Nucleic Acids Research, 1993, Vol. 21, No. 3 435–438

Formation of AraC-DNA sandwiches

John H.Carra and Robert F.Schleif*
Biology Department, The Johns Hopkins University, 34th and Charles Streets, Baltimore, MD 21218, USA

Received November 12, 1992; Revised and Accepted December 28, 1992

ABSTRACT

With the use of special DNA binding sites, but not the natural *araI* binding site, the dimeric AraC protein can be forced to make sandwich structures in which two DNA molecules are joined by two AraC protein dimers. Apparently one subunit from each dimer contacts each DNA molecule in an extended structure. These sandwich structures form only in the absence of arabinose. This behavior is consistent with the protein's ability to form DNA loops by binding to separated half sites in the absence of arabinose and its preference for binding to adjacent half-sites in the presence of arabinose.

INTRODUCTION

The AraC protein of *E.coli* is a DNA-binding regulator of transcription which controls expression of genes involved in the catabolism of arabinose (1). At the *araCBAD* regulatory region, in the absence of arabinose, AraC protein forms a DNA loop by binding to both the $araI_1$ half site and the $araO_2$ half site (Fig.1), which are separated by 210 base pairs (2–4). One subunit of the dimeric protein contacts $araI_1$ and the other subunit contacts $araO_2$ (4). When arabinose is added, AraC protein shifts the location of its DNA contacts without dissociation from the DNA, now binding to the adjacently located $araI_1$ and $araI_2$ half sites (4). Induction of transcription from the promoter P_{BAD} overlapping $araI_2$ follows. The arabinose-bound form of AraC protein is apparently restricted to the recognition of closely spaced pairs of half-sites (5). The binding of AraC protein to the *araI* site has been studied extensively (5–8), and can be measured by an electrophoretic retardation assay in which the AraC-*araI* complex migrates as a single band of slower mobility than the unbound DNA.

Upon the discovery of DNA looping in the *araCBAD* system, vigorous attempts were made to demonstrate the joining of two DNA molecules by AraC protein, but without success (9). It was postulated that if AraC protein can bind simultaneously to the well-separated *araI* and $araO_2$ sites on one DNA molecule, then it should also be able to join two sites on separate DNA molecules, under some conditions. Such structures, where a protein links two DNA molecules, have been called DNA sandwiches by Müller-Hill and colleagues and have been demonstrated in the case of *lac* repressor (10), a tetrameric protein which forms loops (10–15). It is now apparent that the AraC protein does not readily form DNA sandwiches because it is a dimer, and unlike the tetravalent *lac* repressor, a divalent dimer bound to two half-sites on the same DNA molecule cannot join two DNA molecules to form a sandwich structure. On the natural *araI* site, binding of a protein dimer to adjacent half-sites on the same molecule is favored over binding to half-sites on two separate DNA molecules. Sandwiches do not form from two DNA molecules with only one half-site each, probably because a sufficient number of protein–DNA contacts cannot be made to stabilize the structure. Herein we show however, that sandwiches can be made if both subunits of the AraC dimer can be prevented from binding to two half-sites on the same DNA molecule. We synthesized an artificial AraC binding site containing two half-sites angularly separated such that a dimer of AraC protein cannot easily contact both on the same DNA molecule. As shown in Fig. 2, sandwiches might then form that contain two dimers connecting two DNA molecules. This work demonstrates the formation of such structures, which are analogous to DNA loops in that they involve the protein binding to distally located half-sites, and are destabilized by the presence of arabinose.

MATERIALS AND METHODS

Plasmids and binding sites

DNA oligonucleotides containing various binding sites for AraC protein were synthesized and cloned as (5) into the plasmid pES51 (16), replacing the *araI* site in the context of the *ara* P_{BAD} regulatory region. Ligation, transformation, and sequencing of recombinant clones was accomplished by standard DNA manipulation techniques (17). The sequences of the constructed mutant sites are shown below, with each I_1 half-site in bold. A block of A_5 bases in the spacer region of the I_1-*14*-I_1 site is underlined. Both strands of the evI_1-I_1 site are shown for clarity.

I_1-*14*-I_1
TAGCAT TTTT ATCCATA GATC GATCTG AC<u>AAAAA</u>T
TAGCAT TTTT ATCCATA

evI_1-I_1
TATGGAT AAAA ATGCTA GATC **TAGCAT TTTT ATCCATA ATACCTA TTTT TACGAT** CTAG ATCGTA AAAA TAGGTAT

* To whom correspondence should be addressed

Gel retardation assays and DNase footprinting

DNA molecules for use in the assays performed in this study were created *in vitro* using the polymerase chain reaction, PCR (5,18). Gel retardation assays performed on DNA sandwich complexes were done essentially as Hendrickson and Schleif (6), except that the concentrations of acrylamide and methylene-bis-acrylamide used in the gels were changed from 6% and 0.1% to 4% and 0.067% (17). Pure AraC protein was added to approximately 0.2 pmoles 5′-^{32}P labeled I_1-14-I_1 site DNA in the binding buffer of Hendrickson and Schleif (6) with 50 mM KCl, no arabinose, and 0.5 μg calf thymus DNA (non-specific competitor) in a reaction volume of 20 μl. After a 90′ incubation at 37°C, the reaction was loaded on a retardation assay gel. DNase footprinting was done as described (5,7).

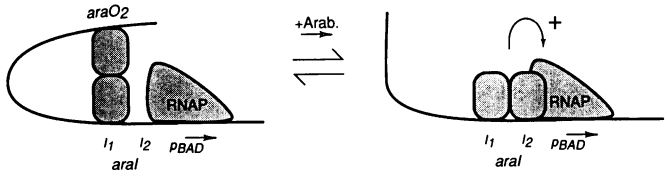

Figure 1. AraC protein forms a DNA loop in the absence of arabinose, and activates RNA polymerase in the presence of arabinose.

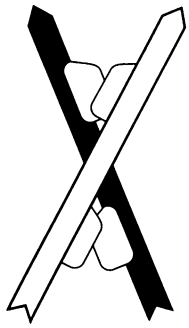

Figure 2. Diagram of a hypothetical AraC-DNA sandwich structure.

RESULTS

AraC-DNA sandwiches form on the I_1-14-I_1 site

At the natural *araI* site, the two 17 base pair half-sites $araI_1$ and $araI_2$ are present in a direct repeat orientation and are separated by 4 base pairs (5,8). Thus, corresponding positions of the two half-sites are separated by 21 base pairs, two helical turns of the DNA, and the protein-DNA contacts between AraC and *araI* can all be made from one face of the DNA double helix. When the half-sites are placed on opposite faces of the DNA, the two subunits of an AraC protein dimer cannot both bind simultaneously bind to both half-sites (5). Here we describe properties of an artificial binding site called I_1-14-I_1 that consists of two I_1 half-sites separated by 14 more bases than separate I_1 and I_2 in the normal *araI* site. This places the two I_1 half-sites about 100 degrees out of helical phase. Additionally, the spacing nucleotides between the two I_1 half-sites contains a stretch of 5 adenines followed by a thymine, a sequence which is naturally bent (19).

Titration of AraC protein against the I_1-14-I_1 site revealed that this site forms several different protein-DNA complexes (Fig. 3a). On the basis of their protein concentration dependencies, their DNase footprints, and the pathway of dissociation of complexes, we identify one of these as containing two AraC dimers per DNA molecule, P_2D, and the other as containing one dimer per DNA, PD. The DNase footprint of the P_2D complex shows complete occupancy of both I_1 half-sites by AraC protein (data not shown). In the P_2D complex, one subunit of each protein dimer specifically contacts one I_1 half-site, while the other presumably interacts with adjacent non-specific DNA. In experiments assaying the dissociation of protein-DNA complexes in the presence of an excess of unlabeled competitor DNA, the P_2D complex was observed to dissociate through the

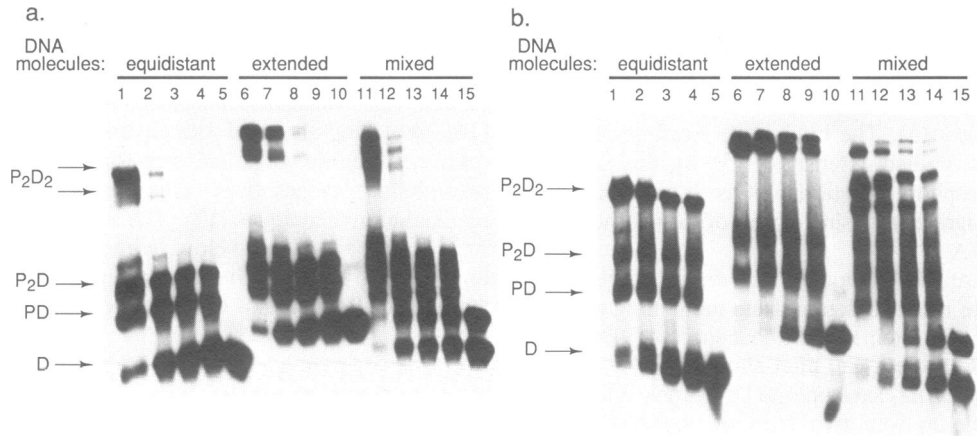

Fig. 3. a. Bandshift assay showing protein−DNA complexes formed on the I_1-14-I_1 site. Lanes 1−5 contained DNA molecules with the I_1-14-I_1 site equidistant from the ends of the molecule. Lanes 6−10 contained DNA molecules with one arm extended by 86 base-pairs. Lanes 11−15 contained an equimolar mixture of the two. Approximate amounts of AraC protein (dimeric) in pmoles added were as follows: Lane 1: 0.4, 2: 0.25, 3: 0.2, 4: 0.15, 5: 0, 6: 0.4, 7: 0.25, 8: 0.2, 9: 0.15, 10: 0, 11: 0.4, 12: 0.25, 13: 0.2, 14: 0.15, 15: 0. **b.** Bandshift assay showing protein−DNA complexes formed on the evI_1-I_1 site. The DNA molecules and protein concentrations used were the same as in Fig. 3a, except for the use of the evI_1-I_1 site.

PD complex as an intermediate (data not shown). The ratios of the P_2D, PD, and D complexes as protein concentration is varied are also consistent with independent binding of protein dimers to each half-site. Previous experiments have shown that AraC can bind cooperatively to half-sites spaced one or two helical turns apart (5), when angular phasing is maintained. An angular separation of 100 degrees is enough then to prevent an AraC dimer from being able to bind simultaneously to two half-sites.

At much higher protein concentrations, approximately 2×10^{-8} M, and relatively high DNA concentrations, 1×10^{-8} M, an additional two bands of very slow mobility were observed in the gel retardation assay (Fig. 3a lane 2). These bands are found only in the absence of arabinose. In the presence of arabinose we observed only the P_2D and PD species. The migration rates of the two very slow bands, as well as their dependence on high protein and DNA concentrations, are consistent with their containing DNA sandwich structures (10).

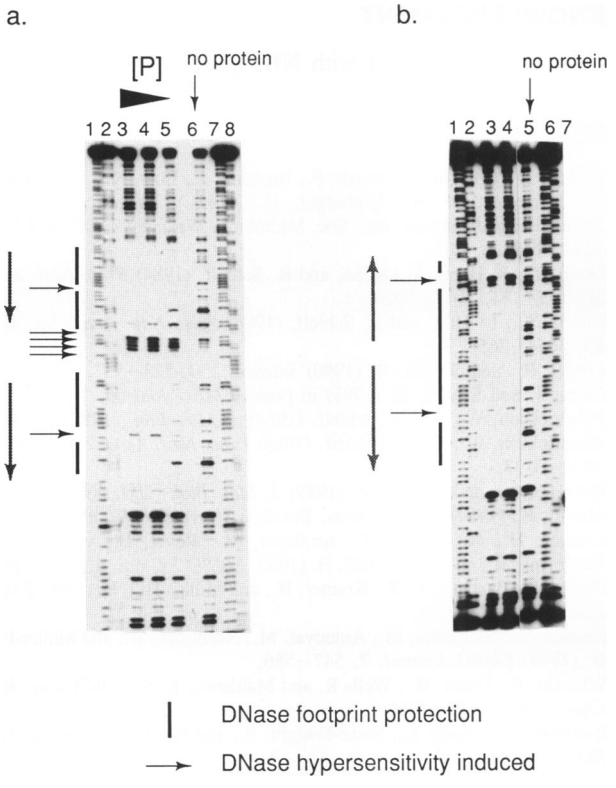

Figure 4. a. DNase footprint of AraC-DNA sandwiches formed on the I_1-14-I_1 site. Pure AraC protein was added to approximately 1 pmole 5'-^{32}P labeled I_1-14-I_1 site DNA in the binding buffer of Hendrickson and Schleif (16) with 50 mM KCl, 2.5 mM $MgCl_2$, 1 mM $CaCl_2$, no arabinose, and 3.5 μg calf thymus DNA (non-specific competitor) in a reaction volume of 150 μl. After a 90' incubation at 37°C, 5 μl of the reaction was loaded on a retardation assay gel (not shown), and the rest treated with 5 ng DNase I for 30 seconds. The retardation assay gel revealed that about 90% of the DNA in lane 3 was in the P_2D_2 complex. Lanes 1 and 7 are G+A sequencing reactions and lanes 2 and 8 are C+T sequencing reactions. The bottom strand (as written in Materials and Methods) was labeled. Approximate amounts of AraC protein (dimeric) in pmoles added were as follows: lane 3: 6, 4: 3, 5: 1, 6: 0. **b.** DNase footprint of AraC-DNA sandwiches formed on the evI_1-I_1 site. Done as in part a, except that AraC protein concentrations in pmoles were as follows: lane 3: 6, 4: 3, 5: 0. Lanes 1 and 6 are G+A sequencing reactions and lanes 2 and 7 are C+T sequencing reactions. A retardation assay gel revealed that about 80% of the DNA in lane 3 was in the P_2D_2 complex.

AraC-DNA sandwich isomers

Modeling indicates that the most plausible AraC-DNA sandwich structures have the DNA between the two half-sites of the I_1-14-I_1 site somewhat bent and twisted. Some bending of this DNA may be expected because of the presence of an A_5T tract. DNase footprinting of the sandwiches, Fig. 4a lanes 3 and 4, shows several points of induced DNase hypersensitivity between half-sites, also suggesting that the structure of this DNA has been further distorted when the complex forms. Two different conformational isomers of sandwiches are shown in Fig. 5, a and b. These we propose are the molecular species in the observed bands. In the *cis* isomer the two DNA molecules bend into the same directions in space. In the *trans* isomer they bend into opposite directions. These two conformers have very different shapes and therefore could well possess different electrophoretic mobilities.

Other potential isomers of sandwiches are structures in which the directly repeated half-sites on each molecule are orientated parallel or anti-parallel, Fig. 5, c and d. These two structures are roughly equivalent only if the DNA sequences on either side of the half-sites are equal in length. Fig. 3 shows the result of gel retardation assays in which in which the arm lengths of the DNA molecules were made unequal. The DNA used in the experiment was either I_1-14-I_1 containing the AraC binding site equidistant from the ends, non-equidistant with a downstream extension of 86 bp, or a mixture of the two. Two sandwich bands (*cis* and *trans*) are observed in the first set of lanes (1−5), but these two bands split into four when instead the extended DNA molecules are used (lanes 6−10). This result is consistent with the formation of parallel and antiparallel isomers. When the two different size DNA molecules are mixed in the binding reaction (lanes 11−15), new bands appear, of intermediate mobility compared to the other complexes. These bands apparently represent some of the eight possible complexes having one DNA molecule of each type, in different relative orientations.

AraC-DNA sandwiches form on the evI_1-I_1 site

Sandwich structures are also formed using DNA containing two $araI_1$ half-sites in a tail-to-tail, or everted orientation, evI_1-I_1 (Fig. 3b, lanes 1−5). By placing four bases between such everted I_1 half-sites, the known contacts between AraC and the DNA are retained on one face of the DNA (5). The ratios of the P_2D, PD and D species formed with evI_1-I_1 indicate that, on this site

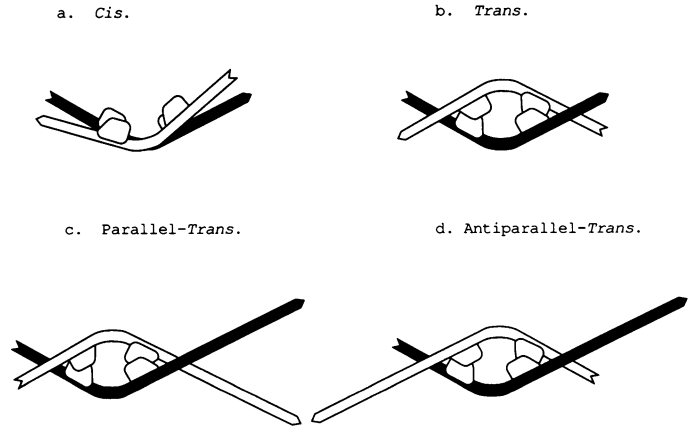

Figure 5. Diagrammatized structures of possible AraC-DNA sandwiches.

also, AraC protein dimers bind independently to each half-site. This non-cooperativity may be a consequence of unfavorable steric interactions arising from the close juxtaposition of half-sites in opposite orientations. The DNase footprint of the evI_1-I_1 site sandwiches (Fig. 4b) shows complete occupancy of both half-sites and no DNase hypersensitivities between half-sites. Because the evI_1-I_1 DNA has no A_5T tract and is not expected to be bent between half-sites, the *cis* and *trans* isomers of sandwiches should not be observed. In fact only one sandwich band is observed with this site when DNA molecules of equal arm lengths are used (Fig. 4b lanes 1–5). As with the I_1-*14*-I_1 site, use of DNA molecules of unequal arm lengths doubles the number of observed sandwich bands (Fig. 4b lanes 6–10), again consistent with the formation of parallel and antiparallel isomers. In addition, mixture of the two types of DNA molecules (Fig. 4b lanes 11–15), again yields more bands of intermediate mobility.

DISCUSSION

We have observed AraC protein to form DNA sandwich complexes on two artificial binding sites. The use of asymmetric DNA molecules of different types in the gel retardation assay allowed demonstration of the presence of different conformational isomers of these species. While DNA sandwiches are not known to form on *ara* regulatory sequences *in vivo*, study of these structures may provide useful analogies for the study of other phenomena involving the joining of two DNA molecules by proteins, including DNA recombination (20) and transvection (21).

DNA–protein–DNA sandwiches have been well studied with components from the *lac* operon (10). These sandwiches were observed to have slow electrophoretic migration rates as well as to exist in parallel and antiparallel isomers, which was confirmed by electron microscopy. Sandwich formation in the *lac* system differs in a fundamental way from that in the *ara* system. The tetrameric *lac* repressor contacts its operator binding site on one DNA molecule using two subunits. The other two subunits of the protein are available for binding to another DNA molecule and forming a sandwich. Thus in the *lac* system binding to adjacent half-sites is not in competition with DNA looping or sandwich formation as it is in the *ara* system. Formation of *ara* sandwiches therefore requires the destabilization of dimeric binding to adjacent half-sites, which we have shown can be achieved.

DNase footprinting of sandwiches formed on the *ara* I_1-*14*-I_1 site showed that each half-site is fully occupied by AraC protein. While we have not directly measured the stiochiometry of the sandwich complexes, the facts that these complexes are dependent upon very high protein and DNA concentrations, that mixing of two DNA molecules in one reaction yields new species, and that all four half-sites are fully protected from DNase, indicate that the stiochiometry of these complexes is very likely P_2D_2.

The DNase footprint of the I_1-*14*-I_1 sandwich complexes (Fig. 4a) also showed several induced DNase hypersensitivities in the region between the half-sites. Model building suggests that the DNA between the half-sites should be slightly distorted to best bring the half-sites together in space. Such distortion of the DNA structure in the spacer likely is the cause of the observed DNase hypersensitivities between half-sites. The protein–DNA complexes formed on this site in the presence of arabinose, P_2D and PD, show no DNase hypersensitivities between half-sites, consistent with the result from the gel retardation assay (data not shown) that the absence of arabinose is required for sandwich formation.

Formation of the parallel and antiparallel isomers drawn in Fig. 5, c and d, necessitate that AraC protein possess the ability to bind to DNA with the DNA binding domains of a dimer orientated in two different directions in space. Such flexibility has been observed (5), as AraC protein can bind with high affinity to DNA with the binding half-sites in direct repeat orientation or inverted repeat orientation.

Previous experiments show that *in vivo* and *in vitro* AraC protein will bind to well-separated half-sites to form DNA loops (2–5,16) in the absence of arabinose. The addition of arabinose to the $araI_1$-$araO_2$ loop, causes the loop to open and AraC protein to bind to the two adjacent half-sites $araI_1$ and $araI_2$ (4). Arabinose prevented the formation of sandwich structures. Thus, the destabilization of sandwiches by arabinose may proceed by the same mechanism as the opening of DNA loops: a conformational change in the protein that causes it to prefer binding to adjacent half-sites.

ACKNOWLEDGEMENT

This work was supported with NIH grant GM18277.

REFERENCES

1. Schleif, R. (1987) In Neidhardt, F., Ingrahm, J., Low, K.B., Magasanik, B., Schaecter, M. and Umbarger, H.E. (eds.), *Escherichia coli and Salmonella typhimurium*, Am. Soc. Microbiol., Washington DC, Vol 2, pp. 1473–1481.
2. Dunn, T., S. Hahn, S. Ogden, and R. Schleif. (1984) *Proc. Natl. Acad. Sci. USA*, 81, 5017–5020.
3. Martin, K., L. Huo, and R. Schleif. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3654–3658.
4. Lobell, R., and Schleif, R. (1990) *Science*, 250, 528–532.
5. Carra, J. and Schleif, R. (1993) in press, *EMBO Journal*.
6. Hendrickson, W., and R. Schleif. (1984) *J. Mol. Biol.*, **178**, 611–628.
7. Hendrickson, W., and R. Schleif. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3129–3133.
8. Brunelle, A., and Schleif, R. (1989) *J. Mol. Biol.*, 209, 607–622.
9. Martin, K. (1987) Ph. D. thesis, Brandeis University, Waltham, MA.
10. Kramer, H., Niemoller, M., Amouyal, M., Revet, B., von Wilcken-Bergmann, B., and Müller-Hill, B. (1987) *EMBO Journal*, **6**, 1481–1491.
11. Oehler, S., Eismann, E. R., Kramer, H., and Müller-Hill, B. (1990) *EMBO Journal*, **9**, 973–979.
12. Kramer, H., Niemoller, M., Amouyal, M., Nordheim, A., and Müller-Hill, B. (1988) *EMBO Journal*, **7**, 547–556.
13. Whitson, P., Hsieh, W., Wells R. and Matthews, K. S. (1987) *Jour. Biol. Chem.*, **262**, 4943–4946.
14. Borowiec, J., Zhang, L., Sasse-Dwight, S., and Gralla, J. (1987) *J. Mol. Biol.*, **196**, 101–111.
15. Cann, J. R. (1990) *J. Mol. Biol.*, **216**, 1067–1075.
16. Huo, L., Martin, K. and Schleif, R. (1988) *Proc. Natl. Acad. Sci. USA*, 85, 5444–5448.
17. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor University Press, Cold Spring Harbor.
18. Carra, J. (1992) Ph. D. thesis, Brandeis University, Waltham, MA.
19. Wu, H. M., and Crothers, D. M. (1984) *Nature*, **308**, 509–513.
20. Kim, S., and Landy, A. (1992) *Science*, **256**, 198–203.
21. Bickel, S., and Pirrotta, V. (1990) *EMBO Journal*, **9**, 2959–2967.