Probing key DNA contacts in AraR-mediated transcriptional repression of the *Bacillus subtilis* arabinose regulon

Irina Saraiva Franco¹, Luís Jaime Mota¹, Cláudio Manuel Soares² and Isabel de Sá-Nogueira¹,3,*

¹Laboratory of Microbial Genetics and ²Laboratory of Protein Modeling, Instituto de Tecnologia Química e Biomédica, Universidade Nova de Lisboa. Av. da República, Apt. 127, 2781-901 Oeiras and ³Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Quinta da Torre, 2829-516 Caparica, Portugal

Received May 11, 2007; Revised June 7, 2007; Accepted June 12, 2007

ABSTRACT

In the absence of arabinose, the AraR transcription factor represses the expression of genes involved in the utilization of arabinose, xylose and galactose in *Bacillus subtilis*. AraR exhibits a chimeric organization: the N-terminal DNA-binding region belongs to the GntR family and the C-terminal effector-binding domain is homologous to the GalR/LacI family. Here, the AraR–DNA-binding interactions were characterized in vivo and in vitro. The effect of residue substitutions in the AraR N-terminal domain and of base-pair exchanges into an AraR–DNA-binding operator site were examined by assaying for AraR-mediated regulatory activity in vivo and DNA-binding activity in vitro. The results showed that residues K4, R45 and Q61, located in or near the winged-helix DNA-binding motif, were the most critical amino acids required for AraR function. In addition, the analysis of the various mutations in an AraR palindromic operator sequence indicated that bases G₉, A₁₁ and T₁₆ are crucial for AraR binding. Moreover, an AraR mutant M34T was isolated that partially suppressed the effect of mutations in the regulatory cis-elements. Together, these findings extend the knowledge on the nature of AraR nucleoprotein complexes and provide insight into the mechanism that underlies the mode of action of AraR and its orthologues.

INTRODUCTION

The transcription factor AraR controls the utilization of carbohydrates in *Bacillus subtilis*. The control exerted by AraR is modulated by the presence of the effector molecule arabinose leading to induction of expression of at least 13 genes, comprising the arabinose (ara) regulon, which includes the *araR* gene (1–4). The products of these genes (ara*ABDLNPQ-abfA, araE, abnA and xsa) include extracellular and intracellular catabolic enzymes involved in the degradation of arabinose, galactose and xylose containing polysaccharides, uptake of these sugars into the cell and further catabolism of l-arabinose and arabinose oligomers (1–3,5).

A key property of AraR is its ability to bind specific DNA sequences in the absence of the inducer l-arabinose, as determined by DNAse I footprinting analysis (4,6,7). AraR recognizes and binds at least eight palindromic operator sequences, located in the five known arabinose-inducible promoters. Three of these promoters contain two *ara* boxes: the promoter of the *ara*ABDLNPQ-abfA operon (boxes OR₁ and OR₂), of *araE* (OR₃ and OR₄) and of *xsa* (OR₅ and OR₆). In the cases of the genes *araR* and abnA, a single *ara* box is present (OR₇ and OR₈). AraR binding to the promoters displaying two *ara* boxes is cooperative, requiring in phase and properly spaced operators, and involves the formation of a small loop in the DNA. These two mechanistically diverse modes of action of AraR result in distinct levels of transcriptional regulation, as cooperative binding to two *ara* boxes results in a high level of repression while interaction with a single operator allows a more flexible control (4,6,7).

AraR is a 362 amino acid homodimeric protein that shows a chimeric organization, consisting of two functional domains with different phylogenetic origins (1,6,8): a small N-terminal DNA-binding domain (DBD) comprising a winged helix–turn–helix (HTH) motif belonging to the GntR family of transcriptional regulators (9)
and a larger C-terminal domain homologous to that of the GalR/LacI family of bacterial regulators and sugar-binding proteins (10). AraR typifies one of the six GntR-subfamilies of proteins (11,12). Currently, there are 54 members of this rapidly growing class of proteins, which can be found in prokaryotes [CDART database; (13)].

Previously, a model for AraR was derived using comparative modelling based on crystal structures of FadR (DBD) and PurR (COOH domain) from Escherichia coli (8). We have used random and site-directed mutagenesis to map the functional domains of AraR required for DNA binding, dimerization and effector binding. The arabinose-binding pocket is composed of polar and charged residues, whereas the dimerization interface has a hydrophobic nature. In both cases, the residues are distributed along the primary sequence of the C-terminal domain (8). Based on crystallographic studies of structurally and functionally related proteins, binding of the effector to the COOH region in AraR is predicted to elicit a conformational change in the N-terminal region, leading to inhibition of binding to operator sequences, and allowing transcription from the arabinose-responsive promoters. This allosteric signal involves a switching mechanism for communicating structural changes triggered in the sensor domain to the regulatory domain, decreasing the affinity of the latter for DNA.

Winged helix motifs are functionally and mechanistically versatile (14). They are primarily involved in DNA binding, but cases have been reported in which they participate in protein–protein interactions. Monomeric, homo- or heterodimeric protein–DNA complexes have been characterized and revealed quite distinct modes of binding to DNA, which can involve interactions between the recognition helix and the wing with the major and minor groove (15). Although the level of amino acid identity for the DBD of all members of the GntR superfamily is low (~25%) they share this conserved structural topology (11). Global analysis of the conservation of amino acid sequences in DNA-binding proteins concluded that residues interacting with the DNA backbone establish a set of core contacts that provide stability for homologous protein–DNA complexes, and consequently are well conserved across all protein families. On the other hand, residues that interact with DNA bases have more variable levels of conservation (16). Previous mutagenic studies showed that AraR residues in the N-terminal region were required for DNA binding because mutations in these residues abolished its regulatory function in vivo (8). However, the precise contribution of the mutated amino acids to DNA-binding activity was unclear.

To understand the specific properties of the interaction AraR-operator sequences, we substituted amino acids, in or near the HTH motif, which according to the model were predicted to contact DNA. We determined the effects of these substitutions on the ability of AraR to function in vivo and on the DNA-binding affinities in vitro. Conversely, mutational analysis of the AraR-binding sites was used to determine the base-specific requirements for transcriptional regulation in vivo and DNA binding in vitro. These experiments gave both expected and unexpected results, which together showed that specific AraR residues and operator bases are crucial to achieve a high level of regulatory activity, while others display variable contributions to DNA binding. In addition, an AraR mutant was isolated, which partially suppresses the loss of regulation observed in certain mutated DNA operators.

**MATERIALS AND METHODS**

**Strains and growth conditions**

Bacillus subtilis strains used in this work (Table 1) were grown in Luria–Bertani (LB) medium (17) or C minimal medium (18) and solid sugar-free agar (SFA) medium (LabM) or LB solidified with 1.6% agar. Chloramphenicol (5 μg ml⁻¹), kanamycin (10 μg ml⁻¹) and erythromycin (1 μg ml⁻¹) were added when appropriate. The Amy phenotype was tested by detection of starch hydrolysis on tryptose blood agar base medium (Difco) plates, containing 1% of potato starch, with a β–Kl solution as described previously (3). Escherichia coli strains were grown on LB medium, with ampicillin (100 μg ml⁻¹), chloramphenicol (20 μg ml⁻¹), kanamycin (30 μg ml⁻¹) and IPTG (isopropyl-β-D-thiogalactopyranoside) (1 mM) added as appropriate. The B. subtilis and E. coli cells were transformed as described previously (7).

**DNA manipulations and sequencing**

DNA manipulations were carried out as described previously (20). Restriction enzymes were purchased from MBI Fermentas, New England Biolabs or Roche, and used according to manufacturer’s instructions. DNA was eluted from agarose gels using the GeneCleanII kit (Bio101) or the GFX DNA purification kit (GE Healthcare). PCRs were performed in a GeneAmp PCR system 2400 (Perkin-Elmer) and PCR products purified using QIAquick PCR purification kit (QIAGen). DNA was sequenced using an ABI PRIS BigDye terminator ready reaction cycle sequencing kit (Applied Biosystems).

**Site-directed mutagenesis of araR**

Amino acid substitutions in AraR were made by the QuikChange (Stratagene) site-directed method using as template plasmid pLS30 (8) and mutagenic oligonucleotides carrying the modified codon in the centre (listed in Supplementary Table 1). For R41A and H42A, a 486-bp template plasmid pLS30 (8) and mutagenic oligonucleotides carrying the modified codon in the centre (listed in Supplementary Table 1). For R41A and H42A, a 486-bp template plasmid pLS30 (8) and mutagenic oligonucleotides carrying the modified codon in the centre (listed in Supplementary Table 1). For R41A and H42A, a 486-bp template plasmid pLS30 (8) and mutagenic oligonucleotides carrying the modified codon in the centre (listed in Supplementary Table 1). For R41A and H42A, a 486-bp template plasmid pLS30 (8) and mutagenic oligonucleotides carrying the modified codon in the centre (listed in Supplementary Table 1). For R41A and H42A, a 486-bp template plasmid pLS30 (8) and mutagenic oligonucleotides carrying the modified codon in the centre (listed in Supplementary Table 1). For R41A and H42A, a 486-bp template plasmid pLS30 (8) and mutagenic oligonucleotides carrying the modified codon in the centre (listed in Supplementary Table 1).
### Table 1. *Bacillus subtilis* strains used in this work

| Strain   | Genotype                  | Relevant phenotype  | Source                  |
|----------|---------------------------|---------------------|-------------------------|
| IQB101   | araAB-lacZ erm            | Ara" LacZ"          | (1)                     |
| IQB351   | ΔaraR::km araAB-lacZ erm  | Ara" LacZ"          | pLM8→IQB101<sup>a,b</sup> |
| IQB352   | ΔaraR::km araAB-lacZ erm  | LacZ"                | pLS24→IQB350           |
| IQB353   | ΔaraR::km araAB-lacZ erm  | LacZ"                | pLS30→IQB350           |
| IQB355   | ΔaraR::km araAB-lacZ erm  | ΔamyE::araR cat     | pFL1→IQB350            |
| IQB356   | ΔaraR::km araAB-lacZ erm  | ΔamyE::araR Q73S cat| pFL2→IQB350            |
| IQB357   | ΔaraR::km araAB-lacZ erm  | ΔamyE::araR L335 cat| pFL3→IQB350            |
| IQB358   | ΔaraR::km araAB-lacZ erm  | ΔamyE::araR Δ131-65 cat| pFL8→IQB350         |
| IQB505   | ΔaraR::km araAB-lacZ erm  | ΔamyE::araR H42A cat| pFL4→IQB350            |
| IQB513   | ΔaraR::km araAB-lacZ erm  | ΔamyE::araR R41A cat| LacZ"                  |
| IQB563   | ΔaraR::km araAB-lacZ erm  | ΔamyE::araR R454A cat| LacZ"                  |
| IQB564   | ΔaraR::km araAB-lacZ erm  | ΔamyE::araR Q61A cat| LacZ"                  |
| IQB568   | ΔaraR::km araAB-lacZ erm  | ΔamyE::araR E30A cat| LacZ"                  |
| IQB571   | ΔaraR::km araAB-lacZ erm  | ΔamyE::araR Y5F cat  | LacZ<sup>b</sup>       |
| IQB712   | ΔaraR::km araAB-lacZ erm  | ΔamyE::araR K4A cat  | LacZ<sup>b</sup>       |
| IQB530   | ΔaraR::kmΔamyE::ORA1 (G9→T) lacZ cat | LacZ<sup>b</sup> | pLM32→IQB215<sup>6</sup> |
| IQB531   | ΔaraR::kmΔamyE::ORA1 (A11→T) lacZ cat | LacZ<sup>b</sup> | pLM56→IQB215<sup>6</sup> |
| IQB532   | ΔaraR::kmΔamyE::ORA1 (T16→G) lacZ cat | LacZ<sup>b</sup> | pLM65→IQB215<sup>6</sup> |
| IQB533   | ΔaraR::kmΔamyE::ORA1 (T9→G) lacZ cat | LacZ<sup>b</sup> | pLM66→IQB215<sup>6</sup> |
| IQB534   | ΔaraR::kmΔamyE::ORA1 (T10→G) lacZ cat | LacZ<sup>b</sup> | pLM67→IQB215<sup>6</sup> |
| IQB535   | ΔaraR::kmΔamyE::ORA1 (T16→G) lacZ cat | LacZ<sup>b</sup> | pLM68→IQB215<sup>6</sup> |
| IQB536   | ΔaraR::kmΔamyE::ORA1 (A11→C) lacZ cat | LacZ<sup>b</sup> | pLM69→IQB215<sup>6</sup> |
| IQB537   | ΔaraR::kmΔamyE::ORA1 (G9→T) lacZ cat | LacZ<sup>b</sup> | pLM78→IQB215<sup>6</sup> |
| IQB538   | ΔaraR::kmΔamyE::ORA1 (G9→T) lacZ cat | LacZ<sup>b</sup> | pLM70→IQB215<sup>6</sup> |
| IQB539   | ΔaraR::kmΔamyE::ORA1 (C8→A) lacZ cat | LacZ<sup>b</sup> | pL76→IQB530<sup>6</sup> |
| IQB540   | ΔaraR::kmΔamyE::ORA1 (T10→G) lacZ cat | LacZ<sup>b</sup> | pL76→IQB531<sup>6</sup> |
| IQB541   | ΔaraR::kmΔamyE::ORA1 (A11→C) lacZ cat | LacZ<sup>b</sup> | pL76→IQB532<sup>6</sup> |
| IQB542   | ΔaraR::kmΔamyE::ORA1 (T9→G) lacZ cat | LacZ<sup>b</sup> | pL76→IQB533<sup>6</sup> |
| IQB543   | ΔaraR::kmΔamyE::ORA1 (T10→G) lacZ cat | LacZ<sup>b</sup> | pL76→IQB534<sup>6</sup> |
| IQB544   | ΔaraR::kmΔamyE::ORA1 (G9→T) lacZ cat | LacZ<sup>b</sup> | pL76→IQB535<sup>6</sup> |
| IQB545   | ΔaraR::kmΔamyE::ORA1 (G9→T) lacZ cat | LacZ<sup>b</sup> | pL76→IQB536<sup>6</sup> |
| IQB546   | ΔaraR::kmΔamyE::ORA1 (G9→T) lacZ cat | LacZ<sup>b</sup> | pL76→IQB537<sup>6</sup> |
| IQB547   | ΔaraR::kmΔamyE::ORA1 (G9→T) lacZ cat | LacZ<sup>b</sup> | pL76→IQB538<sup>6</sup> |

*The arrows indicate transformation and point from donor DNA to recipient strain.
Transformation was always carried out with linearized DNA.
Mutagenized pLS30 DNA was used as donor DNA (see Materials and Methods section).

---

Mutagenesis of pLS30, the amplified products were ligated, linearized and used to transform *B. subtilis* (Table 1). Substitution R451L was obtained by chance when attempting to create mutation R45A (8) using an identical procedure. Presence of the mutations was verified by sequencing the araR allele in the resulting plasmids or strains.

**Site-directed mutagenesis of operator regions**

Plasmid pLM51 (6) a pBluescript II KS (+) derivative carrying the wild-type *araABDLMPQ-apaBA* promoter, was used as template for generating single-nucleotide substitutions in OR<sub>A1</sub> or OR<sub>A2</sub>, using the QuikChange (Stratagene) site-directed method and pairs of mutagenic oligonucleotides (listed in Supplementary Table 1).

Resulting plasmids contained the following mutations in OR<sub>A1</sub>: A<sub>1</sub>→C (pLM61), G<sub>3</sub>→T (pLM62), T<sub>9</sub>→G (pLM58), C<sub>3</sub>→A (pLM63), G<sub>9</sub>→T (pLM54; (6)), T<sub>10</sub>→G (pLM59), A<sub>11</sub>→C (pLM57), T<sub>16</sub>→G (pLM60); or G<sub>9</sub>→T in OR<sub>A2</sub> (pLM77; (7)). The 204-bp BamHI–EcoRI DNA fragment from these plasmids, containing the mutagenized operator region, was then subcloned in the same sites of pSN32 (6). This procedure generated respectively pLM68, pLM69, pLM70, pLM56 (6), pLM66, pLM64, pLM67 and pLM78 (7), which bear transcriptional fusions of the *araABDLMPQ-apaBA* promoter with single-point mutations to *lacZ*. After linearization, plasmids were used to transform *B. subtilis* IQB215, giving rise to strains where these fusions were integrated at the *amyE* locus (Table 1). To analyse the
repression exerted by AraR on the ara boxes, integration of the araR allele at the thrC locus was accomplished by transformation with pIF76.

Isolation of AraR suppressor mutants

The insertion of a 1446-bp EcoRI–BamHI fragment from pLS30 (8), containing the araR allele, into pDG1664 (EcoRI–BamHI) (21) yielded plasmid pIF76. This plasmid was used as DNA template in random PCR mutagenesis, according to the method described by Leung et al. (22). Random PCR mutagenesis with oligos ARA6 and ARA73 amplified a 650-bp 5′-end region of the araR allele. After digestion with BamHI–Eco47III, the fragment was sub-cloned in the same plasmid leading to the replacement of the equivalent region of the wild-type araR allele, from sites −227 to +251 relative to the transcription start site (containing the promoter and the first 75 codons of the araR gene). The recombinant plasmids were transformed into E. coli DH5α yielding a library of araR mutations (contained in ~2200 transformants). A plasmid pool was used in separate experiments to transform B. subtilis strains with a araR::amyE::OR_A-lacZ background, in which the OR_A operator sequence carried the mutations described above (Table 1), leading to integration of AraR mutants at the thrC locus via double crossing-over. The constitutive expression of lacZ due to the presence of the mutated ara boxes leads to a Lac + phenotype in the receptor strains, reflected by a blue color in SFA medium with X-gal. To isolate mutant araR alleles suppressing the deleterious effect of the operator mutations, we screened for colonies displaying a weaker Lac phenotype (white/dim blue/colourless). Aplasticity of colonies was confirmed by a quantitative measure of AraR repression in the analysed strains (Represion Index).

β-Galactosidase assays

Bacillus subtilis strains were grown in C minimal medium supplemented with 1% (w/v) casein hydrolysate in the presence and in the absence of l-arabinose 0.4% (w/v) as previously reported (1). Samples of cell culture were collected and analysed 2 h after the addition of l-arabinose. The ratio of β-galactosidase activity, determined as described (17) from cultures grown for 2 h in the presence and absence of inducer was taken as a measure of AraR repression in the analysed strains (Repression Index).

Immunoblotting of cell extracts

Bacillus subtilis strains were grown as for β-galactosidase assays. Preparation of cell extracts and immunoblotting were performed as described (8). Blots were developed with anti-AraR-MBP2+ serum (6) using the ECL detection system (Amersham Biosciences). Protein concentration was determined using a Bio-Rad kit.

Construction of plasmids for overexpression of AraR mutants and protein purification

Fusion of the C-terminus of AraR variants to six histidines in the plasmid pET30a(+) (Novagen) was engineered, placing the genes under the control of a T7 promoter. The construction of plasmid pLS16, carrying the wild-type allele, was described previously (7). Construction of plasmids carrying the AraR substitutions F37S, Q61R and L33S was accomplished in a similar manner. Briefly, the alleles containing these mutations were amplified with oligos ARA50 and ARA51 from the pLS30-derivatives pIF1, pIF2 and pIF3 obtained previously (8). The 1112-bp PCR product was separately digested with Aval–NdeI and Aval–HindIII, the resulting 282-bp and a 805-bp fragments were inserted in pET30a(+) restricted with NdeI and HindIII, yielding plasmids pIF5, pIF6 and pIF7, respectively. To introduce mutations S53P, H42A or M34T, regions BglII–XhoI from plasmids pIF7 were obtained from plasmids pIF17, pIF41 and pIF85 described above, and used to substitute the same region in pIF7 generating pIF111, pIF123 and pIF121. For K4A, Y5F, E30A, R45A and Q61A B. subtilis chromosomal DNA from strains IQB712, IQB571, IQB568, IQB563 and IQB564 was used (Table 1). PCR products were digested with appropriate enzymes (NdeI–KpnI, BglII–KpnI or HindIII–HindIII) and used to substitute the corresponding region in pIF7. These procedures yielded pIF124, pIF112, pIF78, pIF74 and pIF75, respectively. The presence of the mutations was verified by sequencing the araR alleles. For the purification of these AraR-his6 variants, E. coli BL21 (DE3) pLysS (19) cells transformed with the corresponding pET30 derivatives were grown at 37°C to an optical density at 600 nm of 0.6 in 11 of LB medium, and then expression of the fusion proteins was induced by addition of IPTG to 1 mM. Incubation in the same conditions continued for additional 2 h. All subsequent steps were carried out similarly to the method described previously (7).

Electrophoretic mobility shift assays (EMSA)

A DNA fragment carrying the OR_A1–OR_A2 region was amplified from pLM51 using primers ARA262 and ARA263 (Supplementary Table 1). After purification, the 126-bp PCR product was labelled with T4 Polynucleotide Kinase (MBI Fermentas) and [γ-32P]dATP, followed by extraction with phenol/chloroform and precipitation with ethanol. Binding reactions contained 12 mM HEPES-KOH pH 7.6, 10 mM MgCl2, 0.5% (w/v) BSA, 1 mM DTE, 10% (v/v) Glycerol, 200 mM NaCl, 4 mM Na2HPO4, 4 mM NaH2PO4, 0.4 mM EDTA, a 200-fold molar excess of competitor DNA (polydlIdC), 1 nM of labelled DNA and increasing concentrations of wild-type or mutant AraR proteins. After incubation for 30 min at room temperature, the mixture was loaded onto a pre-run 8% polyacrylamide gel in 25 mM Tris 200 mM Glycine (pH 8.9) and run at 100 V for ~1 h. Gels were dried under vacuum and exposed to a Phosphorimager screen before analysis with a Molecular Dynamics Storm 860 Imager and ImageQuant version 5.0. To determine the dissociation constants, protein concentrations were used
according to previous in vitro results (6) and \( K_d \) values were obtained using the GraphPad Prism software. For competition DNA-binding experiments, various amounts of cold double-stranded oligonucleotides containing single mutations in the operator sequences (Supplementary Table 1) were added to the reaction in the presence of 40 nM AraR. As controls, we used oligonucleotides carrying the wild-type operator (ARA288 and ARA289) or a non-specific DNA sequence with the same length (ARA244 and ARA245). The following procedures were made as described above. The percentage inhibition in the presence of competitor DNA was determined similarly to the method described by Bera et al. (23). The radioactivity of bound DNA was quantified in the control without competitor, and in samples containing 500-fold molar excess of the distinct competitors. Inhibition (\%) = 100 \times \left( \frac{\text{[bound]control} - \text{[bound]sample}}{\text{[bound]control}} \right).

**RESULTS AND DISCUSSION**

In vivo effect of amino acid substitutions in the DBD of AraR

We have previously established that AraR interaction with DNA is achieved via a 70 amino acid N-terminal domain. These results were obtained by random and site-directed mutagenesis based on a 3D model of the DBD derived from the crystal structure of the E. coli regulator FadR (8). However, many of these mutations resulted in changes that could alter the protein structure or interfere with DNA binding. In order to conduct a more clarifying characterization of the role of specific AraR residues on its DNA-binding activity, we made several amino acid substitutions in or near the HTH motif. Residues were chosen based on the 3D model and/or primary sequence alignment of AraR-like proteins (8). The majority of the substituted amino acids was predicted to contact directly the bases of the DNA and consequently, would account for the specificity of the interaction with operator sequences. Positions K4, E30, R41, H42 and Q61 were substituted amino acids was predicted to contact directly the DNA backbone and in dark blue the ones contacting the DNA bases, through site-directed mutagenesis and green through random mutagenesis (8). The DNA-binding domain of AraR and localization of mutations. (A) Sequence alignment of the N-terminal region of AraR, FadR and GntR. Residues that are conserved in the entire family GntR are shaded in black and residues characteristic of AraR homologous proteins in grey (8). Positions of mutations leading to a constitutive phenotype are boxed in yellow (8,26,29). A substitution yielding a suppressor phenotype is boxed in red. The introduced residues in AraR are shown above the sequence, coloured in orange when obtained through site-directed mutagenesis and green through random mutagenesis (8). In FadR are coloured in light blue residues contacting the DNA backbone and in dark blue the ones contacting the DNA bases, according to crystallographic data (28); an asterisk below the sequence indicates amino acids within contact distance of DNA (24). The secondary structure (arrows representing beta-strands and bars alphahelices) of FadR (amino acid residues 1–73) is shown below the alignment according to van Aalten et al. (24). The microorganisms of source and accession numbers are: AraR from B. subtilis (P09671); GntR from B. subtilis (P10585); FadR from E. coli (P09371). (B) Structure of the modelled N-terminal domain of AraR (depicted in green ribbons; only one monomer is represented for clarity; see (8) for details) together with the DNA segment (depicted in orange ribbons) crystallized with FadR, highlighting the site-directed (left) and random (right) mutations displaying a constitutive phenotype. Site-directed mutations: K4→A, Y5→F, E30→A, H42→A, R45→A and Q61→A. Random mutations: L33→S, F37→S, S53→P and P61→R.

![Figure 1](image)

Figure 1. The DNA-binding domain of AraR and localization of mutations. (A) Sequence alignment of the N-terminal region of AraR, FadR and GntR. Residues that are conserved in the entire family GntR are shaded in black and residues characteristic of AraR homologous proteins in grey (8). Positions of mutations leading to a constitutive phenotype are boxed in yellow (8,26,29). A substitution yielding a suppressor phenotype is boxed in red. The introduced residues in AraR are shown above the sequence, coloured in orange when obtained through site-directed mutagenesis and green through random mutagenesis (8). In FadR are coloured in light blue residues contacting the DNA backbone and in dark blue the ones contacting the DNA bases, according to crystallographic data (28); an asterisk below the sequence indicates amino acids within contact distance of DNA (24). The secondary structure (arrows representing beta-strands and bars alphahelices) of FadR (amino acid residues 1–73) is shown below the alignment according to van Aalten et al. (24). The microorganisms of source and accession numbers are: AraR from B. subtilis (P09671); GntR from B. subtilis (P10585); FadR from E. coli (P09371). (B) Structure of the modelled N-terminal domain of AraR (depicted in green ribbons; only one monomer is represented for clarity; see (8) for details) together with the DNA segment (depicted in orange ribbons) crystallized with FadR, highlighting the site-directed (left) and random (right) mutations displaying a constitutive phenotype. Site-directed mutations: K4→A, Y5→F, E30→A, H42→A, R45→A and Q61→A. Random mutations: L33→S, F37→S, S53→P and P61→R.
Analysis of the wild-type AraR and mutant proteins
DNA-binding affinity in vitro

The apparent affinity constants ($K_d$) of AraR mutants for operator sequences were determined by EMSAs using a $^{32}$P-labelled 126-bp DNA fragment, which carried both operators of the metabolic operon, OR$_{A1}$ and OR$_{A2}$ (depicted in Figure 5). Binding of AraR to this DNA fragment was specific, as the presence of the inducer arabinose but not xylose prevented the formation of the protein–DNA complex (Figure 3A). Titration of 1 nM of DNA with increasing concentrations of wild-type repressor (Figure 3B) allowed the determination of an apparent $K_d$ $3.9 \times 10^{-8}$ M, which is defined as the amount of protein necessary to shift 50% of the labelled probe (25). This value is comparable to that previously calculated for each individual box using DNase I quantitative footprinting experiments, $3.4 \times 10^{-8}$ M and $4.7 \times 10^{-8}$ M for OR$_{A1}$ and OR$_{A2}$, respectively (6).

All the mutant proteins that displayed an effect in vivo were overexpressed in E. coli and purified to homogeneity (see Materials and Methods section). Binding to DNA was assayed by EMSA and their respective apparent affinity constants determined (Figure 3C). Variant H42A, which showed the minimal loss of repression in vivo (2-fold) bound DNA with an apparent $K_d$ of $4.1 \times 10^{-8}$ M, similar to the wild-type protein. The most severe effects were displayed by Q61A ($K_d \sim 5.5 \times 10^{-7}$ M), K4A and R45A, both showing an apparent $K_d > 1.5 \times 10^{-6}$ M. These three mutants were also unable to perform a regulatory activity in vivo. Residues K4 and H42 are completely conserved among AraRs (8) and in a contact distance of the DNA according to the model (Figure 1A and B), however, mutation of these residues had different outcomes. Since these two residues are not conserved among the members of the entire GntR family they may contribute in very different extents to the DNA-binding specificity of AraR-like proteins.

R45 is a conserved amino acid in the GntR family members. In the regulator FadR from E. coli, the...
substitution of the corresponding residue (R49) has also a
drastic effect in vivo (26). Moreover, the crystal structure
of the FadR–DNA complex (27,28) shows that R49
locates in the recognition helix of the winged HTH and
interacts with a phosphate group, not specifically a base.
According to the predictions of the tertiary structure of
AraR, R45 is also located in the recognition helix
(Figure 1B), which is generally more responsible for the
interaction with DNA, in particular the positively charged
residues. Q61 belongs to the predicted wing of the DNA-
binding motif (Figure 1A and B). The corresponding
residue in both FadR and in GntR from B. subtilis
is also positively charged (Figure 1A) and substitutions
led to loss of DNA-binding ability (26,29). In FadR, H65

Figure 3. Binding of AraR to araABDLMNPO-abfA promoter (operators OR1-OR3) in EMSA. AraR was incubated with the 5'-end-labelled
probe (1-2 nM) and the protein–DNA complexes resolved by electrophoresis on native 8% polyacrylamide gels. Protein concentrations were
calculated considering a pure dimeric protein. (A) Specificity controls for AraR binding. AraR (60 nM) was incubated with the DNA probe in the
presence of L-arabinose or D-xylose (15 mM). (B) The indicated amounts of wild-type AraR were used in the binding reactions (left). Densitometric
quantification of the bands corresponding to free DNA and protein–DNA complex allowed the determination of the affinity constant (right)
(see Materials and Methods section). The values shown represent the average and standard deviation of at least three independent assays, with an
intrinsic error <30%. (C) Effect of AraR substitutions on binding to DNA probe. The indicated concentrations of the mutant proteins were used in
EMSA and determination of $K_d$ was made as described above.
is part of the wing and makes specific contacts with an adenine (24,28).

Intermediate decreases of DNA binding were observed with Y5F and E30A, with $K_d$ $2.0 \times 10^{-7} \text{M}$ and $2.3 \times 10^{-7} \text{M}$, respectively, similarly to that seen in vitro (8). Therefore, these exchanges led to a comparable effect both in vivo and in vitro (Figures 2 and 3). Moreover, the nature of the mutation Y5F revealed the importance of the OH group in the interaction with DNA. Both residues are conserved in the GntR-family proteins, and the corresponding residues in FadR, A9 and E34, were shown to contact the DNA backbone (24,28). The latter also contacts nearby amino acids, contributing presumably to the stabilization of residues that interact specifically with the DNA bases.

Additionally, four other AraR mutants, L33S, F37S, S53P and Q61R, obtained by random mutagenesis and characterized in vivo in a previous work (8); Figure 1B), were studied by EMSA. Both substitutions F37S and S53P had led to derepression in vivo, of ~24- and 2-fold, and in mutant L33S the regulatory activity was almost completely abolished (8). These values could be explained by the observed instability of the proteins (8). However, purified mutants F37S and S53P showed decreased DNA-binding affinities, $K_d$ 7.4 $\times 10^{-7} \text{M}$ and 6.5 $\times 10^{-7} \text{M}$, respectively, that may also contribute to the deregulation in vivo. These results could be explained by the nature and localization of these substitutions, which suggest implications in the folding of the DBD. Overexpression and purification of L33S yielded only small amounts of protein. Nevertheless, at the maximal concentration that we could use in EMSA assays, 50nM of the mutant, no DNA binding was observed (data not shown).

Interestingly, while the Q61A substitution completely abolished regulation in vivo and DNA binding in vitro, the change to arginine in the same position showed only a 1.6-fold decrease of regulatory activity in vivo (8), and the affinity to the DNA probe in vitro, $K_d$ 3.1 $\times 10^{-7} \text{M}$, was even slightly higher than that displayed by wild-type AraR. Noteworthy, in GntR the inverse of AraR mutation Q61R (i.e. GntR R75Q) leads to a significant loss of regulation in vivo (29). Based on these observations, we may speculate that the rise of positive charge as a result of AraR substitution Q61R increased the overall (non-specific) affinity for DNA, leading in vivo to a titration of the protein.

In summary, K4, R45 and Q61, were the most critical AraR residues in achieving specific DNA binding, and Y5 and E30 also play an important role, and overall there was a good correlation between the effects of the mutations in the binding affinities to the ara operon promoter in vitro and in the regulatory activity in vivo.

**Effect of base-pairs mutations in the operator sequences on transcriptional regulation by AraR**

AraR recognizes and binds at least eight palindromic operator sequences, located in the five known arabinose-inducible promoters. Three of these promoters contain two ara boxes: the promoter of the ara metabolic operon (boxes OR$_{A1}$ and OR$_{A2}$), of araE (OR$_{E1}$ and OR$_{E2}$) and of xsa (OR$_{X1}$ and OR$_{X2}$). In the cases of the genes araR and $abrA$, a single ara box is present (OR$_{RA3}$ and OR$_{RA4}$). AraR binding to the promoters displaying two ara boxes is cooperative and involves the formation of a small loop in the DNA. In fact, for full in vivo repression, communication between repressor molecules bound to two properly spaced operators is required, as shown by the analysis of mutations designed to prevent cooperative binding of AraR (6,7). An alignment of the eight ara boxes, identified by DNase I footprinting and/or mutagenesis, showed the 16-bp consensus sequence 5’-TTTGTACGTACAAAT-3’ and highlighted the conserved nucleotides at each position (Figure 4A). This operator consensus presents the typical signature for cis-acting elements recognized by GntR family members 5’-(N)x-GT-N(0-15)-A-C-(N)x-3’ (11).

In a previous work, we showed that G$_9$ is important for AraR binding because the substitution G$_9$→T in both boxes OR$_{A1}$ and OR$_{A2}$ caused defect in the regulatory activity of AraR in vivo and prevented cooperative binding (6). To further investigate which nucleotides within the consensus sequence were necessary for protein binding, single-nucleotide exchanges were made in OR$_{A1}$ at the promoter of the ara operon. The most conserved bases were substituted and mutations were designed to introduce transversions from AT to CG and CG to AT: A$_1$→C, G$_5$→T, T$_6$→G, C$_8$→A, T$_{10}$→G, A$_{11}$→C and T$_{16}$→G. The mutated promoters transcriptional fused to the lacZ gene were independently integrated at the amyE locus of the B. subtilis receptor strain (see Materials and Methods section and Table 1).

Strain IQB572, bearing a transcriptional fusion to the wild-type operator ($\Delta$araR $\Delta$amyE::OR$_{A(wt)}$-lacZ thrC::araR) was used as a control to assay the repression exerted by AraR. The levels of accumulated $\beta$-galactosidase activity measured in all strains are shown in Figure 4B. Mutations having the most drastic effect on AraR binding were G$_9$→T, both in OR$_{A1}$ and OR$_{A2}$ [as previously determined; (7)], A$_{11}$→C and T$_{16}$→G, leading to a decrease in the regulatory activity >9-fold compared to the control. A moderate effect of deregulation, varying from 2.4- to 4.4-fold, was observed for A$_1$→C, G$_5$→T and T$_{6}$→G, and substitution T$_{10}$→G had no effect in vivo.

Surprisingly, C$_8$→A abolished lacZ expression both in inducing and non-inducing conditions. One possibility to explain this result would be an increase in the affinity of AraR for the mutated operator leading to a tight binding of the repressor (even in the presence of arabinose), thus preventing transcription by RNA polymerase. To test this hypothesis, we investigated the Lac phenotype in araR-null mutants ($\Delta$araR) bearing the transcriptional fusion OR$_{A1}$/(C$_8$→A)$_{3}$-lacZ and OR$_{A(wt)}$-lacZ, strains IQB257 and IQB530 (Table 1), respectively. The results obtained in solid medium with X-gal and arabinose indicated that the lack of lacZ expression in mutant C$_8$→A (Position +4, relative to the transcriptional start site) was independent of the presence of AraR (data not shown) suggesting that the mutation may affect transcription initiation by RNA polymerase.
Figure 4. Ara operator sequences. (A) Alignment of the eight AraR boxes and a picture representing the conservation of bases at each position in the inferred consensus 16-bp palindromic operator sequence [generated by WebLogo 2.8.2 software (33)]. AraR boxes are located in the promoter of the araABDLMNQP-abfA metabolic operon (ORA1 and OR A2), araE (ORA3 and OR X2), araR (OR R3), xsa (OR X1 and OR X2) and abnA (OR R4). The conserved nucleotides in at least seven boxes are shaded. Bases substituted in OR A1 or OR A2 by site-directed mutagenesis used for araB analysis are indicated, and the new base is shown above. (B) β-Galactosidase activities were determined in B. subtilis strains carrying OR 5-lacZ fusions integrated at the amyE locus and wild-type araR at the thrC locus grown in the absence or presence of arabinose (upper panel, in white and black bars, respectively). The repression index (lower panel) reflects the regulation exerted on each ara box (Figure 2). Nucleotide substitutions (obtained by site-directed mutagenesis) are indicated for the mutated position in the ara consensus box (Figure 4A). Values are the average of three independent experiments, each assayed in duplicate. Error bars represent the SD. M.U.—Miller Units.

Analysis of operator mutations on AraR–DNA affinity in vitro

The effect of the operator mutations was also analysed in vitro by EMSA competition assays. The experiments were performed in the presence of 1 nM of the OR A1– OR A2 DNA probe described above, 40 nM of AraR, and increasing concentrations of a double-stranded 38-bp competitor oligonucleotide (50–500 nM) containing the wild-type or the mutated sequences in OR A1 or OR A2. In addition, oligonucleotides carrying all possible substitutions at the highly conserved base pairs, G9, A11 and T16 were also used. We compared the ability of these cold DNAs to titrate binding of AraR to the labelled probe, reflected in the decrease of the intensity of the protein–DNA complex band. The wild-type OR A1 box was able to compete for AraR binding in a concentration-dependent manner, with a 79% loss of band shift at 500-fold excess competitor DNA (Figure 5A). In contrast, a non-specific oligonucleotide (equivalent in length; Supplementary Table 1) used as control disrupted only 18% of the binding (data not shown). Inhibition of binding in the presence of 500 nM of the different competitors was quantified and the results are summarized in Figure 5B. The DNA containing mutation T16→G in OR A1 competed in levels similar to that obtained for the wild-type box (68 and 79% inhibition, respectively). In contrast, AraR was unable to bind the boxes with single base-pair substitutions in G9, either to T (the mutation tested in vivo, previously), to A and C (inhibition values between 21 and 16%). A notorious decrease in binding to A11→C was also observed, which was more pronounced when A was exchanged for G or T. However, oligonucleotides containing a mutation at T16 (either to G, A or C) were still able to partially compete for the repressor. The three mutations, A1→C, G5→T and T6→G, leading to a partial de-repression in vitro also showed an intermediate effect. Taken together, the results indicate a good correlation between the in vivo and in vitro, but the exchanges at T16 and the mutation A11→C, comparatively to the regulatory activity in vivo were expected to bind to AraR less tightly. This could be due to the more reduced sensitivity of the competition assays. Interestingly, the mutation C8→A, that affected transcription even in the presence of inducer, did not compete (23%) for AraR, indicating that the mutation has an effect on AraR binding that could not be tested in vivo.

In previous work, a search for AraR operator sequences in the B. subtilis genome (6,30) identified a putative binding sequence in the open reading frame ydfK [identified as a myo-inositol transporter, ioIT; (31)]. The sequence 5′-TTTTACGTACATTT-3′ [+27 relative to the translation start site; (31)] displayed only two deviations (underlined) from the consensus sequence: A1 and G5. Construction and analysis of transcriptional fusions of the promoter region and 5′-end of ydfK to lacZ showed that expression is not AraR dependent, thus the potential operator is not functional in vivo (Inácio J.M. and I.S.-N., unpublished data). To determine the ability of AraR to bind this sequence in vitro, competition assays were performed as described above but no competition was detected (Figure 5B). Since T at Position 1 is present in functionally active AraR boxes (Figure 4A) these observations suggest, in accordance to the mutagenic
analysis described above, that G at Position 5 plays an important role in AraR binding.

In conclusion, we found that bases in both arms of the palindrome of the AraR boxes are involved in AraR–DNA contacts. The in vivo and in vitro studies together with sequence analysis of the eight functionally active AraR B. subtilis boxes indicated that bases G₉, A₁₁ and T₁₆ are crucial for AraR binding, and that A₁ and T₆ play also an important role. Furthermore, Position 5 required a purine (Pu) for functionality in vivo, and sequence analysis suggested that the corresponding mirror base (Position 12) in the other arm of the palindrome is always a pyrimidine (Py). An alignment of all putative AraR-binding sequences based on a search of the consensus 5₀-ATTTGTACGTACAAAT-3₀ in genomes of bacteria from the Bacillus/Clostridium group that contain AraR orthologues also highlighted the majority of the invariable positions: Pu₅, T₆ and the correspondent mirror A₁₁, Py₁₂ and G₉ at the centre of the palindrome (Supplementary Figure 1).

Mutation M34T partially restores AraR binding to mutated DNA operators both in vivo and in vitro

In order to isolate AraR mutants that could suppress the loss of regulation caused by the single nucleotide substitutions in the ara boxes, an in vivo screening method was developed (see Materials and Methods section). Briefly, random mutagenesis of the 5₀-end of the araR allele was performed by PCR and the resulting library of plasmids carrying the mutated alleles was used to transform B. subtilis strains, allowing its integration at the non-essential thrC locus via a double-crossover event. The receptor strains possessed a /C₁ araR/C₁ amyE::ORA- lacZ background, and carried different mutations in the ORA operator sequence (Table 1). The constitutive expression of lacZ due to the inability of the wild-type AraR to bind the mutated operator leads to a Lac⁺ phenotype in the absence of inducer. However, if the integrated mutant araR allele encodes a protein that suppressed the deleterious effect of the operator mutation...

Figure 5. In vitro analysis of AraR binding to mutated ara boxes. (A) Competition EMSA experiments using double-stranded oligonucleotides containing mutated ara boxes. AraR (40 nM) was incubated with the ³²P-labelled ORA₁–ORA₂ region (1 nM) in the presence or absence of the indicated molar excess of 38-bp ds oligonucleotide competitors with wild type or the mutated ara boxes shown above. Protein–DNA complexes were resolved by electrophoresis in 8% polyacrylamide gels. Representative results are shown. (B) Quantification of the inhibition of AraR binding to wild-type operator sequence in the presence of competitor DNA. The values represent the percentage of inhibition of AraR binding to the labelled DNA probe observed in the presence of 500-fold molar excess of competitor DNA. For quantification, the intensity of the bands corresponding to protein-DNA complexes in EMSA, obtained in the presence or absence of competitor, were quantified in a densitometer. The percentage of inhibition was calculated as described in the Materials and Methods section. The results are the average and SD of at least three independent experiments, with a maximal associated error of 30%.
a Lac− phenotype is displayed indicating recovery of regulation. Thus, we screened for transformants with decreased β-galactosidase production in the absence of arabinose.

One transformant of strain IQB533, containing the fusion ORA1(T6→G)–lacZ, displayed a gain-of-function phenotype, and the sequencing of the ara operon promoter revealed the substitution M34T, located in the HTH motif of the protein. To determine if this effect was specific or also affected AraR binding to the other mutated promoters, the allele was integrated at the thrC locus of the corresponding B. subtilis strains and β-galactosidase activities were measured (Figure 6A). Interestingly, the repression exerted by mutant M34T was higher than that exerted by wild-type AraR in almost all operators, in particular with mutant boxes G9→T (both in ORA1 and ORA2), T16→G, A11→C and T16→G. Since a higher level of intracellular accumulation of this AraR variant could explain this phenotype, we determined the accumulation of the wild-type AraR and mutant M34T in strains IQB572 and IQB583, respectively (Table 1). The observed cellular levels of protein were similar in both strains (Figure 6B), indicating that the phenotype displayed by M34T was not due to increased concentration of protein. In fact, EMSA assays performed as described above showed that the mutant displays an increased affinity to the ara operon promoter probe, with an apparent Kd of 1.0 × 10⁻⁸ M (Figure 6C), which is almost 4-fold lower than that of the wild-type protein (Kd = 3.9 × 10⁻⁸ M). The substitution M34T is located in the first helix of the winged HTH motif (Figure 1A). According to a study on protein–DNA interactions based on structures of 129 complexes (32), threonine is responsible for a far larger number of protein–DNA bonds than methionine, although almost all are made with the DNA backbone and not the bases. This in agreement with the results we obtained, that show an increase in the repression exerted over all mutated boxes, suggesting an increased DNA affinity of M34T through non-specific contacts.

CONCLUDING REMARKS

Previous studies have mapped the functional domains of AraR and characterized the C-terminal region involved in effector binding and dimerization (8). In this work, we focused on two additional and crucial components of the transcription process, the DBD and cis-acting elements. Guided by molecular modelling combined with multiple primary sequence alignment of AraR orthologues and GntR family members, we identified amino acids potentially involved in DNA binding. The effect of their substitution was analysed in vivo and in vitro and revealed key residues necessary for the regulatory activity. In addition, important bases for AraR–DNA interactions in both arms of the palindromic operator sequences were also identified. We obtained both expected and unexpected results highlighting the uniqueness of protein–DNA interactions in each particular system. A future determination of the structure of AraR, in its unbound form or in complex with the inducer or DNA, would allow a more detailed analysis of the mechanism by which AraR binds its cognate operator sequences and how the conformational change triggered by the binding of arabinose prevents this interaction.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This work was partially supported by grants POCTI/ BME/36164/00 and POCI/BIA-MIC/61140/04 from Fundação para a Ciência e Tecnologia (FCT) and FEDER to I.S.-N. I.S.F. is the holder of Ph.D. fellowship SFRH/BDL/61140/04 from FCT. Funding to pay the Open Access publication charges for this article was provided by Fundação para a Ciência e Tecnologia.

Conflict of interest statement. None declared.
REFERENCES

1. Sá-Nogueira, I., Nogueira, T.V., Soares, S. and de Lencastre, H. (1997) The Bacillus subtilis L-arabinose (ara) operon: nucleotide sequence, genetic organization and expression. Microbiology, 143 (Pt 3), 957–969.

2. Sá-Nogueira, I. and Ramos, S.S. (1997) Cloning, functional analysis, and transcriptional regulation of the Bacillus subtilis araE gene involved in L-arabinose utilization. J. Bacteriol., 179, 7705–7711.

3. Sá-Nogueira, I. and Mota, L.J. (1997) Negative regulation of L-arabinose metabolism in Bacillus subtilis: characterization of the araR (araC) gene. J. Bacteriol., 179, 1598–1608.

4. Raposo, M.P., Inacio, J.M., Mota, L.J. and de Sá-Nogueira, I. (2004) Transcriptional regulation of genes encoding arabino-degrading enzymes in Bacillus subtilis. J. Bacteriol., 186, 1287–1296.

5. Krispin, O. and Allmansberger, R. (1998) The Bacillus subtilis AraE protein displays a broad substrate specificity for several different sugars. J. Bacteriol., 180, 3250–3252.

6. Mota, L.J., Tavares, P. and Sá-Nogueira, I. (1999) Mode of action of AraR, the key regulator of L-arabinose metabolism in Bacillus subtilis. Mol. Microbiol., 33, 476–489.

7. Mota, L.J., Sarmento, L.M. and de Sá-Nogueira, I. (2001) Control of the arabinose regulon in Bacillus subtilis by AraR in vivo: crucial roles of operators, cooperativity, and DNA looping. J. Bacteriol., 183, 4190–4201.

8. Franco, I.S., Mota, L.J., Soares, C.M. and de Sá-Nogueira, I. (2006) Functional domains of the Bacillus subtilis transcription factor AraR and identification of amino acids important for nucleoprotein complex assembly and effector binding. J. Bacteriol., 188, 3024–3036.

9. Haydon, D.J. and Guest, J.R. (1991) A new family of bacterial regulatory proteins. FEMS Microbiol. Lett., 63, 291–295.

10. Weickert, M.J. and Adhya, S. (1992) A family of bacterial regulators homologous to Gal and Lac repressors. J. Biol. Chem., 267, 15869–15874.

11. Rigali, S., Derouaux, A., Giannotta, F. and Dusart, J. (2002) Subdivision of the helix-turn-helix GntR family of bacterial regulators in the FadR, HutC, MoeR and YtrA subfamilies. J. Biol. Chem., 277, 12507–12515.

12. Lee, M.H., Scherer, M., Rigali, S. and Golden, J.W. (2003) PlmA, a new member of the GntR family, has plasmid maintenance functions in Anaerbaena sp. strain PCC 7120. J. Bacteriol., 185, 4315–4325.

13. Geer, L.Y., Domrachev, M., Lipman, D.J. and Bryant, S.H. (2002) CDART: protein homology by domain architecture. Genome Res., 12, 1619–1623.

14. Clark, K.L., Halay, E.D., Lai, E. and Burley, S.K. (1993) Co-structural DNA-recognition motifs resemble histone H5. Nature, 364, 412–420.

15. Gajiwala, K.S. and Burley, S.K. (2000) Winged helix proteins. Curr. Opin. Struct. Biol., 10, 110–116.

16. Luscombe, N.M. and Thornton, J.M. (2002) Protein-DNA interactions: amino acid conservation and the effects of mutations on binding specificity. J. Mol. Biol., 320, 991–1009.

17. Miller, J. (1972) Experiments in Molecular Genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.

18. Pascal, M., Lai, S. and Dedonder, R. (1971) Characterization of two sucrase activities in Bacillus subtilis. Biochimie, 53, 1059–1066.

19. Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorf, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol., 185, 60–89.

20. Sambrook, J., Frisch, E. and Maniatis, T. (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York, USA.

21. Guerout-Fleury, A.M., Frandsen, N. and Stragier, P. (1996) Plasmids for ectopic integration in Bacillus subtilis. Gene, 180, 57–61.

22. Leung, D., Chen, E. and Goeddel, D.V. (1989) A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. Technique, 1, 11–15.

23. Bera, A.K., Zhu, J., Zalkin, H. and Smith, J.L. (2003) Functional dissection of the Bacillus subtilis pur operator site. J. Bacteriol., 185, 4099–4109.

24. van Aalten, D.M., DiRusso, C.C. and Knudson, J. (2001) The structural basis of acyl coenzyme A-dependent regulation of the transcription factor FadR. EMBO J., 20, 2041–2050.

25. Riggs, A.D., Suzuki, H. and Bourgeois, S. (1970) Lac repressor-operator interaction. I. Equilibrium studies. J. Mol. Biol., 48, 67–83.

26. Raman, N., Black, P.N. and DiRusso, C.C. (1997) Characterization of the fatty acid-responsive transcription factor FadR: biochemical and genetic analyses of the native conformation and functional domains. J. Biol. Chem., 272, 30645–30650.

27. van Aalten, D.M., DiRusso, C.C., Knudson, J. and Wierenga, R.K. (2000) Crystal structure of FadR, a fatty acid-responsive transcription factor with a novel acyl coenzyme A-binding fold. EMBO J., 19, 5167–5177.

28. Xu, Y., Heath, R.J., Li, Z., Rock, C.O. and White, S.W. (2001) The FadR–DNA complex. Transcriptional control of fatty acid metabolism in Escherichia coli. J. Biol. Chem., 276, 17373–17379.

29. Yoshida, K., Fujita, Y. and Sarai, A. (1993) Missense mutations in the Bacillus subtilis gnt repressor that diminish operator binding affinity. J. Mol. Biol., 231, 167–174.

30. Rodionov, D.A., Mironov, A.A. and Gelfand, M.S. (2001) Transcriptional regulation of pentose utilisation systems in the Bacillus/Clostridium group of bacteria. FEMS Microbiol. Lett., 205, 305–314.

31. Yoshida, K., Yamamoto, O., Omac, K., Yamamoto, M. and Fujita, Y. (2002) Identification of two myo-inositol transporter genes of Bacillus subtilis. J. Bacteriol., 184, 983–991.

32. Luscombe, N.M., Laskowski, R.A. and Thornton, J.M. (2001) Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. Nucleic Acids Res., 29, 2860–2874.

33. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. Genome Res., 14, 1188–1190.