Regulation of the L-arabinose operon in *Escherichia coli*

Robert Schleif

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

GENERAL BACKGROUND

Escherichia coli can grow on L-arabinose as a source of carbon and energy (Fig. 1) (Schleif, 1996). Not surprisingly, because the amounts of enzymes required to convert arabinose into a component of the pentose phosphate shunt are significant, the levels of the enzymes are regulated. In the absence of arabinose, the uninduced or basal level of expression of the proteins is about 1/300 the induced level. It is a product of one of the genes of the arabinose operon, the AraC protein, that controls the expression level and hence is a sensor of the presence of arabinose. AraC protein transduces information about the sugar's presence into an induced synthesis rate of the AraE, AraF and AraG proteins, which are required for arabinose uptake, as well as the AraB, AraA and AraD enzymes, which are required for the catabolism of arabinose.

Genetic analysis of the arabinose operon was begun more than 40 years ago by Ellis Englesberg (Gross & Englesberg, 1959). In a series of genetic experiments of increasing rigour and elegance, Englesberg and his collaborators then provided rather convincing evidence that the primary activity of AraC protein was in inducing the expression of the other arabinose specific proteins (Sheppard & Englesberg, 1967; Englesberg *et al.*, 1965, 1969b; Gielow *et al.*, 1971); that is, AraC acted positively to turn on expression rather than acting negatively to turn off expression like the *lac* repressor turns off expression of the *lac* operon. That means that the intrinsic set state of the *ara*-specific promoters is off and AraC turns them on, whereas the set state of the *lac* operon promoter is on and the *lac* repressor turns it off. Subsequently, definitive biochemical

SGM symposium 61: Signals, switches, regulons and cascades: control of bacterial gene expression.
Editors D. A. Hodgson, C. M. Thomas. Cambridge University Press. ISBN 0 521 81388 3

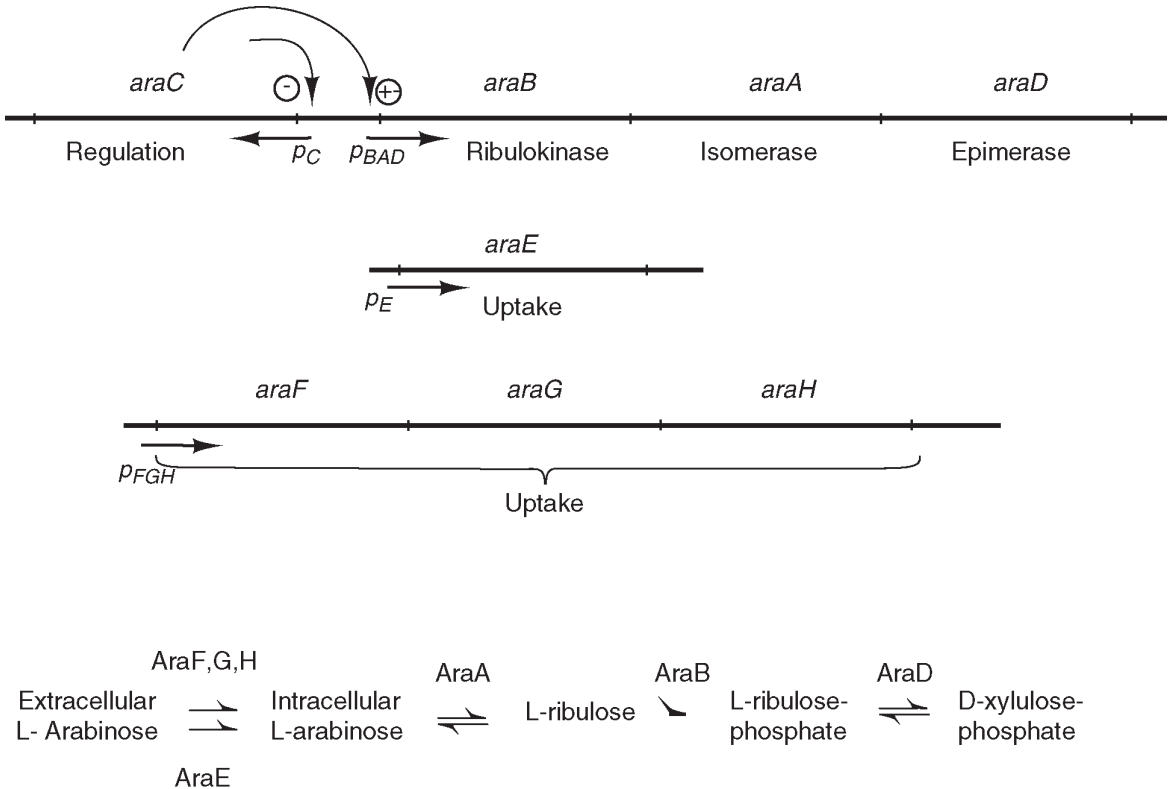

Fig. 1. Organization of the genes of the L-arabinose operon of *Escherichia coli*. The *araCBAD* gene cluster is at 1.45 min on the genetic map, *araE* is at 64.2 min, and *araFGH* is at 42.7 min. Also shown is the metabolic pathway for conversion of L-arabinose to D-xylulose 5-phosphate.

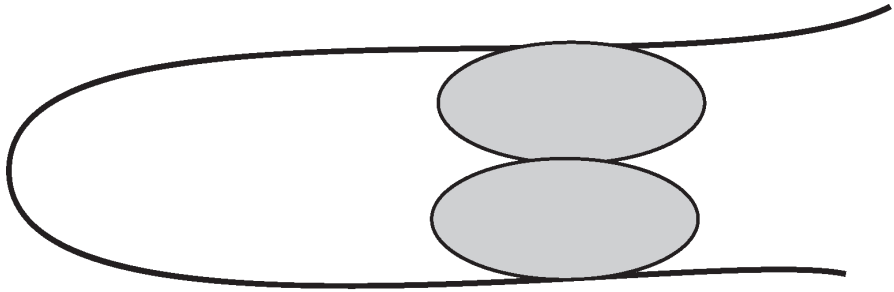

Fig. 2. Schematic diagram of DNA looping where a protein or proteins simultaneously binds to two well-separated sites on DNA and generate a DNA loop.

experiments proved that, indeed, AraC acted positively to turn on expression of the promoter that serves the *araB*, *araA* and *araD* genes, p_{BAD} (Greenblatt & Schleif, 1971).

Englesberg's genetic studies also revealed a secondary activity of AraC protein, that of acting negatively to repress the p_{BAD} promoter (Englesberg *et al.*, 1969a). Due to the implausibility of this negative activity, which by Englesberg's experiments appeared to act from upstream of the promoter, an extensive deletion screen was carried out to confirm the existence of the repressive phenomenon (Schleif, 1972; Schleif & Lis, 1975). This work indicated not only that repression existed, but that the site required for repression lay at least several hundred nucleotides upstream from the promoter. These experiments to verify the existence of the repression from upstream then led to the discovery of DNA looping (Dunn *et al.*, 1984). In looping, AraC protein simultaneously binds to two half-sites that are separated on the DNA by more than 200 bp (Fig. 2). Consequently, a loop is formed in the DNA. The discovery and demonstration of DNA looping came just in time to explain the puzzling properties posed by eukaryotic enhancer elements whose action at a distance properties had just been discovered. DNA looping provided a simple mechanism by which enhancers could act from sites located hundreds or thousands of nucleotides away to influence transcription from a promoter.

MECHANISMS OF REGULATION, DNA LOOPING AND LIGHT SWITCH RESPONSE TO ARABINOSE

Much work has gone into the study of DNA looping and into the mechanism by which AraC protein activates transcription from the *ara* p_{BAD} promoter. A few of the highlights of the earlier work on DNA looping were the development of *in vivo* footprinting and the demonstration of cooperativity in AraC protein binding to the two well-separated DNA-binding sites (Martin *et al.*, 1986), the subsequent discovery of DNA looping in other prokaryotic systems (Dandanell & Hammer, 1985; Eismann *et al.*,

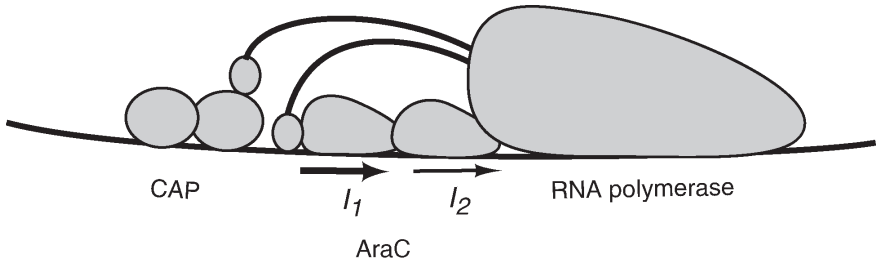

Fig. 3. Interactions thought to occur at the *ara* p_{BAD} promoter between RNA polymerase and AraC protein and CAP protein.

1987), the measurement of the *in vivo* helical pitch of DNA (Lee & Schleif, 1989), where it was found to be a little over 11 bp per turn, size limits on the DNA loop in the *ara* system (Lee & Schleif, 1989), and the construction of an *in vitro* system displaying DNA looping and unlooping caused by the presence of arabinose (Lobell & Schleif, 1990). More recently, AraC protein has been shown, by a gel electrophoretic assay that was developed for the study of multi-protein transcription complexes (Zhang *et al.*, 1996), to stimulate transcription by both assisting the binding of RNA polymerase to the p_{BAD} promoter and speeding the isomerization of RNA polymerase to an open complex. Also, the energetics of AraC protein binding to its half-sites of slightly different sequences and its energetic preference for looping DNA when there is no arabinose present, and its preference for binding to two adjacent sites when arabinose is present have been described thermodynamically (Seabold & Schleif, 1998). One of the two C-terminal domains on the two alpha subunits of RNA polymerase appears to bind to the polymerase distal subunit of AraC and the second likely binds to the cAMP receptor protein (Fig. 3). Mutations have been found in CAP that identify the contact region used by CAP when stimulating p_{BAD} as well as the promoter that is activated by the AraC homologue RhaR (Zhang & Schleif, 1998).

Much of our current understanding of the mechanism of regulation of the arabinose operon in *E. coli* is summarized in Fig. 4, which shows the light switch mechanism that lies at the heart of the protein's arabinose response (Saviola *et al.*, 1998). A monomer of the homodimeric AraC protein consists of a dimerization domain that also binds arabinose. This is loosely connected to a DNA-binding domain that also interacts with RNA polymerase to activate transcription. In the absence of arabinose, AraC protein loops the DNA by binding to the I_1 and O_2 DNA half-sites that are located 210 bp apart. This loop both restricts RNA polymerase access to the p_C and p_{BAD} promoters as well as keeps a DNA-binding domain of AraC from binding to the I_2 half-site. When AraC is bound in this state, the p_{BAD} is off. When arabinose is present, however, p_{BAD} is on.

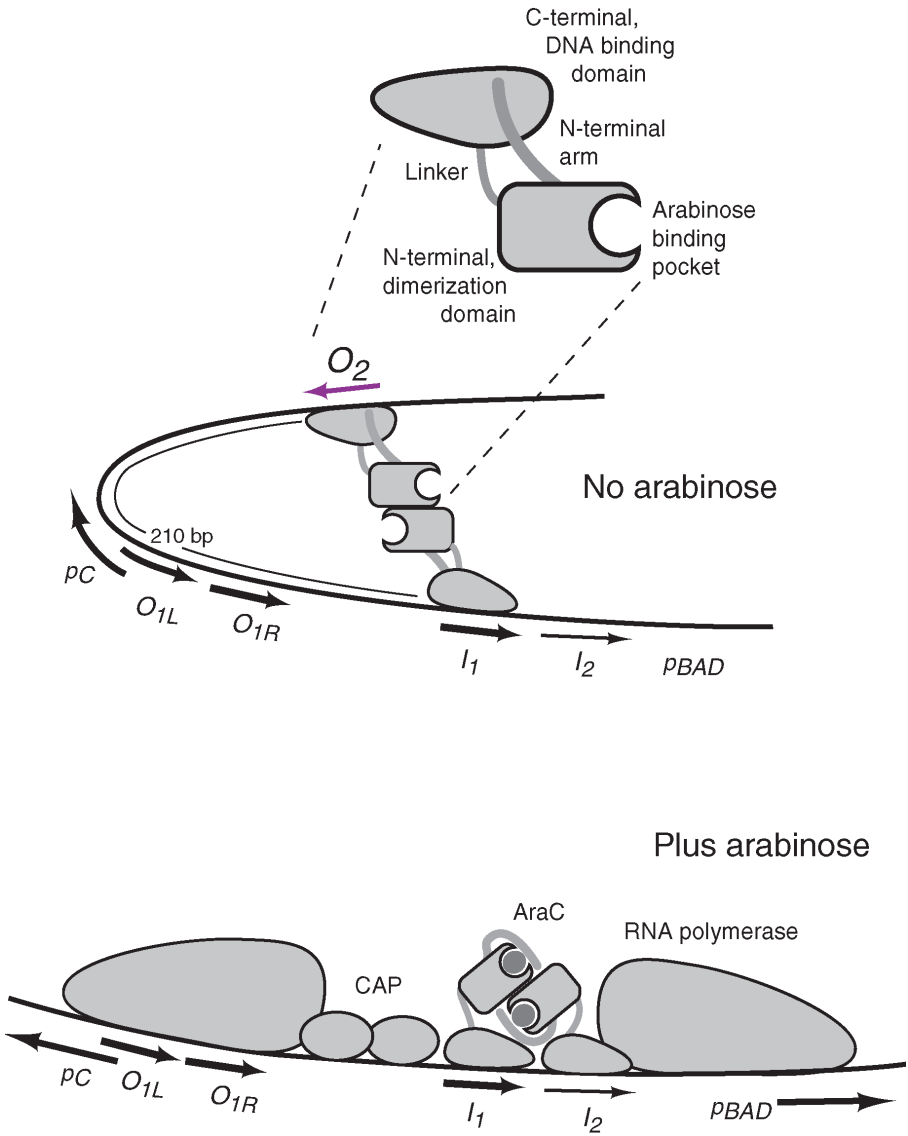

Fig. 4. Light switch mechanism of AraC protein action. In the absence of arabinose, the N-terminal arm interacts with the DNA-binding domains and holds the latter such that the protein loops between *araO*$_2$ and *araI*$_1$. This represses both p_C and p_{BAD}. In the presence of arabinose, interactions between arabinose and the N-terminal arms hold the latter over the arabinose, thus freeing the DNA-binding domains. These domains are then free to bind to I_1 and I_2, which then leads to induction of the p_{BAD} promoter. The binding of AraC to O_{1L} and O_{1R}, left and right, directly represses the p_C.

AraC then prefers to bind to the adjacent I_1 and I_2 half-sites. Binding to the I_2 half-site probably results in direct interactions with RNA polymerase, whereas RNA polymerase also contacts the DNA-binding domain bound at I_1 by reaching over the polymerase proximal subunit of AraC. Hence induction ensues because AraC is no longer looping the DNA and because AraC actively promotes transcription by its interactions with RNA polymerase.

AraC protein prefers to loop in the absence of arabinose because an N-terminal arm of about 15 amino acids extends from the dimerization domain and binds to the DNA-binding domain. The combination of the direct connection between the dimerization domains and the DNA-binding domains and the noncovalent arm-mediated connection holds the DNA-binding domains in a relative orientation that favours DNA looping between the O_2 and I_1 half-sites (Seabold & Schleif, 1998). In order for AraC to bind to the adjacent I_1 and I_2 half-sites, AraC would have to be significantly distorted or at least one of its two arms would have to be removed from a DNA-binding domain. Hence the arm–DNA-binding domain interactions make it energetically more favourable for the protein to engage in DNA-looping interactions between the distally located half-sites rather than utilize the adjacent half-sites. When arabinose is present, however, the arms find it energetically more favourable to bind over the arabinose that is bound to the dimerization domains than to bind to the DNA-binding domains. Thus the arms shift position from the DNA-binding domains to the dimerization domains. As a result of the loss of the interactions with the arms, the DNA-binding domains are no longer constrained in their relative orientations. As a consequence of this, the AraC dimer now finds it energetically more favourable to bind to the adjacent half-sites I_1 and I_2 than to bind to I_1 and O_2.

How does the arm determine that arabinose is present on the dimerization domain if it is bound to the DNA-binding domain, and how is the DNA-binding domain at O_2 directed to the I_2 half-site? These two questions are generated due to the static nature of the representation in Fig. 4. In reality, part of the time the arm is bound to the DNA-binding domain, but some of the time it must be bound to the dimerization domain and part of the time it must be free in solution. The fact that its lowest energy state occurs when it is bound to the DNA-binding domain merely means that it spends most of the time at that position. It is in equilibrium with the other states, and likely shifts from one to another on a millisecond to microsecond timescale. Hence, despite the fact that most of the time the arm is bound to the DNA-binding domain, the arm samples arabinose occupancy of the dimerization domain, and can almost instantaneously respond to the presence of arabinose. This same diffusional motion is also the reason that the DNA-binding domain that occupies the O_2 site in the absence of arabinose does not have to be explicitly directed to the I_2 half-site when it is freed from the constraint provided by

the arm. Very quickly, probably in the order of milliseconds, the DNA-binding domain can shift from O_2 to I_2.

What binds the arm to the dimerization domain in the presence of arabinose? The binding of arabinose to the dimerization domain could bind the arm to the dimerization domain by either of two mechanisms. Arabinose could change the structure of the dimerization domain and the arm could bind to this altered structure. Alternatively, direct interactions between the arm and arabinose could hold the arm in place. The structures of the dimerization domain that were determined by X-ray crystallography of crystals grown in both the presence and absence of arabinose provide some information on this question, but less than might have been hoped (Soisson *et al.*, 1997). While there is virtually no change in the structure of the dimerization domain in the presence and absence of arabinose, when arabinose is absent, pairs of subunits engage in a French kiss, with each subunit inserting a tyrosine from near the base of the arm into the arabinose-binding pocket of its partner. Conceivably, the tyrosine also generates the same conformational change that is induced by arabinose. More likely, however, is the possibility that arabinose (and tyrosine) does not induce a significant change in the structure of the dimerization. If this is the case, the main determinant of arm position would then be direct arm–arabinose interactions. The interaction energies of such interactions calculated from molecular dynamics simulations of the dimerization domain are consistent with this notion.

EVIDENCE FOR THE LIGHT SWITCH MECHANISM

A number of lines of evidence are consistent with the light switch mechanism for AraC action. The simplest is the effect of deleting the N-terminal arm from AraC. The light switch model predicts that in the absence of the arm, the DNA-binding domains of AraC will behave the same as they do when arabinose is present; that is, armless AraC should make the arabinose operon constitutive. Indeed, that is what was found (Saviola *et al.*, 1998). Deleting the first five amino acids from the arm did nothing, but deleting more residues made the protein induce p_{BAD} in the absence of arabinose. It was concluded that the first four or five amino acids of AraC do not make important contacts with the DNA-binding domain. These same residues were also not structured in the crystals of the dimerization domain, and hence they appear not to play an important role in the function of AraC.

Several lines of evidence whose common theme was rigidity support the light switch model. The simplest was the effect of reversing the orientation of the O_2 half-site. When this was done, taking care to retain its binding face on the same side of the DNA, AraC could no longer form a DNA loop between I_1 and the reversed O_2 half-site (Seabold &

Schleif, 1998). This indicated that the AraC protein itself is rigid in the absence of arabinose and cannot adapt to the reoriented O_2 half-site.

In vivo, AraC was seen to shift from forming a DNA loop to binding to adjacent half-sites when arabinose was added (Martin *et al.*, 1986). *In vitro*, the consequence of adding arabinose was that AraC bound about 30 times more tightly to adjacent half-sites (Hendrickson & Schleif, 1984). This difference reflects the energetic costs of bending AraC with the arms binding to the DNA-binding domain or of releasing at least one of the arms from the DNA-binding domain. If DNA were provided whose half-sites could be freely positioned and orientated just the way AraC would prefer them, both in the presence and absence of arabinose, this energetic difference should be eliminated. Consequently, AraC would bind to such DNA with the same affinity in the presence and absence of arabinose. Such DNA was constructed by connecting double-stranded regions constituting the half-sites with single-stranded, and hence highly flexible, DNA. The DNA for this experiment was constructed by hybridizing two 25 base oligonucleotides to each end of a 75 base oligonucleotide, leaving a long stretch of single-stranded DNA in the middle connecting two double-stranded I_1 half-sites. AraC dissociated from this DNA at a rate independent of the presence of arabinose (Fig. 5) (Harmer *et al.*, 2001). In a parallel experiment, AraC dissociated nearly 30 times faster from normal double-stranded DNA in the absence of arabinose than in the presence of arabinose. Not only was this experimental result consistent with the light switch mechanism, it ruled out the possibility that the arm changes the intrinsic affinity of the DNA-binding domains for DNA or for special sequences of DNA.

According to the light switch mechanism, the two DNA-binding domains of AraC are comparatively free to move and orient themselves in the presence of arabinose. Hence if the two DNA-binding domains of AraC were connected not by dimerization domains, but merely by a peptide linker so that the single polypeptide chain included two DNA-binding domains and a peptide linker, the resulting protein ought to behave like AraC in the presence of arabinose. Indeed, the double DNA-binding domain protein acted *in vivo* to induce p_{BAD} and did not detectably loop between I_1 and O_2 (Harmer *et al.*, 2001). Another way to look at this result is that a signal must be sent from the dimerization domain of AraC to the DNA-binding domains to inform the DNA-binding domains of the presence or absence of arabinose. Formally, one signal could be sent in the presence of arabinose, and a different signal sent in the absence of arabinose. More likely, however, is the possibility that no signal is sent in one situation, and signal is sent in the other situation. The results with double DNA-binding domain protein showed that the no signal state corresponds to the plus arabinose situation.

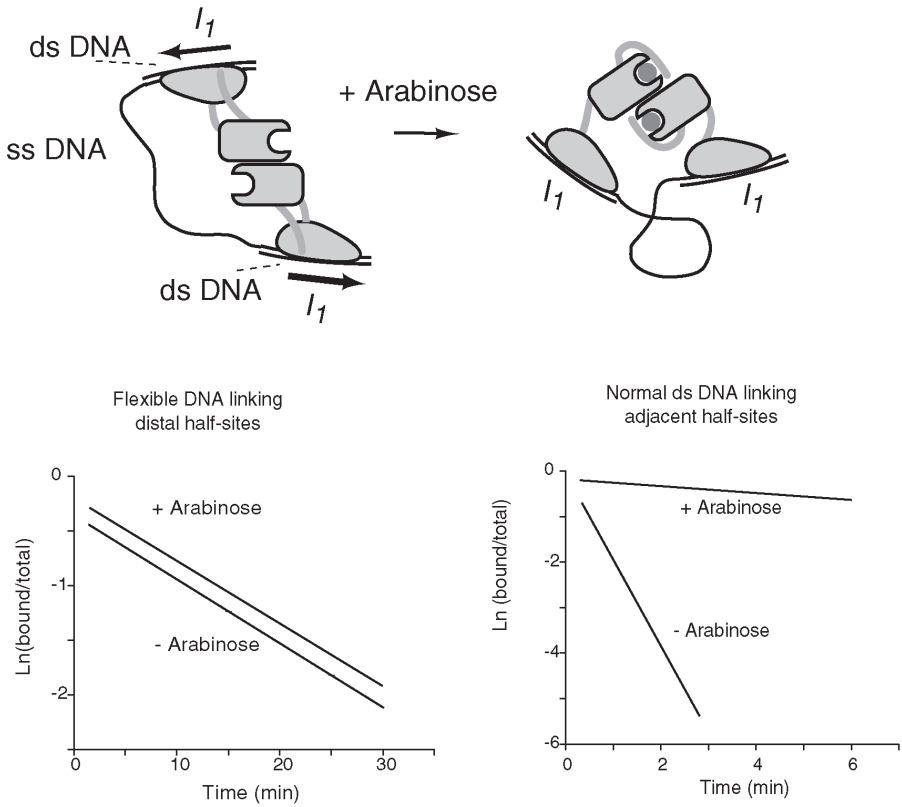

Fig. 5. Dissociation rate of AraC from DNA consisting of flexibly connected DNA binding half-sites is independent of arabinose. The dissociation rate of AraC from normal double-stranded DNA shows a 30-fold arabinose dependence.

Although they were developed for the study of the mechanism of the ligand response of AraC, the two methods of eliminating rigidity in the DNA or in the protein are generally applicable to the study of DNA-binding proteins that respond to ligands. They can determine whether a ligand modulates a protein's DNA-binding affinity by changing the intrinsic affinity of a DNA-binding domain for DNA or whether the ligand changes the relative positioning of two DNA-binding domains with respect to each other. By changing the positioning of DNA subunits, DNA-binding affinity is altered by changing the DNA-binding cooperativity between the two domains. This results from the fact that when one domain binds, the other domain may then be properly positioned for binding itself (positive cooperativity) or improperly positioned (negative cooperativity). Elimination of a ligand-binding domain allows identification of a 'no signal' state of the protein.

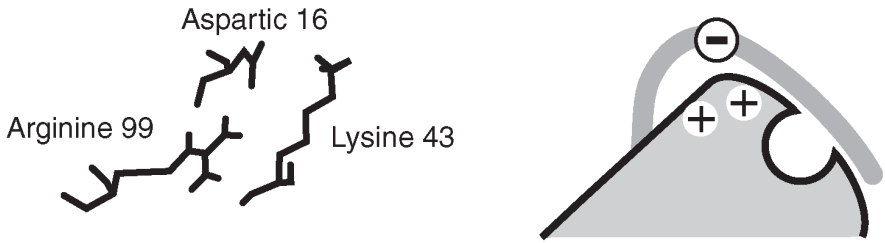

Fig. 6. Schematic drawing of the locations of the aspartic acid residue at position 16 in the N-terminal arm and the charged arginine and lysine residues that lie nearby when the arm binds over the arabinose-binding pocket in the dimerization domain.

The light switch mechanism predicted the existence of a number of different classes of mutations that affect the regulatory properties of AraC. Since the N-terminal arm was postulated to bind to two different sites, it was possible to imagine mutations that either strengthened or weakened its binding to either of these two sites. Furthermore, the mutations could be imagined to lie in either the binding sites on the domains, or in the arm itself. Hence there are eight classes of mutations that might be found that would directly affect arm positioning and which should have easily observable effects on regulation by AraC. Several of these expected classes have been identified (Saviola *et al.*, 1998; Reed & Schleif, 1999).

A mutation in the dimerization domain that binds the N-terminal arm more tightly would tend to make AraC protein constitutive; that is, not to require arabinose for it to be in the state where it binds *cis* and induces transcription from p_{BAD}. The mutation in the arm from asparagine at position 16 to aspartic acid appeared to be such a mutation (Wu & Schleif, 2001). The change made AraC constitutive. Structurally, it was easy to see how the change would lead to the arm binding more tightly to the dimerization domain. In the plus arabinose state, asparagine 16 lies very near the positively charged lysine 43 and arginine 99 residues (Fig. 6). Hence the arm likely is held more tightly to the dimerization domain by electrostatic interactions. Consistent with this idea is the fact that changing any one of the three charged amino acids to uncharged alanine significantly reduced the constitutivity and moved the behaviour of the protein back towards wild-type in that it requires arabinose for induction.

Mutations with a similar effect of binding the arm more tightly have also been identified on the DNA-binding domain (Saviola *et al.*, 1998). A number of these mutations left AraC unable to respond to arabinose, presumably because the arm cannot be pulled off the DNA-binding domain. Many of these introduced a positively charged amino acid, implying that a new electrostatic interaction was created. The arm contains two negatively charged amino acids, aspartic acid at positions 3 and 7. Changing aspartate

7 to alanine relieved the effect of introducing a positive change on the DNA-binding domain. Thus this second site suppressor indicated the existence of a direct interaction between the N-terminal arm and the DNA-binding domain.

The two classes of mutations discussed above both strengthened arm–domain interactions. Mutations that apparently weaken the interactions have also been identified. An example would be a mutation in the DNA-binding domain that leads AraC not to repress by looping DNA, but instead to induce to some degree. Such mutations have also been found (Wu, 2000). Once their existence was shown, an alanine and glutamate scan of the surface of the DNA-binding domain was performed. The scan identified ten different residues whose side chains were required for the full repressive abilities of AraC protein. These residues lie in a path across the side of the domain opposite to the DNA-binding side and suggest that the arm binds along the path.

Surface residues of the DNA-binding domain could be targeted for the scan because two homologues of the DNA-binding domain of AraC, MarA and Rob have been crystallized in the presence of their DNA-binding sites and their structures determined (Rhee *et al.*, 1998; Kwon *et al.*, 2000). The amino acid sequence of AraC could be threaded through these homologous structures to provide an approximate structure of the DNA-binding domain of AraC.

Alanine and glutamic acid were chosen as residues for the scan for two reasons. Alanine possesses a minimal side chain and its introduction rarely disrupts proteins. It might be thought that substitution of alanine for a residue whose side chain is involved in an interaction would weaken the stability of a structure; this is often not the case. The increased flexibility resulting from the absence of a side chain leads to a phenomenon called entropy–enthalpy compensation (Dunitz, 1995), and the resulting free energy differences can be surprisingly small. As a result, residues that are directly involved in interactions may not be identified in alanine scans. The substitution of the large and charged glutamic acid for such a residue almost surely will interfere with an interaction, however. The combination of alanine and glutamate is advantageous because in oligonucleotide-directed mutagenesis, a mixed base oligonucleotide may be used that will introduce alanine or glutamic acid, but no other residues. Then, following mutagenesis, the absence of transformants with altered phenotype means that neither alanine nor glutamic acid at that position generates a change. Any candidates with altered phenotype can be sequenced to determine whether the alteration is due to alanine or glutamic acid.

Finally, direct physical evidence for the light switch mechanism was obtained by observing binding between a peptide with a sequence of the N-terminal arm, and a dimerization

domain that was deleted of the arm. These experiments were carried out with a Biacore machine, which uses plasmon resonance to measure very small concentration changes within about 100 Å of a surface. Peptide with a sequence of the N-terminal arm of AraC was immobilized in the surface, and solution containing the arm-deleted dimerization domain of AraC was flowed past (Ghosh & Schleif, 2001). Binding of the dimerization domain to the immobilized arm was seen only when arabinose was also present. No binding was seen when glucose was substituted for arabinose, and intermediate binding was seen when fucose, an analogue of arabinose, was provided. Similarly, using an arm peptide containing the asparagine to aspartate change at position 16 showed tighter binding between the peptide and the domain. A mutation, phenylalanine 15 to leucine, is uninducible *in vivo* and likely is a result of weaker arm–arabinose binding. Dimerization domain did not detectably bind to this peptide.

SUMMARY

In summary, two important and general biochemical mechanisms have been discovered in the studies of the arabinose system. These are the phenomenon of DNA looping, and the light switch mechanism for regulation of protein activity. The consequences of DNA looping have been extensively discussed (Schleif, 1988, 1992) and will not be further addressed here. More should be said about the light switch mechanism, however. The mechanism is a special case of arm–domain interactions. While such interactions in proteins have not been given much attention, they are being found in an increasing number of systems. Such interactions provide a simple way for evolution to add interactions with one protein to another protein or to a whole class of proteins (Schleif, 1999). An arm sequence that binds to one protein might be added to many proteins, thus enabling each of them to bind to the first protein. Several examples for this phenomenon are that the proteins in the prokaryotic and in the eukaryotic DNA replication apparatus bind to one another via arm–domain interactions. Cell sorting and transport proteins also use this interaction mechanism.

Arm–domain interactions are also attractive for the engineering of new protein–protein interactions. While it is difficult to imagine engineering the surfaces of two proteins so that they will now bind to one another, it is simple to imagine finding a peptide that will bind to one of the protein's surfaces. This peptide could then be added to the other protein. Such peptides may prove to be readily identified through the use of phage display libraries. It is even possible to imagine designing an allosteric regulation mechanism based on the light switch mechanism. If part of the peptide arm of AraC activated or inhibited an enzyme to which it bound, then the addition of arabinose would change the enzyme's activity by pulling the arm off the enzyme and onto the dimerization domain of AraC.

REFERENCES

Dandanell, G. & Hammer, K. (1985). Two operator sites separated by 599 base pairs are required for *deoR* repression of the *deo* operon of *Escherichia coli*. *EMBO J* **4**, 3333–3338.

Dunitz, J. D. (1995). Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chem Biol* **2**, 709–712.

Dunn, T., Hahn, S., Ogden, S. & Schleif, R. (1984). An operator at −280 base pairs that is required for repression of *araBAD* operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. *Proc Natl Acad Sci U S A* **81**, 5017–5020.

Eismann, E., von Wilcken-Bergmann, B. & Müller-Hill, B. (1987). Specific destruction of the second *lac* operator decreases repression of the *lac* operon in *Escherichia coli* fivefold. *J Mol Biol* **195**, 949–952.

Englesberg, E., Irr, J., Power, J. & Lee, N. (1965). Positive control of enzyme synthesis by gene C in the L-arabinose system. *J Bacteriol* **90**, 946–957.

Englesberg, E., Squires, C. & Meronk, F. (1969a). 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* **62**, 1100–1107.

Englesberg, E., Sheppard, D., Squires, C. & Meronk, F., Jr (1969b). An analysis of "revertants" of a deletion mutant in the C gene of the L-arabinose gene complex in *Escherichia coli* B/r: isolation of initiator constitutive mutants (Ic). *J Mol Biol* **43**, 281–298.

Ghosh, M. & Schleif, R. (2001). Biophysical evidence of arm-domain interactions in AraC. *Anal Biochem* **295**, 107–112.

Gielow, L., Largen, M. & Englesberg, E. (1971). Initiator constitutive mutants of the L-arabinose operon (OIBAD) of *Escherichia coli* B/r. *Genetics* **69**, 289–302.

Greenblatt, J. & Schleif, R. (1971). Regulation of the arabinose operon *in vitro*. *Nat New Biol* **233**, 166–170.

Gross, J. & Englesberg, E. (1959). Determination of the order of mutational sites governing L-arabinose utilization in *Escherichia coli* B/r by transduction with phage P1bt. *Virology* **9**, 314–331.

Harmer, T., Wu, M. & Schleif, R. (2001). The role of rigidity in DNA looping-unlooping by AraC. *Proc Natl Acad Sci U S A* **98**, 427–431.

Hendrickson, W. & Schleif, R. (1984). Regulation of the *Escherichia coli* L-arabinose operon studied by gel electrophoresis DNA binding assay. *J Mol Biol* **178**, 611–628.

Kwon, H. J., Bennik, M. H., Demple, B. & Ellenberger, T. (2000). Crystal structure of the *Escherichia coli* Rob transcription factor in complex with DNA. *Nat Struct Biol* **7**, 424–430.

Lee, D. & Schleif, R. (1989). *In vitro* DNA loops in *araCBAD*: size limits and helical repeat. *Proc Natl Acad Sci U S A* **86**, 476–480.

Lobell, R. & Schleif, R. (1990). DNA looping and unlooping by AraC protein. *Science* **250**, 528–532.

Martin, K., Huo, L. & Schleif, R. (1986). The DNA loop model for *ara* repression: AraC protein occupies the proposed loop sites in vivo and repression-negative mutations lie in these same sites. *Proc Natl Acad Sci U S A* **83**, 3654–3658.

Reed, W. & Schleif, R. (1999). Hemiplegic mutations in AraC protein. *J Mol Biol* **294**, 417–425.

Rhee, S., Martin, R. G., Rosner, J. L. & Davies, D. R. (1998). A novel DNA-binding motif in

MarA: the first structure for an AraC family transcriptional activator. *Proc Natl Acad Sci U S A* **95**, 10413–10418.

Saviola, B., Seabold, R. & Schleif, R. (1998). Arm-domain interactions in AraC. *J Mol Biol* **278**, 539–548.

Schleif, R. (1972). Fine-structure deletion map of the *Escherichia coli* L-arabinose operon. *Proc Natl Acad Sci U S A* **69**, 3479–3484.

Schleif, R. (1988). DNA looping. *Science* **240**, 127–128.

Schleif, R. (1992). DNA looping. *Annu Rev Biochem* **61**, 199–223.

Schleif, R. (1996). Two positively regulated systems, *ara* and *mal*. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, pp. 1300–1309. Edited by F. C. Neidhardt, R. Curtiss, III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter & H. E. Umbarger. Washington, DC: American Society for Microbiology.

Schleif, R. (1999). Arm-domain interactions in proteins: a review. *Proteins* **34**, 1–3.

Schleif, R. & Lis, J. T. (1975). The regulatory region of the L-arabinose operon: a physical, genetic and physiological study. *J Mol Biol* **95**, 417–431.

Seabold, R. & Schleif, R. (1998). Apo-AraC actively seeks to loop. *J Mol Biol* **278**, 529–538.

Sheppard, D. & Englesberg, E. (1967). Further evidence for positive control of the L-arabinose system by gene *araC*. *J Mol Biol* **25**, 443–454.

Soisson, S., MacDougall-Shackleton, B., Schleif, R. & Wolberger, C. (1997). Structural basis for ligand-regulated oligomerization of AraC. *Science* **276**, 421–425.

Wu, M. (2000). *Arm domain interactions in AraC*. PhD thesis, Johns Hopkins University.

Wu, M. & Schleif, R. (2001). Strengthened arm-dimerization domain interactions in AraC. *J Biol Chem* **276**, 2562–2564.

Zhang, X. & Schleif, R. (1998). Catabolite gene activator protein mutations affecting activity of the *araBAD* promoter. *J Bacteriol* **180**, 195–200.

Zhang, X., Reeder, T. & Schleif, R. (1996). Transcription activation parameters at *ara* p_{BAD}. *J Mol Biol* **258**, 14–24.