugar and waters deleted was used as a search model. After rigid-body refinement of the structure, manual corrections of the model were made, including deletion of the NH_2-terminal 12 residues. The structure was then refined by torsion angle dynamics refinement (*32, 33*) and conventional conjugate gradient positional refinement in X-PLOR. Both the arabinose-bound and unliganded AraC models show good stereochemistry and no Ramachandran plot outliers.

	Native 1	Native 2	EMTS	Unliganded
Wavelength (Å)	1.54	0.689	1.006	1.072
Resolution limit (Å)	1.8	1.5	1.5	2.7
Measured reflections	102,201	219,475	340,323	51,550
Unique reflections	29,234	63,588	59,145	4,676
Completeness (%)	86.0	91.0	90.0	99.0
Overall $I/\sigma(I)$	17.7	10.9	7.1	16.7
R_{sym}(%)*	4.2	5.0	6.7	6.7
SIRAS phasing—AraC complexed with arabinose (20.0 to 1.8 Å)				
f(calc.) electrons			−14.4	
f″(calc.) electrons			10.13	
R_{iso}† (%)			24.7	
R_{cullis}‡			0.61/0.64§	
Mean figure of merit‖			0.50	
Phasing power¶			1.57/1.97§	
Refinement—AraC complexed with arabinose				
Resolution range (Å)		12 to 1.5		
R_{factor}/R_{free} (%)#**		17.9/23.2		
rmsd				
Bond length (Å)		0.010		
Bond angles (degrees)		1.50		
Molecular replacement solution—unliganded AraC				
Correlation coefficient (%)				63.3
R_{factor} (%)				37.3
Refinement–unliganded AraC				
Resolution range (Å)				8.0 to 2.8
R_{factor}/R_{free} (%)				21.8/33.8
rmsd				
Bond length (Å)				0.017
Bond angles (degrees)				1.95

*$R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma\langle I\rangle$, where I is the observed intensity, and $\langle I\rangle$ is the average intensity of multiple observations of symmetry-related reflections. †$R_{iso} = \Sigma|F_{ph} - F_p|/\Sigma F_p$, where F_{ph} and F_p are the observed derivative and native structure factor amplitudes. ‡$R_{cullis} = \Sigma\|F_{ph} \pm F_p| - F_{h(calc.)}|/\Sigma|F_{ph} - F_p|$ §Centric/acentric. ‖Mean figure of merit = $\langle|P(\alpha)e^{i\alpha}/\Sigma P(\alpha)|\rangle$, where α is the phase, and $P(\alpha)$ is the phase probability distribution. ¶Phasing power = $\{\Sigma|F_{ph(calc.)}|^2/\Sigma[|F_{ph(obs.)}| - |F_{p(calc.)}|]^2\}^{1/2}$. #$R_{factor} = \Sigma|F_p - F_{p(calc.)}|/\Sigma F_p$. **$R_{free}$ is the R factor for a subset of 10% of the reflection data that were not included in the crystallographic refinement.

inose and packs against the indole ring of Trp^{95} (Fig. 3B). In the arabinose-bound AraC molecule, Tyr^{31} is in a slightly different conformation and is solvent exposed. The position of Tyr^{31} in each arabinose-free monomer is stabilized by a hydrogen bond to Tyr^{82} from the opposite monomer, as well as by van der Waals interactions with Trp^{95}, Met^{42}, and Ile^{36}. Association of two AraC monomers by way of the β-barrel interface is further stabilized by hydrophobic interactions mediated by Val^{20}, Leu^{32}, Phe^{34}, and Ile^{46}.

The ability of unliganded AraC to engage simultaneously in both the β-barrel and the coiled-coil interactions in the crystal lattice suggests that, at elevated concentrations, both interactions should occur in solution. If they do, the uncomplexed protein should aggregate. In contrast, the sugar-bound form should not aggregate because in the presence of arabinose, the NH_2-terminal arms fold over the arabinose-binding pockets and block the β-barrel interaction, leaving the protein capable of coiled-coil dimerization only. We tested this prediction with velocity sedimentation of the AraC NH_2-terminal domain at 0.5 mg/ml (Fig. 4). The domain sediments in the absence of arabinose as a range of species with sedimentation coefficients in the range of 3.2 to 6S, consistent with the formation of dimers and higher ordered oligomers. In contrast, the domain in the presence of arabinose behaves as a homogeneous dimer, sedimenting at 2.9 to 3.1S and showing no signs of aggregation. Small-angle x-ray scattering data (*23*) also lead to the same conclusions.

The structural and solution studies of AraC presented here show that a new oligomerization interface is opened when arabinose is absent from its binding pocket and the NH_2-terminal arm of the protein consequently becomes disordered. Association by way of this interface is completely disrupted when arabinose binds within the sugar-binding pocket, causing occlusion of the β-barrel oligomerization interface by displacing the tyrosine residue inserted by the opposing monomer and triggering the folding of the NH_2-terminal arm. This mechanism of ligand-regulated oligomerization contrasts with that of another structurally characterized regulatory protein, *E. coli* arginine repressor (ArgR), whose oligomeric state is also controlled by ligand. In that case, arginine binding directly promotes the association of ArgR trimers to form a hexamer (*24*).

How might the arabinose-induced change in the ability of AraC to self-associate through the β-barrel interface alter the protein's DNA-binding properties? An intriguing possibility is that the AraC dimerization interface may change in response to the presence or absence of arabinose. By this mechanism, the protein would dimerize by the antiparallel coiled-coil in the presence of arabinose and, because AraC always functions as a dimer (*4*, *5*, *13*), by the more extensive β-barrel interface in the absence of arabinose.

In the β-barrel dimer, the DNA-binding domain attachment points to the COOH-terminus of the dimerization domain are 60 Å apart (Fig. 3A), whereas in the coiled-coil dimer the distance is 37 Å (Fig. 3B). The reduction in distance could make DNA looping energetically unfavorable in the presence of arabinose and simultaneously increase the affinity of the protein for adjacently located DNA half-sites, as is observed experimentally (*6*, *14*, *25*). The shorter distance between the DNA-binding domains in the liganded versus unliganded protein would also account for the observation that AraC in the absence of arabinose can bind to DNA with an additional 10 base pairs inserted between half-sites, as compared with AraC with bound arabinose (*14*). Unclear at this time is how, in the absence of arabinose, formation of higher ordered oligomers is inhibited because the coiled-coil interface is still available for interactions; however, it is possible that the structural rearrangements at the coiled-coil interface observed in the absence of arabinose could decrease its affinity.

Dissociation of arabinose from AraC, in addition to exposing a new oligomerization interface, also frees the NH_2-terminal arm. Although studies with protein chimeras do not indicate the presence of a significant interaction between the sugar-binding and dimerization domain and DNA-binding domain of AraC (*9*), it remains possible that the freed NH_2-terminal arm may modulate the effective distance or orientation between DNA-binding domains.

The regulation by arabinose of the NH_2-terminal arm's conformation and of the availability of the sugar-binding pocket for protein-protein interactions represents a versatile way to couple ligand binding with changes in conformation and multimeric state. Definitive determination of the mechanism by which the binding of arabinose changes the way AraC interacts with DNA awaits further testing by mutagenesis experiments, solution studies, and additional structural studies of the AraC protein.

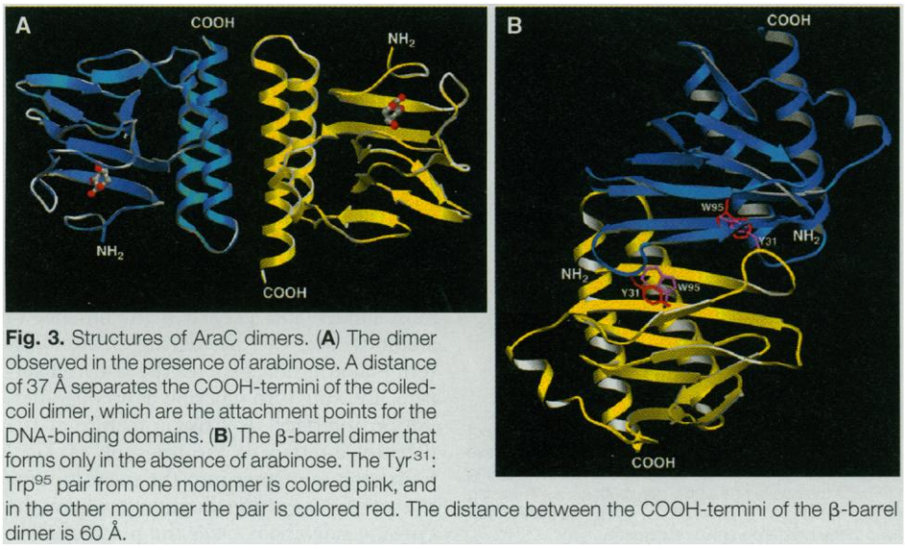

Fig. 3. Structures of AraC dimers. (**A**) The dimer observed in the presence of arabinose. A distance of 37 Å separates the COOH-termini of the coiled-coil dimer, which are the attachment points for the DNA-binding domains. (**B**) The β-barrel dimer that forms only in the absence of arabinose. The Tyr^{31}:Trp^{95} pair from one monomer is colored pink, and in the other monomer the pair is colored red. The distance between the COOH-termini of the β-barrel dimer is 60 Å.

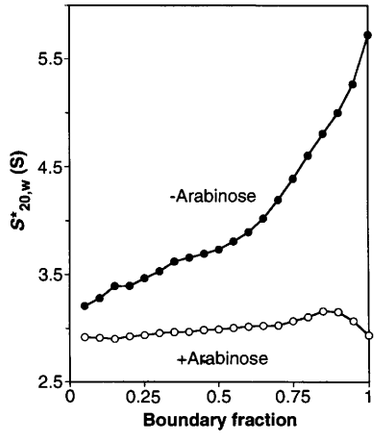

Fig. 4. Sedimentation velocity data for the AraC sugar-binding and dimerization domain in the presence (○) and absence (●) of arabinose. Samples were prepared in 15 mM tris-HCl (pH 7.5) and 75 mM KCl in the presence and absence of 0.2% (w/v) L-arabinose and analyzed at 42,000 rpm in a Beckman XL-A analytical ultracentrifuge. Van Holde–Weischet analysis of the resulting data (*36*) produced the distribution plot shown.

REFERENCES AND NOTES

1. R. Schleif, in Escherchia coli *and* Salmonella Typhimurium, F. Neidhardt *et al.*, Eds. (American Society for Microbiology, Washington, DC, 1996), pp. 1300–1309.
2. T. Dunn, S. Hahn, S. Ogden , R. Schleif, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5017 (1984).
3. K. Martin, L. Huo, R. Schleif, *ibid.* **83**, 3654 (1986).
4. R. Lobell and R. Schleif, *Science* **250**, 528 (1990).
5. W. Hendrickson and R. Schleif, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3129 (1985).
6. ———, *J. Mol. Biol.* **178**, 611 (1984).
7. X. Zhang, T. Reeder, R. Schleif, *ibid.* **258**, 14 (1996).
8. R. Schleif, *Annu. Rev. Biochem.* **61**, 199 (1992).
9. S. Bustos and R. Schleif, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5638 (1993).
10. R. Eustance, S. Bustos, R. Schleif, *J. Mol. Biol.* **242**, 330 (1994).
11. H. Lauble, Y. Georgalis, U. Heinemann, *Eur. J. Biochem.* **185**, 319 (1989).
12. R. Eustance and R. Schleif, *J. Bacteriol.* **178**, 7025 (1996).
13. G. Wilcox and P. Meuris, *Mol. Gen. Genet.* **145**, 97 (1976).
14. J. Carra and R. Schleif, *EMBO J.* **12**, 35 (1993).
15. Intact AraC was expressed in *E. coli* and purified [R. Schleif and A. Favreau, *Biochemistry* **21**, 778 (1982)]. After dialyzing purified protein into buffer A [15 mM tris-HCl (pH 8.0), 75 mM KCl, 0.2% (w/v) L-arabinose], we prepared a tryptic fragment comprising the NH_2-terminal two-thirds of the protein by digesting intact AraC overnight with 0.1% trypsin by mass. The cleavage mixture was concentrated by ultrafiltration, and the NH_2-terminal domain was purified by anion-exchange chromatography on a Mono-Q column (Pharmacia, Uppsala, Sweden) developed with a gradient of buffer A + 1 M KCl. Peak fractions were dialyzed into buffer A + 2 mM sodium azide, concentrated by ultrafiltration to 8 to 12 mg/ml, and stored at 4°C. Arabinose was added to a final concentration of 0.2% (w/v) immediately before cocrystallization by hanging-drop vapor diffusion. Amino-terminal sequencing and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry indicated that the tryptic fragment was >99% pure and contained residues 2 to 178 of AraC (S. Soisson, unpublished data). Small, blocky crystals appear infrequently with a reservoir solution of 30% (w/v) polyethylene glycol (PEG) 8000, 100 mM tris-HCl (pH 7.0), and 40 mM $MgCl_2$. Optimal results were obtained when crystals were grown by microseeding, with a reservoir solution containing 18% PEG 8000, 100 mM tris-HCl (pH 7.25), and 40 mM magnesium acetate. All cocrystals of AraC and arabinose were initially stabilized in a solution of 24% PEG 8000, 100 mM tris-HCl (pH 7.5), 40 mM magnesium acetate, and 0.2% (w/v) L-arabinose. For data collection, crystals were transferred to the above stabilizing solution plus 10% PEG 400 for 5 to 10 min and then flash-frozen in a small monofilament nylon loop placed in a cold nitrogen stream maintained at 100 K.
16. Crystals of the sugar-binding and dimerization domain of AraC in the absence of arabinose were grown by hanging-drop vapor diffusion with a reservoir solution of 20% PEG 4000, 0.1 M tris-HCl (pH 9.0), 5 mM KCl, and 0.2 M sodium acetate. Crystals were stabilized by sequential transfer into reservoir solutions containing increasing amounts of PEG 4000 in 4% increments (10 min in each step) until a concentration of 40% PEG 4000 was reached. Crystals were then directly flash-frozen in a 100 K nitrogen stream for data collection.
17. Simulated-annealing omit maps calculated with X-PLOR consistently showed the presence of only the α-anomer of L-arabinose in the AraC binding site. The equilibrium dissociation constant of *E. coli* AraC and arabinose is $\sim 10^{-3}$ M [G. Wilcox, *J. Biol. Chem.* **249**, 6892 (1974)].
18. Accessible surface area calculations were performed with X-PLOR, and probe sizes of 1.4, 1.5, and 1.6 Å gave similar results.
19. F. H. C. Crick, *Acta Crystallogr.* **6**, 689 (1953).
20. Modeled as a water, the molecule has a refined thermal B factor of 14.4 $Å^2$ and an occupancy of 1.0, as compared with an average B factor of 26.6 $Å^2$ for all water molecules. This does not exclude the possibility that the "water" is a sodium or potassium ion.
21. W. S. Somers and S. E. V. Phillips, *Nature* **359**, 387 (1992).
22. J. Janin and F. Rodier, *Proteins* **23**, 580 (1995).
23. S. M. Soisson, B. MacDougall-Shackleton, R. Schleif, C. Wolberger, data not shown.
24. G. D. V. Duyne, G. Ghosh, W. K. Maas, P. B. Sigler, *J. Mol. Biol.* **256**, 377 (1996).
25. N. Lee, C. Francklyn, E. Hamilton, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8814 (1987).
26. Z. Otwinowski, Oscillation Data Reduction Program, L. Sawyer, N. Isaacs, S. Bailey, Eds., *Proceedings of the CCP4 Study Weekend: Data Collection and Processing* (SERC Daresbury Laboratory, 1993).
27. Collaborative Computational Project Number 4, *Acta Crystallogr.* **D50**, 760 (1994).
28. T. A. Jones, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, *ibid.* **A47**, 110 (1991).
29. A. T. Brünger, *X-PLOR v3.1, A System for X-Ray Crystallography and NMR, Manual* (Yale Univ. Press, New Haven, CT, 1992).
30. After two rounds of simulated annealing and positional refinement in X-PLOR, one molecule of α-L-arabinose per monomer was added. The model was completed by alternating rounds of Powell positional refinement in X-PLOR followed by manual adjustments. The rmsd for superimposed monomers (A and B) of AraC in the asymmetric unit is 0.39 Å. Waters were added in shells only to positive $F_o - F_c$ peaks greater than 3σ and that form at least one potential hydrogen bond with the protein or with other waters already hydrogen-bonded to the protein. Occupancies of the waters were refined during the last stages of refinement, and alternate conformations were included for residues Glu^{27}, Ile^{36}, Leu^{47}, Ser^{112}, and Leu^{151} of monomer A and Ile^{36} and Leu^{47} of monomer B. All waters in the refined model have B factors of less than 50 $Å^2$.
31. J. G. Navaza, *Acta Crystallogr.* **A50**, 157 (1994).
32. A. T. Brünger, personal communication.
33. L. M. Rice and A. T. Brünger, *Proteins* **19**, 277 (1994).
34. S. V. Evans, *J. Mol. Graph.* **11**, 134 (1993).
35. The electron density for the NH_2-terminal arm of AraC is better defined in monomer B; therefore, that monomer was used for generation of all sugar-binding figures.
36. K. E. van Holde and W. O. Weischet, *Biopolymers* **17**, 1387 (1978).

37. We thank A. Batchelor, E. Reisinger, J. Aishima, and M. Greisman for help with data collection; C. Ogata and R. Abramowitz of beamline X4A at the National Synchrotron Light Source for advice and technical support; and J. Berg, M. Amzel, and E. Lattman for comments on the manuscript. C. Becker, M. Nagypal, S. Jenkins, A. Favreau, M. Williams, and J. Withey helped with purification and crystallization efforts. We thank C. Turgeon and J. Hansen of the University of Texas Health Sciences Center at San Antonio for analyzing samples by analytical ultracentrifugation and A. Brünger and L. Rice for a prerelease version of X-PLOR for torsion angle dynamics refinement. Supported by the Howard Hughes Medical Institute (C.W. and beamline X4A), the David and Lucile Packard Foundation (C.W.), and grant GM18277 from the National Institutes of Health (R.S.). Atomic coordinates have been deposited at the Brookhaven Protein Data Bank (accession numbers 2ARC and 2ARA).

16 October 1996; accepted 18 February 1997

Quantitative Trait Loci for Refractoriness of *Anopheles gambiae* to *Plasmodium cynomolgi* B

Liangbiao Zheng, Anton J. Cornel, Rui Wang, Holger Erfle, Hartmut Voss, Wilhelm Ansorge, Fotis C. Kafatos, Frank H. Collins*

The severity of the malaria pandemic in the tropics is aggravated by the ongoing spread of parasite resistance to antimalarial drugs and mosquito resistance to insecticides. A strain of *Anopheles gambiae*, normally a major vector for human malaria in Africa, can encapsulate and kill the malaria parasites within a melanin-rich capsule in the mosquito midgut. Genetic mapping revealed one major and two minor quantitative trait *loci* (QTLs) for this encapsulation reaction. Understanding such antiparasite mechanisms in mosquitoes may lead to new strategies for malaria control.

Melanotic encapsulation, an immune reaction in which invading parasites are enclosed and destroyed within a melanin-rich capsule, is widespread among insects. Malaria parasites, which must develop into oocysts in the mosquito midgut, can also be encapsulated in some refractory vector strains, resulting in a block to disease transmission (*1*). The mechanism of parasite rejection is a key to the biology of interaction between *Plasmodium* and its vector, and an understanding of this mechanism may ultimately be useful in malaria control strategies such as mosquito population replacement using robust refractory strains.

Fully refractory and susceptible strains of *A. gambiae* have been selected for the ability to encapsulate or tolerate, respectively, oocysts of *Plasmodium cynomolgi*, a simian parasite. These strains respond similarly to most *Plasmodium* species, including the human pathogen *P. falciparum* (*1*). Many dif-

L. Zheng, R. Wang, H. Erfle, H. Voss, W. Ansorge, F. C. Kafatos, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany.
A. J. Cornel and F. H. Collins, Centers for Disease Control and Prevention, 4770 Buford Highway, Mail Stop F22, Chamblee, GA 30341, USA.

*To whom correspondence should be addressed. E-mail: fhc1@cdc.gov